A STRUCTURE PRESERVING SHIFT-INVERT INFINITE ARNOLDI ALGORITHM FOR A CLASS OF DELAY EIGENVALUE PROBLEMS WITH HAMILTONIAN SYMMETRY *

PIETER APPELTANS† AND WIM MICHIELS†

Abstract. This work considers a class of delay eigenvalue problems that admit a spectrum similar to that of a Hamiltonian matrix, in the sense that the spectrum is symmetric with respect to both the real and imaginary axis. We propose a method to iteratively approximate the eigenvalues closest to a given purely real or imaginary shift, while preserving the symmetries of the spectrum. To this end, our method exploits the equivalence between the considered delay eigenvalue problem and the eigenvalue problem associated with a linear but infinite-dimensional operator. To compute the eigenvalues closest to the given shift, we apply a specifically chosen shift-invert transformation to this linear operator and compute the eigenvalues with the largest modulus of the new shifted and inverted operator using an (infinite) Arnoldi procedure. The advantage of the chosen shift-invert transformation is that the spectrum of the transformed operator has a "real skew-Hamiltonian"-like structure. Furthermore, it is proven that the Krylov space constructed by applying this operator satisfies an orthogonality property in terms of a specifically chosen bilinear form. By taking this property into account during the orthogonalization process, it is ensured that, even in the presence of rounding errors, the obtained approximation for, e.g., a simple, purely imaginary eigenvalue is simple and purely imaginary. The presented work can thus be seen as an extension of [V. Mehrmann and D. Watkins, *Structure-Preserving Methods for Computing Eigenpairs of Large Sparse Skew-Hamiltonian/Hamiltonian Pencils*, SIAM J. Sci. Comput. (22.6), 2001], to the considered class of delay eigenvalue problems. Although the presented method is initially defined on function spaces, it can be implemented using finite-dimensional linear algebra operations. The performance of the resulting numerical algorithm is verified for two example problems: the first example illustrates the advantage of proposed approach in preserving purely imaginary eigenvalues when working in finite precision, while the second one demonstrates its applicability to a large scale problem.

Key words. nonlinear eigenvalue problems, delay eigenvalue problems, Hamiltonian eigenvalue problems, structure preserving method, infinite Arnoldi algorithm.

AMS subject classifications. 65H17, 34K06.

1. Introduction.

In this manuscript, we consider nonlinear eigenvalue problems (NEPs) of the form

$$M(\lambda)v = 0, \tag{1.1}$$

with $\lambda \in \mathbb{C}$ an eigenvalue and $v \in \mathbb{C}^{2n} \setminus \{0\}$ a right eigenvector, for which the characteristic matrix has the following structure:

$$M(\lambda) := \lambda I_{2n} - H_0 - \sum_{k=1}^{K} \left(H_{-k} e^{-\lambda \tau_k} + H_k e^{\lambda \tau_k} \right), \tag{1.2}$$

with $0 < \tau_1 < ... < \tau_K < \infty$ discrete delays and I_{2n} the identity matrix of size $2n$. The matrices $H_0, H_1, H_{-1}, \ldots, H_K$ and H_{-K} belong to $\mathbb{R}^{2n \times 2n}$ and satisfy the following assumption.

Assumption 1.1. *Firstly, the matrix H_0 is Hamiltonian meaning that*

$$\left(JH_0\right)^\top = JH_0 \tag{1.3}$$

*Submitted to the editors on November 3, 2022.

Funding: This work was supported by the project C14/17/072 of the KU Leuven Research Council and the project G092721N of the Research Foundation-Flanders (FWO - Vlaanderen).

†NUMA Section, Department of Computer Science, KU Leuven, Leuven (B-3001), Belgium (pieter.appeltans@kuleuven.be,wim.michiels@kuleuven.be).

in which the matrix J is defined as

$$J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Secondly, the matrices H_k and H_{-k} are related via

$$\left(JH_{-k}\right)^\top = JH_k \text{ for } k = 1, \ldots, K, \tag{1.4}$$

with J as defined above.

A motivation for considering such NEPs stems from a popular approach to compute the $\mathcal{H}_\infty$-norm of a time-delay system [6]. More specifically, consider the following state-space system with delays

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & A_0 x(t) + \sum_{k=1}^K A_k x(t-\tau_k) + Bw(t) \\ z(t) & = & Cx(t), \end{array} \right. \tag{1.5}$$

with $x(t) \in \mathbb{R}^n$ the state vector, $w(t) \in \mathbb{R}^m$ the performance input, $z(t) \in \mathbb{R}^p$ the performance output, $0 < \tau_1 < ... < \tau_K < \infty$ discrete delays, and A_0, ..., A_K, B and C real-valued matrices of appropriate dimensions. The corresponding transfer matrix,

$$T(s) := C\left(sI - A_0 - \sum_{k=1}^K A_k e^{-s\tau_k}\right)^{-1} B,$$

describes the system's input-output map in the frequency domain. If system (1.5) is exponentially stable, its $\mathcal{H}_\infty$-norm is given by

$$\|T(\cdot)\|_{\mathcal{H}_\infty} := \max_{\omega \in \mathbb{R}^+} \|T(\jmath\omega)\|_2,$$

with $\jmath$ the imaginary unit. The $\mathcal{H}_\infty$-norm is an important performance measure in the robust control framework as it can be used to quantify both the input-to-output noise suppression of the system as well as the distance to instability of the system [26, 11]. To compute the $\mathcal{H}_\infty$-norm of a time-delay system, the methods presented in [19, 6], which can be seen as an extension of the well-known Boyd-Balakrishnan-Bruinsma-Steinbuch algorithm [3, 4], use a level set approach. An important component of such level set algorithms is to check whether for a given $\gamma > 0$ the inequality $\|T(\cdot)\|_{\mathcal{H}_\infty} \geq \gamma$ holds. For systems of the form (1.5) the following equivalence from [19, Lemma 2.1] can be used: for $\omega \in \mathbb{R}$ the matrix $T(\jmath\omega)$ has a singular value equal to γ if and only if $\jmath\omega$ is an eigenvalue of the NEP associated with the following characteristic matrix,

$$\lambda \begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix} - \begin{bmatrix} A_0 & \gamma^{-1}BB^\top \\ -\gamma^{-1}C^\top C & -A_0^\top \end{bmatrix} - \sum_{k=1}^K \left(\begin{bmatrix} A_k & 0 \\ 0 & 0 \end{bmatrix} e^{-\lambda\tau_k} + \begin{bmatrix} 0 & 0 \\ 0 & -A_k^\top \end{bmatrix} e^{\lambda\tau_k} \right). \tag{1.6}$$

Notice that this characteristic matrix fits the structure given in (1.2). Verifying whether $\|T(\cdot)\|_{\mathcal{H}_\infty} \geq \gamma$ is thus equivalent with checking whether the NEP associated with (1.6) has purely imaginary eigenvalues. For this application, it is thus important to accurately compute the purely imaginary eigenvalues.

The spectrum, i.e., the set of eigenvalues, of the considered NEP has some interesting features. Firstly, notice that (1.1) bears some similarities with a retarded

delay eigenvalue problem (RDEP) [10, 5, 21], but in contrast has both positive and negative delays. It is therefore not surprising that (1.1) generally has infinitely many eigenvalues. Secondly, due to the structure of the characteristic matrix in (1.2) and Assumption 1.1, the considered class of NEPs can be seen as an extension of the linear eigenvalue problem

$$(I_{2n}\lambda - H)\, v = 0, \tag{1.7}$$

with $H \in \mathbb{R}^{2n\times 2n}$ a Hamiltonian matrix, as studied in, among others, [2, 24, 16]. It is well known that the spectrum of such eigenvalue problems is symmetric with respect to both the real and imaginary axis, meaning that eigenvalues appear either in quadruplets (λ, $\bar{\lambda}$, $-\lambda$ and $-\bar{\lambda}$) or in purely real or imaginary pairs (λ and $-\lambda$) [18]. In the next section we will see that the spectrum of the considered NEP possesses the same symmetries.

The goal of this paper is to develop a method to accurately compute the eigenvalues of (1.1) that lie close to a given shift σ while preserving the symmetry of the spectrum. We are particularly interested in the case for which the characteristic matrix (1.2) is large. For large-scale linear eigenvalue problems, Krylov subspace methods, such as the (shift-invert) Arnoldi method, are well established [22]. These methods have recently been generalized to NEPs, see for example [13, 12, 7, 25, 14]. In the case of RDEPs, the method in [13] can be interpreted in two ways: either as applying the Arnoldi method to a sufficiently large linearisation of the NEP with vectors of increasing length or as applying the Arnoldi method to an associated linear but infinite-dimensional operator, resulting in a Krylov subspace spanned by functions instead of vectors, hence its name, the infinite Arnoldi method. The main benefit of this method is that it is in some sense discretization-free as the error made by approximating the nonlinear delay term can be reduced by applying more Arnoldi iterations. Here we want to apply a similar approach. However, for eigenvalue problems with a Hamiltonian structure, as considered here, the regular shift-invert (infinite) Arnoldi method destroys the particular structure of the spectrum. For linear finite-dimensional eigenvalue problems, a modified shift-invert Arnoldi approach was therefore developed in [16], which preserves the Hamiltonian structure. This method was subsequently generalized to both matrix pencils and polynomial eigenvalue problems in [15] and [17], respectively. The goal of this paper is to extend this structure-preserving method to the considered class of NEP.

More specifically, the adopted approach can be interpreted as the algorithm of [16] applied to a linear (but infinite dimensional) operator whose spectrum corresponds to the one of (1.1). An alternative approach could consist of constructing a rational or polynomial approximation of the characteristic matrix (1.2), followed by a direct application of the algorithm of [17] (a so-called discretize-and-solve approach). As we are interested in computing selected eigenvalues of large-scale problems close to a shift σ (on the real or imaginary axis) this would require an approximation that preserves the Hamiltonian symmetry of the spectrum while being accurate around $\pm\sigma$, precluding an approximation around zero if $|\sigma| \gg 0$. As we shall see, the latter property is naturally embedded in the presented shift-invert transformation. Furthermore, when using the discretize-and-solve framework, it is not trivial to a priori determine the degree of the discretization for which the approximation error is sufficiently small (e.g., such that the imaginary axis eigenvalues are preserved for the $\mathcal{H}_\infty$-norm computation). It should be noted, though, that the discretize-and-solve framework does not exclude dynamic methods, for which the degree of the discretisation does not need to be specified beforehand and can change during the Arnoldi process (see, e.g., the

dynamic variants of NLEIGS [8]). These methods are however closely related to the adopted infinite Arnoldi approach.

The remainder of this work is structured as follows. Section 2 recalls some preliminary results: first some properties of the spectrum of (1.1) are highlighted, next the most important components of the structure preserving shift-invert Arnoldi method from [16] are reviewed and finally the equivalence of the considered NEP and the eigenvalue problems associated with two linear but infinite-dimensional operators (one related to the left eigenspace and one related to the right eigenspace) is demonstrated. In Section 3, the structure preserving shift-invert Arnoldi method from [16] is generalized to the considered NEP based on the aforementioned equivalence with a linear infinite-dimensional eigenvalue problem, giving rise to an algorithmic implementation of the structure preserving shift-invert *infinite* Arnoldi method. Subsequently Sections 4 and 5 discuss how this method can be implemented using finite-dimensional operations for the shift σ employed in the shift-invert Arnoldi method equal to zero and different from zero, respectively. Next, Section 6 applies the resulting numerical implementations on two example problems. The first example illustrates that the presented method indeed preserves the Hamiltonian structure of the spectrum, while the second example demonstrates its applicability to large-scale problems. Finally, Section 7 draws some concluding remarks.

2. Preliminary results.

In this section we will recall some important preliminary results.

2.1. Properties of the considered NEP.

Throughout the paper a nonzero column vector $w \in \mathbb{C}^{2n} \setminus \{0\}$ is a left eigenvector of the characteristic matrix (1.2) associated with the eigenvalue $\lambda \in \mathbb{C}$ if it satisfies

$$w^\top M(\lambda) = 0.$$

Note that this definition differs from the most commonly used definition for the left eigenvector, namely, a nonzero column vector w is a left eigenvector if there exists a λ such that $w^H M(\lambda) = 0$. The definition used here is however common for real Hamiltonian eigenvalue problems and simplifies the notation in the remainder of this text.

Now we prove two important properties of the spectrum of the NEP (1.1). Firstly, as mentioned before, the considered NEP generally has infinitely many eigenvalues. However the number of eigenvalues in any vertical strip around the imaginary axis is finite, as stated in the following proposition.

PROPOSITION 2.1. *For any $c > 0$, the characteristic matrix (1.2) has only a finite number of eigenvalues in the vertical strip $\{z \in \mathbb{C} : -c < \Re(z) < c\}$.*

Proof. This result follows from a similar argument as in [19, Corollary 2.6]. ∎

Secondly, we show that its spectrum is symmetric with respect to both the real and imaginary axis. Symmetry with respect to the real axis follows from the fact that all matrices and delays in (1.2) are real-valued, while symmetry with respect to the origin follows from the following proposition.

PROPOSITION 2.2. *If $\lambda \in \mathbb{C}$ is an eigenvalue of (1.2) with associated right eigenvector v_+ and left eigenvector w_+, then $-\lambda$ is also an eigenvalue of (1.2) and the corresponding right and left eigenvectors are Jw_+ and Jv_+, respectively.*

Proof. Since (λ, w_+) is a left eigenpair of (1.2) and J is nonsingular, we have that $w_+^\top M(\lambda)J = 0$. Transposing this equation and using Assumption 1.1 leads to $M(-\lambda)Jw_+ = 0$, meaning that $(-\lambda, Jw_+)$ is a right eigenpair of (1.2). In a similar way it can be shown that if (λ, v_+) is a right eigenpair of (1.2) then $(-\lambda, Jv_+)$ is a left eigenpair of (1.2). ∎

2.2. The structure preserving shift-invert Arnoldi method for linear finite-dimensional Hamiltonian eigenvalue problems. In this subsection we review the method presented in [16], as our paper requires a good understanding of the results in the finite-dimensional case, which will be mirrored by their infinite-dimensional counterparts in Section 3. The method in [16] allows to compute the eigenvalues of (1.7) closest to a given purely real or imaginary shift σ while preserving the Hamiltonian structure of the spectrum. Below we will assume σ purely imaginary (the discussion for σ real is however similar).

Let us first consider the traditional shift-invert Arnoldi method which considers the shifted and inverted matrix

$$(H - \sigma I_{2n})^{-1}.$$

Note that the transformed matrix is no longer real nor Hamiltonian and the purely imaginary eigenvalue $\jmath\omega$ of H is mapped to the purely imaginary eigenvalue $\mu = \frac{1}{\jmath\omega - \sigma}$ of $(H - \sigma I_{2n})^{-1}$. However, when applying the standard Arnoldi procedure to the complex matrix $(H - \sigma I_{2n})^{-1}$, the obtained approximation for μ typically does not lie on the imaginary axis since the Arnoldi procedure does not preserve the eigenvalue symmetry with respect to the imaginary axis in the constructed Hessenberg matrix. Furthermore, even if the Hessenberg matrix would contain the desired purely imaginary Ritz value, then computing the eigenvalues of this Hessenberg matrix in finite precision would introduce a small real component due to rounding errors inside the eigenvalue solver. As a consequence, the approximations for purely imaginary eigenvalues of H obtained using this approach typically have a small but non-zero real part. This means that in applications for which the detection of purely imaginary eigenvalues is important, such as the application mentioned in the introduction, additional postprocessing to determine whether an eigenvalue is purely imaginary, is necessary.

The structure preserving shift-invert Arnoldi method from [16] is therefore preferred. The Arnoldi method is now applied to the matrix

$$R_\sigma^{-1} := (H + \sigma I_{2n})^{-1} (H - \sigma I_{2n})^{-1}. \tag{2.1}$$

The transformed matrix R_σ^{-1} is real-valued and skew-Hamiltonian[1] and both the eigenvalues λ and $-\lambda$ of H are mapped to the eigenvalue $\mu = \frac{1}{\lambda^2 - \sigma^2}$ of R_σ^{-1} (which consequently has multiplicity two). For the traditional Arnoldi method such multiple eigenvalues would hamper the convergence behavior. Yet, as we will see below, each eigenvalue of R_σ^{-1} will only appear once in the Hessenberg matrix obtained by applying the Arnoldi procedure to R_σ^{-1}. More specifically, consider the following Krylov subspace generated by R_σ^{-1},

$$\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right) = \mathrm{span}\left\{q_1, R_\sigma^{-1}q_1, R_\sigma^{-2}q_1, \ldots, R_\sigma^{-(m-1)}q_1\right\}, \tag{2.2}$$

[1] A matrix S is skew-Hamiltonian if $(JS)^\top = -JS$.

with $q_1 \in \mathbb{R}^{2n}$ an arbitrary real-valued starting vector. It can be shown that this subspace is J-neutral (sometimes also referred to as isotropic), meaning that for each pair of vectors x and y in this subspace, the equality $x^\top Jy = 0$ holds, see [16, Proposition 3.3]. This has the following important consequence.

PROPOSITION 2.3. *Let q_1 be an arbitrary real-valued vector of length $2n$ and let $\lambda \neq 0$ be a simple eigenvalue of H with corresponding right eigenvector v_+ and let v_- be a right eigenvector associated with the eigenvalue $-\lambda$ of H, then the dimension of the intersection of* $\operatorname{span}\{v_+, v_-\}$ *and* $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$ *is at most 1.*

Proof. The result follows from a similar argument as in [15, Lemma 2.3], which is repeated here to ease the derivations for the infinite-dimensional case. It follows from Proposition 2.2 that if (λ, v_+) is a right eigenpair of H, then $(-\lambda, w_-)$ with $w_- := Jv_+$ is a left eigenpair of H. This implies that $v_-^\top Jv_+ = v_-^\top w_- \neq 0$, in which the last inequality follows from the fact that $-\lambda$ is a simple eigenvalue of H. Thus the intersection of $\operatorname{span}\{v_+, v_-\}$ and $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$ can not have dimension 2 (which implies that both v_+ and v_- are in $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$) as this would contradict the J-neutrality of the constructed Krylov subspace. ∎

This proposition has an important effect on the obtained Ritz values as demonstrated below. For sake of simplicity, assume that all eigenvalues of H are simple and different from zero. In this case, all eigenvalues of R_σ^{-1} have multiplicity two and the right eigenspace corresponding to the eigenvalue $\mu = \frac{1}{\lambda^2 - \sigma^2}$ is spanned by v_+ and v_-, the right eigenvectors of H associated with λ and $-\lambda$, respectively. Now let us introduce the matrix $Q_m = \begin{bmatrix} q_1 & \dots & q_m \end{bmatrix}$, with $m \leq n$, whose columns $q_1, \dots, q_m \in \mathbb{R}^{2n}$ form an orthonormal basis for $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$. Using this notation, the Arnoldi recurrence relation can be written as

$$R_\sigma^{-1} Q_m = Q_m \Psi_m + \Psi_{[m+1,m]} q_{m+1} e_m^\top \tag{2.3}$$

with $\Psi_m \in \mathbb{R}^{m \times m}$ a reduced Hessenberg matrix, $\Psi_{[m+1,m]}$ a real-valued scalar, $q_{m+1} \in \mathbb{R}^{2n}$ and $e_m \in \mathbb{R}^m$ the m^{th} Euclidean basis vector. When the Arnoldi procedure breaks down, i.e., $\Psi_{[m+1,m]} = 0$, let $\Psi_m V_m = V_m \Sigma_m$ be an eigenvalue decomposition of Ψ_m with $\Sigma_m = \operatorname{diag}(s_1, \dots, s_m)$ a diagonal matrix containing the eigenvalues of Ψ_m (the so called Ritz values) and $V_m \in \mathbb{C}^{m \times m}$ a matrix containing the associated right eigenvectors. It then follows from (2.3) that $R_\sigma^{-1}(Q_m V_m) = (Q_m V_m)\Sigma_m$, meaning that $s_1, \dots, s_m$ are also eigenvalues of R_σ^{-1} with as corresponding right eigenvectors the columns of $Q_m V_m$. However, although each eigenvalue of R_σ^{-1} has multiplicity two, there appear no doubles in $s_1, \dots, s_m$ as this would imply that there exist two linear independent vectors in $\operatorname{span}\{v_+, v_-\}$ that lie in the column space of Q_m what would contradict Proposition 2.3.

Thus, if the eigenvalues $\pm\jmath\omega$ lie sufficiently close to the shifts $\pm\sigma$, it follows from the reasoning above and the convergence behavior of the Arnoldi method that for sufficiently large m the corresponding real eigenvalue $\mu = \frac{1}{-\omega^2 - \sigma^2}$ of R_σ^{-1} will be approximated by a **real and simple Ritz value** of the **real-valued** Hessenberg matrix Ψ_m. Thus even when the eigenvalues of Ψ_m are computed in finite precision, the approximations for $\pm\jmath\omega$ obtained via the transformation $\pm\sqrt{\frac{1}{s} + \sigma^2}$, with s the computed eigenvalues of Ψ_m, are typically strictly imaginary.

Note however that to ensure that only one component of the space $\operatorname{span}\{v_+, v_-\}$ appears in the Krylov subspace, it is important that the Krylov subspace remains J-neutral. Although this property is satisfied by construction when working in exact arithmetic, J-neutrality is typically lost when working in finite precision. [16,

Section 5] therefore suggests to use the following orthogonalization procedure,

$$\Psi_{[m+1,m]} q_{m+1} = R_\sigma^{-1} q_m - Q_m \Psi_{[:,m]} - J Q_m \Upsilon_{[:,m]}, \tag{2.4}$$

with $\Psi_{[:,m]} = Q_m^\top R_\sigma^{-1} q_m$, $\Upsilon_{[:,m]} = (JQ_m)^\top R_\sigma^{-1} q_m$ and $\Psi_{[m+1,m]}$ a normalisation factor such that $\|q_{m+1}\|_2 = 1$, i.e., orthogonalise the new basis vector not only with respect to Q_m but also to JQ_m. (Note that the last term in (2.4) is zero when working in exact arithmetic.)

Remark 2.4. The transformation in (2.1) assumed that σ is either purely real or purely imaginary. However, similar results for a more general shift σ can be obtained by choosing

$$R_\sigma^{-1} = \Big((H - \sigma I_{2n})(H + \sigma I_{2n})(H - \bar{\sigma} I_{2n})(H + \bar{\sigma} I_{2n})\Big)^{-1},$$

see [15, Equation 3.2].

2.3. Two equivalent eigenvalue problems on an infinite-dimensional space. In this subsection we examine the relation between the considered NEP and two linear but infinite-dimensional eigenvalue problems. To this end, we consider the space X of continuous functions that map the interval $[-\tau_K, \tau_K]$ to $\mathbb{C}^{2n}$, denoted by $X := C\big([-\tau_K, \tau_K], \mathbb{C}^{2n}\big)$, and define the linear operator $\mathcal{H} : D(\mathcal{H}) \subseteq X \mapsto X$ as

$$\mathcal{H}\varphi(\theta) := \varphi'(\theta) \text{ for } \theta \in [-\tau_K, \tau_K], \tag{2.5}$$

in which the domain of this operator, $D(\mathcal{H})$, consists of the functions φ in X that are continuously differentiable and fulfill the condition

$$\varphi'(0) = H_0 \varphi(0) + \sum_{k=1}^{K} \Big(H_{-k} \varphi(-\tau_k) + H_k \varphi(\tau_k)\Big), \tag{2.6}$$

or in other words

$$D(\mathcal{H}) := \{\varphi \in X : \varphi' \in X \ \&\ \varphi \text{ satisfies (2.6)}\}.$$

It can be shown that the operator $\mathcal{H}$ only features a point spectrum and a complex number λ is an eigenvalue of this operator if (and only if) there exists a non-trivial function $\varphi \in D(\mathcal{H})$ such that

$$\mathcal{H}\varphi = \lambda\varphi.$$

This function φ is called the eigenfunction of $\mathcal{H}$ associated with the eigenvalue λ. From (2.5), it is clear that these eigenfunctions must have the form $ve^{\lambda \cdot}$ with $v \in \mathbb{C}^{2n}$. By plugging this result into (2.6) it is evident that there exists a correspondence between a right eigenpair (λ, v) of (1.2) and an eigenpair (λ, φ) of $\mathcal{H}$. This relation is stated more rigorously in the following proposition.

PROPOSITION 2.5. *It holds that*

1. *if (λ, v) is a right eigenpair of the NEP associated with (1.2), then $(\lambda, ve^{\lambda \cdot})$ is eigenpair of $\mathcal{H}$, and*
2. *if (λ, φ) is an eigenpair of $\mathcal{H}$, then the eigenfunction φ is of the form $ve^{\lambda \cdot}$ with (λ,v) a right eigenpair of the NEP associated with (1.2).*

Proof. The result follows from a similar argument as in [19, Proposition 2.2]. ∎

As seen in Proposition 2.5, the right eigenpairs of the considered NEP are connected with the eigenpairs of the operator $\mathcal{H}$. Now we will look for an operator which has a similar connection with the left eigenpairs of the considered NEP. To this end, let us introduce another operator $\mathcal{G} : D(\mathcal{G}) \subseteq X \mapsto X$:

$$\mathcal{G}\psi(\theta) := -\psi'(\theta) \text{ for } \theta \in [-\tau_K, \tau_K],$$

with

$$D(\mathcal{G}) := \left\{ \psi \in X : \psi' \in X \;\&\; -\psi'(0) = H_0^\top \psi(0) + \sum_{k=1}^{K} \left(H_k^\top \psi(-\tau_k) + H_{-k}^\top \psi(\tau_k) \right) \right\}.$$

The infinite-dimensional eigenvalue problem associated with this operator has the following relation with the left eigenpairs of the considered NEP.

PROPOSITION 2.6. *It holds that,*

1. *if (λ, w) is a left eigenpair of the NEP associated with (1.2), then $(\lambda, we^{-\lambda \cdot})$ is an eigenpair of $\mathcal{G}$.*
2. *if (λ, ψ) is an eigenpair of $\mathcal{G}$, then the eigenfunction ψ has the form $we^{-\lambda \cdot}$ with (λ, w) a left eigenpair of the NEP associated with (1.2).*

Proof. As before, the assertions follow from a similar argument as in [19, Proposition 2.2]. ∎

Combined with Proposition 2.2, these two propositions imply that the following relation between the eigenpairs of $\mathcal{H}$ and those of $\mathcal{G}$ holds.

COROLLARY 2.7. *If (λ, φ) is an eigenpair of $\mathcal{H}$, then $(-\lambda, \psi)$, with $\psi(\theta) = J\varphi(\theta)$ for $\theta \in [-\tau_K, \tau_K]$, is an eigenpair of $\mathcal{G}$ and visa versa.*

Finally, let us introduce the bilinear form $\mathbf{B}(\cdot, \cdot) : X \times X \mapsto \mathbb{C}$ with

$$\mathbf{B}(\varphi, \psi) := \psi(0)^\top \varphi(0) + \sum_{k=1}^{K} \left(\int_0^{\tau_k} \psi(\theta)^\top H_{-k} \varphi(\theta - \tau_k) \, \mathrm{d}\theta - \int_0^{\tau_k} \psi(\theta - \tau_k)^\top H_k \varphi(\theta) \, \mathrm{d}\theta \right). \tag{2.7}$$

Notice that this bilinear form does not define an inner product as it is neither Hermetian symmetric nor positive definite. However, this bilinear form does induce two important relations between the operators $\mathcal{H}$ and $\mathcal{G}$. Firstly, borrowing terminology from [9, Chapters 7], $\mathcal{G}$ can be seen as the *formal adjoint* of $\mathcal{H}$ with respect to the bilinear form (2.7), since the following result holds.

PROPOSITION 2.8. *For $\varphi \in D(\mathcal{H})$ and $\psi \in D(\mathcal{G})$ the equality*

$$\mathbf{B}(\mathcal{H}\varphi, \psi) = \mathbf{B}(\varphi, \mathcal{G}\psi)$$

holds.

Proof. Using the definition of both operators and partial integration, we find that

$$\begin{aligned}\mathbf{B}(\varphi, \mathcal{G}\psi) = &-\psi'(0)^\top\varphi(0) - \sum_{k=1}^{K}\left(\int_0^{\tau_k}\psi'(\theta)^\top H_{-k}\varphi(\theta-\tau_k)\,\mathrm{d}\theta - \right.\\ &\qquad\left.\int_0^{\tau_k}\psi'(\theta-\tau_k)^\top H_k\varphi(\theta)\,\mathrm{d}\theta\right)\\ =&\,\psi(0)^\top\underbrace{\left(H_0\varphi(0) + \sum_{k=1}^{K}\left(H_{-k}\varphi(-\tau_k) + H_k\varphi(\tau_k)\right)\right)}_{\varphi'(0)} +\\ &\sum_{k=1}^{K}\left(\int_0^{\tau_k}\psi(\theta)^\top H_{-k}\varphi'(\theta-\tau_k)\,\mathrm{d}\theta - \int_0^{\tau_k}\psi(\theta-\tau_k)^\top H_k\varphi'(\theta)\,\mathrm{d}\theta\right)\\ =&\,\mathbf{B}(\mathcal{H}\varphi, \psi).\end{aligned}$$

∎

Secondly, under the bilinear form (2.7) the eigenfunctions of $\mathcal{G}$ are complementary to the eigenfunctions of $\mathcal{H}$, as spelled out in the following proposition.

PROPOSITION 2.9. *If φ_λ is an eigenfunction of $\mathcal{H}$ associated with an eigenvalue λ and if ψ_μ is an eigenfunction of $\mathcal{G}$ associated with a different eigenvalue $\mu \neq \lambda$ then $\mathbf{B}(\varphi_\lambda, \psi_\mu) = 0$.*

Proof. It follows from the result above that $\mathbf{B}(\mathcal{H}\varphi_\lambda, \psi_\mu) = \mathbf{B}(\varphi_\lambda, \mathcal{G}\psi_\mu)$. Using the properties of a bilinear form one finds that

$$\begin{aligned}0 &= \mathbf{B}(\mathcal{H}\varphi_\lambda, \psi_\mu) - \mathbf{B}(\varphi_\lambda, \mathcal{G}\psi_\mu)\\ &= \mathbf{B}(\lambda\varphi_\lambda, \psi_\mu) - \mathbf{B}(\varphi_\lambda, \mu\psi_\mu)\\ &= (\lambda-\mu)\,\mathbf{B}(\varphi_\lambda, \psi_\mu).\end{aligned}$$

As $\lambda - \mu \neq 0$, $\mathbf{B}(\varphi_\lambda, \psi_\mu)$ must equal zero. ∎

3. The structure preserving shift-invert infinite Arnoldi method for Hamiltonian delay eigenvalue problems. In the previous section we saw that the eigenvalues of the NEP (1.1) correspond to those of the linear but infinite-dimensional operator $\mathcal{H}$. To compute the eigenvalues of this operator in the neighborhood of some purely real or imaginary shift σ, a modification of the infinite Arnoldi method for RDEPs will be introduced in this section. The presented method differs from the traditional shift-invert infinite Arnoldi method in that it preserves the Hamiltonian structure of the spectrum. As such, it can be seen as an extension of [16] to linear (infinite-dimensional) operators.

Let us start by deriving an adequate shift-invert transformation. Inspired by (2.1), we consider the following operator:

$$\mathcal{R}_\sigma := (\mathcal{H} - \sigma\mathcal{I}_X)(\mathcal{H} + \sigma\mathcal{I}_X),$$

with $\mathcal{I}_X$ the identity operator on X. By plugging in the definition of $\mathcal{H}$ this implies that

$$\mathcal{R}_\sigma\varphi(\theta) = \varphi''(\theta) - \sigma^2\varphi(\theta) \text{ for } \theta \in [-\tau_K, \tau_K] \text{ and } \varphi \in D(\mathcal{R}_\sigma),$$

with $D(\mathcal{R}_\sigma)$, the domain of $\mathcal{R}_\sigma$, consisting of the functions $\varphi \in X$ which are twice

continuously differentiable and fulfill the following two conditions

$$\varphi'(0) = H_0\varphi(0) + \sum_{k=1}^{K} \left(H_{-k}\varphi(-\tau_k) + H_k\varphi(\tau_k)\right) \text{ and} \tag{3.1}$$

$$\varphi''(0) = H_0\varphi'(0) + \sum_{k=1}^{K} \left(H_{-k}\varphi'(-\tau_k) + H_k\varphi'(\tau_k)\right), \tag{3.2}$$

or in other words

$$D\left(\mathcal{R}_\sigma\right) = \left\{\varphi \in X : \varphi' \in X \ , \ \varphi'' \in X \ \& \ \varphi \text{ satisfies (3.1) and (3.2)}\right\}.$$

If σ is not an eigenvalue of $\mathcal{H}$, then the operator $\mathcal{R}_\sigma$ is invertible. For $\sigma \neq 0$ one finds:

$$\mathcal{R}_\sigma^{-1}\phi\left(\theta\right) = \left(\int_0^\theta \tfrac{\phi(\eta)}{2\sigma} e^{-\sigma\eta} d\eta + \mathrm{C}_\sigma[\phi]\right) e^{\sigma\theta} + \left(-\int_0^\theta \tfrac{\phi(\eta)}{2\sigma} e^{\sigma\eta} d\eta + \mathrm{C}_{-\sigma}[\phi]\right) e^{-\sigma\theta}, \tag{3.3}$$

in which the constants $\mathrm{C}_\sigma[\phi]$ and $\mathrm{C}_{-\sigma}[\phi]$ are uniquely defined by conditions (3.1) and (3.2). After some straight forward (but tedious) calculations we find

$$2\sigma M(\sigma)\,\mathrm{C}_\sigma[\phi] = -\phi(0) + \sum_{k=1}^{K}\left[H_k \int_0^{\tau_k} \phi(\eta) e^{-\sigma(\eta-\tau_k)}\,\mathrm{d}\eta - H_{-k}\int_0^{\tau_k}\phi(\eta-\tau_k)e^{-\sigma\eta}\,\mathrm{d}\eta\right] \tag{3.4}$$

and

$$2\sigma M(-\sigma)\,\mathrm{C}_{-\sigma}[\phi] = \phi(0) - \sum_{k=1}^{K}\left[H_k \int_0^{\tau_k} \phi(\eta) e^{\sigma(\eta-\tau_k)}\,\mathrm{d}\eta - H_{-k}\int_0^{\tau_k}\phi(\eta-\tau_k)e^{\sigma\eta}\,\mathrm{d}\eta\right], \tag{3.5}$$

with $M(\cdot)$ the characteristic matrix as introduced in (1.2). Note that to determine the integration constants $C_\sigma[\phi]$ and $C_{-\sigma}[\phi]$, one has to solve a linear system with both $M(\sigma)$ and $M(-\sigma)$ to preserve the symmetries in the spectrum. For $\sigma = 0$, one finds

$$\mathcal{R}_0^{-1}\phi\left(\theta\right) = \int_0^\theta \int_0^{\eta_2} \phi(\eta_1)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2 + \mathrm{C}_1[\phi]\theta + \mathrm{C}_0[\phi], \tag{3.6}$$

in which $\mathrm{C}_1[\phi]$ and $\mathrm{C}_0[\phi]$ again follow from (3.1) and (3.2):

$$\begin{aligned} M(0)\ \mathrm{C}_1[\phi] &= -\phi(0) + \sum_{k=1}^{K}\left(H_k\int_0^{\tau_k}\phi(\eta)\,\mathrm{d}\eta + H_{-k}\int_0^{-\tau_k}\phi(\eta)\,\mathrm{d}\eta\right), \\ M(0)\ \mathrm{C}_0[\phi] &= -\mathrm{C}_1[\phi] + \sum_{k=1}^{K}\left[H_k\left(\int_0^{\tau_k}\int_0^{\eta_2}\phi(\eta_1)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2 + \mathrm{C}_1[\phi]\tau_k\right) + \right. \\ &\qquad \left. H_{-k}\left(\int_0^{-\tau_k}\int_0^{\eta_2}\phi(\eta_1)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2 - \mathrm{C}_1[\phi]\tau_k\right)\right]. \end{aligned}$$

Next, we state some important properties of this operator $\mathcal{R}_\sigma^{-1}$.

PROPOSITION 3.1. *If $\phi \in X$ is a real-valued function and σ is purely real or purely imaginary and not an eigenvalue of $\mathcal{H}$, then $\mathcal{R}_\sigma^{-1}\phi$ is a real function.*

Proof. For $\sigma \neq 0$, this assertion follows directly from (3.3), (3.4) and (3.5) by noting that $\mathrm{C}_\sigma[\phi]$ and $\mathrm{C}_{-\sigma}[\phi]$ are real if σ is real and that $\mathrm{C}_{-\sigma}[\phi]$ is the complex conjugate of $\mathrm{C}_\sigma[\phi]$ if σ is purely imaginary. For $\sigma = 0$ this result follows from (3.6) and the expressions for $C_1[\phi]$ and $C_0[\phi]$ given above. ∎

LEMMA 3.2. *For σ not an eigenvalue of $\mathcal{H}$, the operator $\mathcal{R}_\sigma^{-1}$ is self-adjoint with respect to $\mathbf{B}(\cdot, J\cdot)$ meaning that the equality $\mathbf{B}\left(\mathcal{R}_\sigma^{-1}\varphi, J\psi\right) = \mathbf{B}\left(\varphi, J{\mathcal{R}_\sigma}^{-1}\psi\right)$ holds.*

Proof. Here we restrict ourselves to the case $\sigma \neq 0$ as the proof for $\sigma = 0$ is similar. Using the definition of the bilinear form $\mathbf{B}(\cdot,\cdot)$ given in (2.7), and using (3.3) we find

$$\begin{aligned}
\mathbf{B}\left(\mathcal{R}_\sigma^{-1}\varphi, J\psi\right) = -\Bigg[&\psi(0)^\top J\Big(\mathrm{C}_\sigma[\varphi] + \mathrm{C}_{-\sigma}[\varphi]\Big) + \\
&\sum_{k=1}^{K}\Bigg(\int_0^{\tau_k}\psi(\theta)^\top JH_{-k}\Bigg[\Bigg(\mathrm{C}_\sigma[\varphi] + \int_0^{\theta-\tau_k}\tfrac{\varphi(\eta)}{2\sigma}e^{-\sigma\eta}\,\mathrm{d}\eta\Bigg)e^{\sigma(\theta-\tau_k)} + \\
&\qquad\Bigg(\mathrm{C}_{-\sigma}[\varphi] - \int_0^{\theta-\tau_k}\tfrac{\varphi(\eta)}{2\sigma}e^{\sigma\eta}\,\mathrm{d}\eta\Bigg)e^{-\sigma(\theta-\tau_k)}\Bigg]\,\mathrm{d}\theta \\
&- \int_0^{\tau_k}\psi(\theta-\tau_k)^\top JH_k\Bigg[\Bigg(\mathrm{C}_\sigma[\varphi] + \int_0^{\theta}\tfrac{\varphi(\eta)}{2\sigma}e^{-\sigma\eta}\,\mathrm{d}\eta\Bigg)e^{\sigma\theta} + \\
&\qquad\Bigg(\mathrm{C}_{-\sigma}[\varphi] - \int_0^{\theta}\tfrac{\varphi(\eta)}{2\sigma}e^{\sigma\eta}\,\mathrm{d}\eta\Bigg)e^{-\sigma\theta}\Bigg]\,\mathrm{d}\theta\Bigg)\Bigg].
\end{aligned}$$

Firstly, observe that

$$\begin{aligned}
&\int_0^{\tau_k}\psi(\theta)^\top JH_{-k}\int_0^{\theta-\tau_k}\tfrac{\varphi(\eta)}{2\sigma}e^{-\sigma\eta}\,\mathrm{d}\eta e^{\sigma(\theta-\tau_k)}\,\mathrm{d}\theta \\
&\qquad= \int_0^{-\tau_k}\int_0^{\tau_k+\eta}\psi(\theta)^\top e^{\sigma(\theta-\tau_k)}\,\mathrm{d}\theta JH_{-k}\tfrac{\varphi(\eta)}{2\sigma}e^{-\sigma\eta}\,\mathrm{d}\eta \\
&\qquad= \int_{\tau_k}^{0}\int_0^{\hat\eta}\psi(\theta)^\top e^{\sigma(\theta-\tau_k)}\,\mathrm{d}\theta JH_{-k}\tfrac{\varphi(\hat\eta-\tau_k)}{2\sigma}e^{-\sigma(\hat\eta-\tau_k)}\,\mathrm{d}\hat\eta \\
&\qquad= -\int_0^{\tau_k}\int_0^{\hat\eta}\tfrac{\psi(\theta)}{2\sigma}^\top e^{\sigma\theta}\,\mathrm{d}\theta e^{-\sigma\hat\eta}JH_{-k}\varphi(\hat\eta-\tau_k)\,\mathrm{d}\hat\eta,
\end{aligned}$$

in which we first interchanged the integrating variables, then made the change of variable $\hat\eta = \tau_k + \eta$ and finally interchanged the integration boundaries and reordered the terms. Using the same procedure we find:

$$-\int_0^{\tau_k}\psi(\theta)^\top JH_{-k}\int_0^{\theta-\tau_k}\tfrac{\varphi(\eta)}{2\sigma}e^{\sigma\eta}\,\mathrm{d}\eta\, e^{-\sigma(\theta-\tau_k)}\,\mathrm{d}\theta = \int_0^{\tau_k}\int_0^{\hat\eta}\tfrac{\psi(\theta)}{2\sigma}^\top e^{-\sigma\theta}\,\mathrm{d}\theta e^{\sigma\hat\eta}JH_{-k}\varphi(\hat\eta-\tau_k)\,\mathrm{d}\hat\eta.$$

In a similar fashion we obtain:

$$-\int_0^{\tau_k}\psi(\theta-\tau_k)^\top JH_k\int_0^{\theta}\tfrac{\varphi(\eta)}{2\sigma}e^{-\sigma\eta}\,\mathrm{d}\eta e^{\sigma\theta}\,\mathrm{d}\theta = \int_0^{\tau_k}\int_0^{\eta-\tau_k}\tfrac{\psi(\hat\theta)}{2\sigma}^\top e^{\sigma\hat\theta}\,\mathrm{d}\hat\theta e^{-\sigma(\eta-\tau_k)}JH_k\varphi(\eta)\,\mathrm{d}\eta$$

and

$$\int_0^{\tau_k}\psi(\theta-\tau_k)^\top JH_k\int_0^{\theta}\tfrac{\varphi(\eta)}{2\sigma}e^{\sigma\eta}\,\mathrm{d}\eta e^{-\sigma\theta}\,\mathrm{d}\theta = -\int_0^{\tau_k}\int_0^{\eta-\tau_k}\tfrac{\psi(\hat\theta)}{2\sigma}^\top e^{-\sigma\hat\theta}\,\mathrm{d}\hat\theta e^{\sigma(\eta-\tau_k)}JH_k\varphi(\eta)\,\mathrm{d}\eta,$$

in which we used the change of variables $\hat{\theta} = \theta - \tau_k$. Next, by grouping the terms with $\mathrm{C}_\sigma[\varphi]$ and using $J^\top = -J$, $JH_{-k} = (JH_k)^\top$ and $J\big(2\sigma M(\sigma)\big)^{-1} = -\big(2\sigma M(-\sigma)\big)^{-T} J$ we find

$$\begin{aligned}
&\left(\psi(0)^\top J + \sum_{k=1}^{K}\left[\int_0^{\tau_k} \psi(\theta)^\top JH_{-k} e^{\sigma(\theta-\tau_k)}\,\mathrm{d}\theta - \int_0^{\tau_k} \psi(\theta-\tau_k)^\top JH_k e^{\sigma\theta}\,\mathrm{d}\theta\right]\right) C_\sigma[\varphi] \\
&\quad = \left(\psi(0)^\top - \sum_{k=1}^{K}\left[\int_0^{\tau_k} \psi(\theta)^\top H_k^\top e^{\sigma(\theta-\tau_k)}\,\mathrm{d}\theta - \int_0^{\tau_k} \psi(\theta-\tau_k)^\top H_{-k}^\top e^{\sigma\theta}\,\mathrm{d}\theta\right]\right) J\mathrm{C}_\sigma[\varphi] \\
&= \mathrm{C}_{-\sigma}[\psi]^\top J\left(\varphi(0) - \sum_{k=1}^{K}\left[-H_{-k}\int_0^{\tau_k} \varphi(\eta-\tau_k)e^{-\sigma\eta}\,\mathrm{d}\eta + H_k\int_0^{\tau_k} \varphi(\eta)e^{-\sigma(\eta-\tau_k)}\,\mathrm{d}\eta\right]\right).
\end{aligned}$$

Similarly for the terms with $C_{-\sigma}[\varphi]$ we find

$$\begin{aligned}
&\left(\psi(0)^\top J + \sum_{k=1}^{K}\left[\int_0^{\tau_k} \psi(\theta)^\top JH_{-k} e^{-\sigma(\theta-\tau_k)}\,\mathrm{d}\theta - \int_0^{\tau_k} \varphi(\theta-\tau_k)^\top JH_k e^{-\sigma\theta}\,\mathrm{d}\theta\right]\right) \mathrm{C}_{-\sigma}[\varphi] \\
&\quad = C_\sigma[\psi]^\top J\left(\varphi(0) + \sum_{k=1}^{K}\left[H_{-k}\int_0^{\tau_k} \varphi(\eta-\tau_k)e^{\sigma\eta}\,\mathrm{d}\eta - H_k\int_0^{\tau_k} \varphi(\eta)e^{\sigma(\eta-\tau_k)}\,\mathrm{d}\eta\right]\right).
\end{aligned}$$

Combining all these relations, we obtain the desired result. ∎

Next, we introduce some properties of the (infinite-dimensional) eigenvalue problem associated with this operator:

$$\mathcal{R}_\sigma^{-1}\phi = \mu\phi. \tag{3.7}$$

PROPOSITION 3.3. *For σ not an eigenvalue of $\mathcal{H}$, the infinite-dimensional eigenvalue problem in (3.7) satisfies the following three properties .*

1. *If λ is an eigenvalue of $\mathcal{H}$ then $\mu = \frac{1}{\lambda^2-\sigma^2}$ is an eigenvalue of $\mathcal{R}_\sigma^{-1}$ and visa versa, if μ is an eigenvalue of $\mathcal{R}_\sigma^{-1}$ then $\pm\sqrt{\frac{1}{\mu}+\sigma^2}$ are eigenvalues of $\mathcal{H}$.*
2. *Let $\lambda \neq 0$ and $-\lambda$ be simple eigenvalues of $\mathcal{H}$ with eigenfunctions φ_+ and φ_-, respectively, then the eigenspace of $\mu = \frac{1}{\lambda^2-\sigma^2}$ is spanned by $\{\varphi_+, \varphi_-\}$.*
3. *The eigenvalues of $\mathcal{R}_\sigma^{-1}$ have even multiplicity.*

Proof. First we will show that if $\varphi \in D(\mathcal{H})$ is an eigenfunction of $\mathcal{H}$ associated with λ, i.e., $\mathcal{H}\varphi = \lambda\varphi$, then this function is also an eigenfunction of $\mathcal{R}_\sigma^{-1}$ associated with $\mu = \frac{1}{\lambda^2-\sigma^2}$. This result follows immediately from the following equality:

$$\mathcal{R}_\sigma\varphi = (\mathcal{H} - \sigma I_X)(\mathcal{H} + \sigma I_X)\varphi = (\lambda+\sigma)(\mathcal{H} - \sigma I_X)\varphi = (\lambda^2 - \sigma^2)\varphi,$$

which implies that $\mathcal{R}_\sigma^{-1}\varphi = \frac{1}{\lambda^2-\sigma^2}\varphi$. Visa versa, if there exists a ϕ such that $\mathcal{R}_\sigma^{-1}\phi = \mu\phi$, then $\left(\mathcal{H} - \sqrt{\frac{1}{\mu}+\sigma^2}\mathcal{I}_X\right)\left(\mathcal{H} + \sqrt{\frac{1}{\mu}+\sigma^2}\mathcal{I}_X\right)\phi = 0$, meaning that either $\mathcal{H} - \sqrt{\frac{1}{\mu}+\sigma^2}\mathcal{I}_X$ or $\mathcal{H} + \sqrt{\frac{1}{\mu}+\sigma^2}\mathcal{I}_X$ has a non-trivial null space. The fact that both $\pm\sqrt{\frac{1}{\mu}+\sigma^2}$ are eigenvalues of $\mathcal{H}$ follows from Proposition 2.2.

Thus if $\lambda \neq 0$ is an eigenvalue of $\mathcal{H}$, then both λ and $-\lambda$ are mapped together to $\mu = \frac{1}{\lambda^2-\sigma^2}$ and the eigenspace associated with μ is spanned by their eigenfunctions.

If $\lambda = 0$ is an eigenvalue of $\mathcal{H}$, then it has even multiplicity. As a consequence, $-\sigma^{-2}$ is an eigenvalue of $\mathcal{R}_\sigma^{-1}$ with the same multiplicity. ∎

Next we examine the following Krylov subspace generated by $\mathcal{R}_\sigma^{-1}$:

$$\mathrm{K}_m\left(\mathcal{R}_\sigma^{-1}, \phi_1\right) = \operatorname{span}\left\{\phi_1, \mathcal{R}_\sigma^{-1}\phi_1, \mathcal{R}_\sigma^{-2}\phi_1, \ldots, \mathcal{R}_\sigma^{-(m-1)}\phi_1\right\} \subseteq X,$$

with $\phi_1 \in X$ an arbitrary real-valued starting function. This subspace has the following important property.

PROPOSITION 3.4. *Let σ not be an eigenvalue of $\mathcal{H}$ and let $\phi_1 \in X$ be a real-valued function, then for any two functions φ and ψ in the Krylov subspace $\mathrm{K}_m\left(\mathcal{R}_\sigma^{-1}, \phi_1\right)$ the equality $\mathbf{B}\left(\varphi, J\psi\right) = 0$ holds, with $\mathbf{B}\left(\cdot,\cdot\right)$ as defined in (2.7).*

Proof. Each function in $\mathrm{K}_m\left(\mathcal{R}_\sigma^{-1}, \phi_1\right)$ can be written as $P_{m-1}(\mathcal{R}_\sigma^{-1})\phi_1$ with $P_{m-1}(\cdot)$ a polynomial of degree (at most) $m-1$. Due to the bilinearity of $\mathbf{B}\left(\cdot, J\cdot\right)$, it thus suffices to prove that $\mathbf{B}\left(\mathcal{R}_\sigma^{-i}\phi_1, J\mathcal{R}_\sigma^{-j}\phi_1\right) = 0$ for $i, j = 0, \ldots, m-1$.

By Lemma 3.2 we have

$$\mathbf{B}\left(\mathcal{R}_\sigma^{-i}\phi_1, J\mathcal{R}_\sigma^{-j}\phi_1\right) = \mathbf{B}\left(\mathcal{R}_\sigma^{-(j+i)}\phi_1, J\phi_1\right) = \mathbf{B}\left(\phi_1, J\mathcal{R}_\sigma^{-(j+i)}\phi_1\right).$$

However, due to the fact that $\mathbf{B}\left(\varphi, J\psi\right) = -\mathbf{B}\left(\psi, J\varphi\right)$, i.e., $\mathbf{B}\left(\cdot, J\cdot\right)$ is anti-symmetric (see Appendix A), we also have

$$\mathbf{B}\left(\mathcal{R}_\sigma^{-(j+i)}\phi_1, J\phi_1\right) = -\mathbf{B}\left(\phi_1, J\mathcal{R}_\sigma^{-(j+i)}\phi_1\right),$$

which implies that $\mathbf{B}\left(\mathcal{R}_\sigma^{-i}\phi_1, J\mathcal{R}_\sigma^{-j}\phi_1\right) = 0$. ∎

This proposition has the following important consequence.

THEOREM 3.5. *For σ not an eigenvalue of $\mathcal{H}$ and $\phi_1 \in X$ a real-valued function, let $\lambda \neq 0$ be a simple eigenvalue of $\mathcal{H}$ (or equivalently of the NEP (1.1)) with associated eigenfuction φ_+ and let φ_- be the eigenfunction associated with the eigenvalue $-\lambda$, then the dimension of the intersection of $\operatorname{span}\{\varphi_+, \varphi_-\}$ and $\mathrm{K}_m(\mathcal{R}_\sigma^{-1}, \phi_1)$ is at most 1.*

Proof. Recall from Proposition 2.5 that $\varphi_+(\theta) = v_+ e^{\lambda\theta}$ and $\varphi_-(\theta) = v_- e^{-\lambda\theta}$, with v_+ and v_- right eigenvectors of the NEP (1.1) associated with λ and $-\lambda$, respectively. Furthermore, from Proposition 2.2 it follows that there exists a vector w_+ which is a left eigenvector of (1.2) associated with λ, such that $v_- = Jw_+$. Combing these properties with the definition of the bilinear form in (2.7), we find that

$$\begin{aligned}\mathbf{B}\left(\varphi_+, J\varphi_-\right) &= -\Bigg[w_+^\top v_+ + \sum_{k=1}^{K}\Bigg(\int_0^{\tau_k} e^{-\lambda\theta} w_+^\top H_{-k} v_+ e^{\lambda(\theta-\tau_k)}\,\mathrm{d}\theta \\ &\qquad\qquad - \int_0^{\tau_k} e^{-\lambda(\theta-\tau_k)} w_+^\top H_k v_+ e^{\lambda\theta}\,\mathrm{d}\theta\Bigg)\Bigg] \\ &= -w_+^\top\left[I + \sum_{k=1}^{K}\left(H_{-k}e^{-\lambda\tau_k}\tau_k - H_k e^{\lambda\tau_k}\tau_k\right)\right] v_+ = -w_+^\top \frac{\partial M(\lambda)}{\partial\lambda} v_+ \neq 0,\end{aligned}$$

in which the last inequality follows from the fact that λ is a simple eigenvalue of the NEP (1.1). Thus if the intersection of $\operatorname{span}\{\varphi_+, \varphi_-\}$ and $\mathrm{K}_m(\mathcal{R}_\sigma^{-1}, \phi_1)$ has dimension two, then both φ_+ and φ_- must lie inside this Krylov subspace, which would contradict Proposition 3.4. ∎

Combining this result with Proposition 3.3, we have the following important result.

COROLLARY 3.6. *Under the conditions of Theorem 3.5 and assuming that* $\lambda \neq 0$ *is a simple eigenvalue of* $\mathcal{H}$*, the dimension of the intersection of the eigenspace of* $\mathcal{R}_\sigma^{-1}$ *associated with* $\mu = 1/(\lambda^2 - \sigma^2)$ *and the Krylov subspace* $\mathrm{K}_m\left(\mathcal{R}_\sigma^{-1}, \phi_1\right)$ *is at most 1.*

We conclude this section with Algorithm 3.1, which gives a high-level description of the structure preserving shift-invert infinite Arnoldi method operating on the infinite-dimensional space X. In this algorithm $\Psi_{[i,j]}$ denotes the element on the i^{th} row and j^{th} column of Ψ, $\Psi_{[1:m,:]}$ is the submatrix of Ψ consisting of its first m rows, $\langle \cdot, \cdot \rangle_X$ is an appropriate inner product on X (which we will define later on), and $\| \cdot \|_X$ is the norm associated with this inner product.

The algorithm works as follows. Starting from a real initial function ϕ_1, an orthonormal basis for the Krylov subspace $\mathrm{K}_m(\mathcal{R}_\sigma^{-1}, \phi_1)$ is constructed iteratively. After obtaining this orthonormal basis, the eigenvalues of $\mathcal{R}_\sigma^{-1}$ are approximated by the eigenvalues of its orthogonal projection onto the Krylov subspace, i.e., the submatrix $\Psi_{[1:m,:]}$, which has a reduced Hessenberg structure. Finally, the approximations for the eigenvalues of $\mathcal{H}$ can be obtained using the transformation from Proposition 3.3. Each iteration in Algorithm 3.1 consists of the following steps. First the Krylov subspace is extended with the candidate function $\varphi_{i+1} = \mathcal{R}_\sigma^{-1}\phi_i$. Next, this function is orthogonalized against the already obtained basis vectors and subsequently normalized to have norm 1. Note that in contrast to the traditional infinite Arnoldi method, one now has to ensure that ϕ_{i+1} is also *orthogonal* against $\phi_1, \ldots, \phi_i$ with respect to $\mathbf{B}(\cdot, J\cdot)$. Although this last orthogonality constraint holds by construction in exact arithmetic (see Proposition 3.4), it is no longer the case when working in finite precision, as we shall illustrate in Subsection 6.1.

Inspired by [13], we define the following inner product on X based on the scaled Chebyshev expansion. More precisely, let $\phi \in X$ and $\psi \in X$ be Lipschitz continuous functions, then for each of these functions there exists a unique Chebyshev expansion series that is absolute and uniformly convergent [23]:

$$\varphi(\theta) = \sum_{l=0}^{\infty} c_l T_l\left(\tfrac{\theta}{\tau_K}\right) \text{ and } \psi(\theta) = \sum_{l=0}^{\infty} d_l T_l\left(\tfrac{\theta}{\tau_K}\right) \text{ for } \theta \in [-\tau_K, \tau_K],$$

with $T_l(\cdot) := \cos\left(l \arccos(\cdot)\right)$ the l^{th} Chebyshev polynomial of the first kind. The inner product of these two functions is then defined as

$$\langle \psi, \varphi \rangle_X := \textstyle\sum_{l=0}^{\infty} c_l^H d_l. \tag{3.8}$$

Furthermore, it can be shown, see [13, Equation 4.10], that this inner product is equivalent with the following expression

$$\langle \psi, \varphi \rangle_X = \tfrac{2}{\pi} I\left[\varphi^H \psi\right] - \tfrac{1}{\pi^2} I[\varphi^H] I[\psi].$$

with $I[f] = \int_{-\tau_K}^{\tau_K} \frac{f(\theta)}{\sqrt{\tau_K^2 - \theta^2}}\, \mathrm{d}\theta$.

As in the finite-dimensional case (see Subsection 2.2), Theorem 3.5 and Corollary 3.6 have an important consequence for the obtained approximations of purely imaginary eigenvalues of (1.1). More specifically, since ϕ_1 is chosen real-valued, Proposition 3.1 and the definition of the inner product in (3.8), imply that $\mathrm{K}_m\left(\mathcal{R}_\sigma^{-1}, \phi_1\right)$ consists of real-valued functions and that the matrix Ψ is real-valued. Now, let $\pm \jmath\omega$

be simple eigenvalues of $\mathcal{H}$ sufficiently close to $\pm\sigma$ and m sufficiently large. These two eigenvalues are mapped together to the purely real eigenvalue $\mu = \frac{1}{-\omega^2-\sigma^2}$ of $\mathcal{R}_\sigma^{-1}$. However, as long as the Krylov subspace satisfies Proposition 3.4 (which is always the case in exact arithmetic), only a single component of the two-dimensional eigenspace associated with μ will be approximated in the Krylov subspace and hence the multiple eigenvalue μ will be approximated by a simple, real Ritz value in the reduced Hessenberg matrix $\Psi[1:m,:]$. After re-transformation the obtained approximations for $\pm\jmath\omega$ are thus typically purely imaginary. In contrast, if orthogonality with respect to $\mathbf{B}(\cdot, J\cdot)$ is lost (due to rounding errors), the Krylov subspace will eventually also approximate a second component in the eigenspace of μ. As a consequence, the double eigenvalue μ will be approximated by a pair of Ritz values: either two nearby real Ritz values or a pair of complex conjugate Ritz values. In the latter case, after re-transformation the obtained approximations for the purely imaginary eigenvalues $\pm\jmath\omega$ of $\mathcal{H}$ thus have a small real part.

Algorithm 3.1 High-level description of the structure preserving shift-invert infinite Arnoldi method for NEP (1.1).

Choose a **real-valued** initial function ϕ_1 with $\|\phi_1\|_X = 1$.
Define a $(m+1)\times m$ matrix Ψ with all elements equal to zero.
for $i = 1, \ldots, m$ **do**
  Extend the Krylov subspace with $\varphi_{i+1} = \mathcal{R}_\sigma^{-1}\phi_i$.
  Orthogonalize φ_{i+1} against the already obtained orthogonal basis for the Krylov subspace and normalize the result:

$$\Psi_{[i+1,i]}\phi_{i+1} = \varphi_{i+1} - \Psi_{[1,i]}\,\phi_1 - \cdots - \Psi_{[i,i]}\,\phi_i$$

  with $\Psi_{[j,i]} = \langle \phi_j, \varphi_{i+1}\rangle_X$ for $j = 1, \ldots, i$ and $\Psi_{[i+1,i]}$ such that $\|\phi_{i+1}\|_X = 1$.
  Ensure orthogonality of ϕ_{i+1} against $\phi_1, \ldots, \phi_i$ with respect to $\mathbf{B}(\cdot, J\cdot)$.
end for
Compute the eigenvalues μ of $\Psi_{[1:m,:]}$.
Return the approximations for the eigenvalues of (1.1): $\pm\sqrt{\frac{1}{\mu}+\sigma^2}$.

The algorithm introduced above is described in terms of functions and therefore not directly implementable using finite-dimensional operations. The following two sections will therefore describe how this algorithm can be implemented using finite-dimensional linear algebra operations and how one can ensure that the built up Krylov subspace satisfies Proposition 3.4, even in the presence of rounding errors.

4. Numerical implementation for the shift equal to zero. We start with the case $\sigma = 0$. First we will see that by choosing an appropriate initial function ϕ_1, a natural finite-dimensional representation for the functions and operations in Algorithm 3.1 appears. To this end, observe from (3.3) that if $\mathcal{R}_0^{-1}$ is applied on a vector-valued polynomial of degree N, the result is a vector-valued polynomial of degree $N+2$. As a consequence, when choosing a vector-valued polynomial as initial function, the constructed Krylov subspace consists entirely of vector-valued polynomials. By choosing an appropriate basis for the space of polynomials, Algorithm 3.1 can thus be carried out using finite-dimensional operations on the coefficient vectors of these polynomials. Due to our choice for the inner product in (3.8), it makes sense to work with scaled Chebyshev functions. The extension step in Algorithm 3.1 can

now be computed using straightforward linear algebra operations, as demonstrated in the following theorem.

THEOREM 4.1. *Let*

$$\phi_i(\theta) = \sum_{l=0}^{N_i} q_l^{(i)} T_l\left(\frac{\theta}{\tau_K}\right)$$

with $q_l^{(i)} \in \mathbb{R}^{2n}$ *for* $l = 0, \ldots, N_i$, *then*

$$\varphi_{i+1}(\theta) = \mathcal{R}_0^{-1}\phi_i(\theta) = \sum_{l=0}^{N_i+2} v_l^{(i+1)} T_l\left(\frac{\theta}{\tau_K}\right),$$

in which the coefficient vectors $v_2^{(i+1)}, \ldots, v_{N_i+2}^{(i+1)}$ *follow from*
(4.1)

$$\begin{bmatrix} v_2^{(i+1)} & \cdots & v_{N_i+2}^{(i+1)} \end{bmatrix} = \tau_K^2 \begin{bmatrix} q_0^{(i)} & \cdots & q_{N_i}^{(i)} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & & & & & \\ & \frac{1}{24} & & & & \\ \frac{-1}{6} & & \frac{1}{48} & & & \\ & \frac{-1}{16} & & \frac{1}{80} & & \\ \frac{1}{24} & & \frac{-1}{30} & & \frac{1}{120} & \\ & \ddots & & \ddots & & \ddots \\ & & \frac{1}{4(N_i-1)(N_i-2)} & & \frac{-1}{2(N_i-1)(N_i+1)} & & \frac{1}{4(N_i+1)(N_i+2)} \end{bmatrix},$$

while $v_1^{(i+1)}$ *and* $v_0^{(i+1)}$ *follow from* (3.1) *and* (3.2)*:*

$$M(0)\, v_1^{(i+1)} = \sum_{l=2}^{N_i+2} \left(H_0 T_l'(0) + \sum_{k=1}^{K} \left(H_k T_l'\left(\tfrac{\tau_k}{\tau_K}\right) + H_{-k} T_l'\left(\tfrac{-\tau_k}{\tau_K}\right)\right) - I_{2n} \tfrac{T_l''(0)}{\tau_K} \right) v_l^{(i+1)}$$

and

$$M(0)\, v_0^{(i+1)} = \sum_{l=1}^{N_i+2} \left(H_0 T_l(0) + \sum_{k=1}^{K} \left(H_k T_l\left(\tfrac{\tau_k}{\tau_K}\right) + H_{-k} T_l\left(\tfrac{-\tau_k}{\tau_K}\right)\right) - I_{2n} \tfrac{T_l'(0)}{\tau_K} \right) v_l^{(i+1)},$$

with $M(\cdot)$ *the characteristic matrix as defined in* (1.2).

Proof. We refer to Appendix B for more details on this result. ∎

Notice that for large n, the main computation cost of the extension step consists of solving the systems for $v_0^{(i+1)}$ and $v_1^{(i+1)}$. However, because the same matrix $M(0)$ is used in each iteration, its factorization needs only to be computed ones, which greatly reduces the computation time.

Secondly, due to the chosen initial function, also the orthogonalisation step can be carried out using standard linear algebra operations. More specifically, as a consequence of Theorem 4.1, the functions encountered in the i^{th} orthogonalisation step of Algorithm 3.1 can be written as
(4.2)

$$\phi_j(\theta) = \sum_{l=0}^{N_j} q_l^{(j)} T_l\left(\frac{\theta}{\tau_K}\right) \text{ for } j = 1, \ldots, i+1 \text{ and } \varphi_{i+1}(\theta) = \sum_{l=0}^{N_{i+1}} v_l^{(i+1)} T_l\left(\frac{\theta}{\tau_K}\right),$$

with $q_l^{(j)}$ and $v_l^{(i+1)}$ belonging to $\mathbb{R}^{2n}$. By introducing $v^{(i+1)}$ as the vector containing the stacked coefficients of φ_{i+1}, $q^{(j,i+1)}$ the vector containing the stacked coefficients

of ϕ_j padded with zeros to length $2n(N_{i+1}+1)$ and $Q_{(i,i+1)}$ the matrix whose columns consists of the vectors $q^{(j,i+1)}$ for $j = 1, \ldots, i$ the orthogonalisation step reduces to

$$\Psi_{[i+1,i]} q^{(i+1,i+1)} = v^{(i+1)} - Q_{(i,i+1)} Q_{(i,i+1)}^\top v^{(i+1)} \tag{4.3}$$

with $\Psi_{[i+1,i]}$ a normalisation constant such that $\|q^{(i+1,i+1)}\|_2 = 1$.

Up till now, the presented approach closely resembles the classic infinite Arnoldi method from [13], but with the degree of the polynomials in the Krylov subspace increasing with two instead of one in each iteration. Recall however that it is important that the build-up Krylov subspace satisfies the condition in Proposition 3.4. Although this condition holds automatically in exact arithmetic, rounding errors cause a loss of orthogonality when working in finite precision. In finite precision arithmetic one therefore needs an *additional orthogonalisation step*, which ensures orthogonality with respect to $\mathbf{B}(\cdot, J\cdot)$. To this end, note that for $\{\phi_j\}_{j=1}^i$ and φ_{i+1} in form (4.2) the bilinear form $\mathbf{B}(\cdot, J\cdot)$ can be evaluated using matrix-vector operations:

$$\mathbf{B}(\phi_j, J\varphi_{i+1}) = {v^{(i+1)}}^\top \left(S^0_{N_{i+1}} \otimes J + \sum_{k=1}^{K} \left(S^{-k}_{N_{i+1}} \otimes \left(JH_{-k}\right) + S^{k}_{N_{i+1}} \otimes \left(JH_k\right) \right) \right) q^{(j,i+1)},$$

in which $\otimes$ is the Kronecker product, the matrices $S^0_{N_{i+1}}$, $S^1_{N_{i+1}}$, $S^{-1}_{N_{i+1}}$, $\ldots$, $S^K_{N_{i+1}}$ and $S^{-K}_{N_{i+1}}$ belong to $\mathbb{R}^{(N_{i+1}+1)\times(N_{i+1}+1)}$ and the elements on position (l_1, l_2) of $S^0_{N_{i+1}}$, $S^{-k}_{N_{i+1}}$ and $S^k_{N_{i+1}}$ are equal to $-T_{l_1}(0)T_{l_2}(0)$, $-\int_0^{\tau_k} T_{l_1}\left(\frac{\theta}{\tau_K}\right) T_{l_2}\left(\frac{\theta-\tau_k}{\tau_K}\right) \mathrm{d}\theta$ and $\int_0^{\tau_k} T_{l_1}\left(\frac{\theta-\tau_k}{\tau_K}\right) T_{l_2}\left(\frac{\theta}{\tau_K}\right) \mathrm{d}\theta$, respectively. Furthermore, observe that the matrix

$$S_{N_{i+1}} = S^0_{N_{i+1}} \otimes J + \sum_{k=1}^{K} \left(S^{-k}_{N_{i+1}} \otimes JH_{-k} + S^{k}_{N_{i+1}} \otimes JH_k \right)$$

is skew-symmetric, i.e., $S^\top_{N_{i+1}} = -S_{N_{i+1}}$.

In the i^{th} iteration step, the vector $q^{(i+1,i+1)}$ must thus be orthogonal with respect to the columns of both $Q_{(i,i+1)}$ and $S_{N_{i+1}} Q_{(i,i+1)}$. Moreover, because ${q^{(j,i+1)}}^\top S_{N_{i+1}} q^{(j,i+1)} = 0$ for $j = 1, \ldots, i$ due to the skew-symmetry of $S_{N_{i+1}}$ and ${q^{(j_1,i+1)}}^\top S_{N_{i+1}} q^{(j_2,i+1)} = 0$ for $j_1, j_2 = 1, \ldots, i$ due to the orthogonality conditions in the previous iteration, the product $Q^\top_{(i,i+1)} S_{N_{i+1}} Q_{(i,i+1)}$ is zero. Thus in order to achieve simultaneous orthogonality with respect to $Q_{(i,i+1)}$ and $S_{N_{i+1}} Q_{(i,i+1)}$ we will use the following expression instead of (4.3):

$$\begin{aligned} \Psi_{[i+1,i]} q^{(i+1,i+1)} = {} & v^{(i+1)} - Q_{(i,i+1)} Q_{(i,i+1)}^\top v^{(i+1)} \\ & - S_{N_{i+1}} Q_{(i,i+1)} \left(\left(S_{N_{i+1}} Q_{(i,i+1)}\right)^\top S_{N_{i+1}} Q_{(i,i+1)} \right)^{-1} \left(S_{N_{i+1}} Q_{(i,i+1)}\right)^\top v^{(i+1)}, \end{aligned} \tag{4.4}$$

with $\Psi_{[i+1,i]}$ a normalisation factor such that $\|q^{(i+1,i+1)}\|_2 = 1$. By using this expression, the resulting Krylov subspace satisfies Proposition 3.4 up to machine precision, even in the presence of rounding errors.

Remark 4.2. Note that the left upper blocks of the matrices $S^0_{N_{i+1}}$, $S^{-k}_{N_{i+1}}$ and $S^k_{N_{i+1}}$ are equal to $S^0_{N_i}$, $S^{-k}_{N_i}$ and $S^k_{N_i}$, respectively. Furthermore, note that these matrices only depend on the delays and not on the system matrices. This observation implies for instance that in the single delay case ($K = 1$), where the delay can be rescaled to one by the substitution $s \leftarrow \tau_1 s$, these matrices can be precomputed.

5. Numerical implementation for non-zero shift. In this subsection, it will be assumed that the shift σ is purely imaginary, i.e., $\sigma = \jmath\omega$. A similar result can be obtained when the shift σ is purely real, as will be explained in Remark 5.2.

Similar to the case $\sigma = 0$ we will derive a compact representation for the functions and operations in Algorithm 3.1 by choosing an appropriate structure for the initial function ϕ_1. More specifically, if

$$\phi(\theta) = P_N(\theta)e^{\jmath\omega\theta} + \overline{P_N(\theta)}e^{-\jmath\omega\theta} \text{ for } \theta \in [-\tau_K, \tau_K], \tag{5.1}$$

with $P_N(\cdot) : [-\tau_K, \tau_K] \mapsto \mathbb{C}^{2n}$ an arbitrary complex vector-valued polynomial of degree N and $\overline{P_N(\theta)}$ the complex conjugate of $P_N(\theta)$, then the function $\mathcal{R}_\sigma^{-1}\phi$ is given by

$$\left(\mathcal{R}_\sigma^{-1}\phi\right)(\theta) = Q_{N+1}(\theta)e^{\jmath\omega\theta} + \overline{Q_{N+1}(\theta)}e^{-\jmath\omega\theta} \text{ for } \theta \in [-\tau_K, \tau_K], \tag{5.2}$$

in which $Q_{N+1}(\cdot)$ is a complex vector-valued polynomial of degree $N+1$. Using this representation, the extension step in Algorithm 3.1 can thus again be expressed in terms of operations on the vector-valued coefficients of these polynomials. This is formalized in the following theorem.

THEOREM 5.1. *If ϕ_i is given by*

$$\phi_i(\theta) := \left(\sum_{l=0}^{N_i} q_l^{(i)}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{\jmath\omega\theta} + \left(\sum_{l=0}^{N_i} \overline{q_l^{(i)}}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{-\jmath\omega\theta} \tag{5.3}$$

then extension step can be preformed using basic linear algebra operations:

$$\varphi_{i+1}(\theta) = \left(\sum_{l=0}^{N_i+1} v_l^{(i+1)}\, T_l\left(\tfrac{\theta}{\tau_K}\right)\right) e^{\jmath\omega\theta} + \left(\sum_{l=0}^{N_i+1} \overline{v_l^{(i+1)}}\, T_l\left(\tfrac{\theta}{\tau_K}\right)\right) e^{-\jmath\omega\theta}, \tag{5.4}$$

in which the coefficient vectors $v_1^{(i+1)}, \ldots, v_{N_i+1}^{(i+1)}$ are given by

$$\begin{aligned}\begin{bmatrix} v_1^{(i+1)} & \ldots & v_{N_i+1}^{(i+1)} \end{bmatrix} &\left(2\jmath\omega\tau_K \begin{bmatrix} \frac{1}{4} & & & & \\ & \frac{1}{6} & & & \\ \frac{-1}{4} & & \frac{1}{8} & & \\ & \ddots & & \ddots & \\ & & \frac{-1}{2N_i} & & \frac{1}{2(N_i+2)} \end{bmatrix} + \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \right) \\ &= \begin{bmatrix} q_0^{(i)} & \ldots & q_{N_i}^{(i)} \end{bmatrix} \tau_K^2 \begin{bmatrix} \frac{1}{4} & & & & & & \\ & \frac{1}{24} & & & & & \\ \frac{-1}{6} & & \frac{1}{48} & & & & \\ & \frac{-1}{16} & & \frac{1}{80} & & & \\ \frac{1}{24} & & \frac{-1}{30} & & \frac{1}{120} & & \\ & \ddots & & \ddots & & \ddots & \\ & & \frac{1}{4(N_i-1)(N_i-2)} & & \frac{-1}{2(N_i-1)(N_i+1)} & & \frac{1}{4(N_i+1)(N_i+2)} \end{bmatrix}\end{aligned} \tag{5.5}$$

and $v_0^{(i+1)}$ *can be found by solving the following system*

$$\jmath\omega\, M(\jmath\omega)\, v_0^{(i+1)} = (H_0 - \jmath\omega I_{2n}) \sum_{l=1}^{N_i+1} \left(\Re\Big(v_l^{(i+1)} \Big) \frac{T_l'(0)}{\tau_K} + \jmath\omega v_l^{(i+1)} T_l(0) \right) + \sum_{k=1}^{K} \Big[H_k \sum_{l=1}^{N_i+1} \left(\Re\big(v_l^{(i+1)} e^{\jmath\omega\tau_k} \big) \frac{T_l'\left(\frac{\tau_k}{\tau_K}\right)}{\tau_K} + \jmath\omega v_l^{(i+1)} T_l\left(\frac{\tau_k}{\tau_K}\right) e^{\jmath\omega\tau_k} \right) + H_{-k} \sum_{l=1}^{N_i+1} \left(\Re\big(v_l^{(i+1)} e^{-\jmath\omega\tau_k} \big) \frac{T_l'\left(-\frac{\tau_k}{\tau_K}\right)}{\tau_K} + \jmath\omega v_l^{(i+1)} T_l\left(-\frac{\tau_k}{\tau_K}\right) e^{-\jmath\omega\tau_k} \right) \Big] - \sum_{l=0}^{N_i} \Re\big(q_l^{(i)} \big) T_l(0),$$

with $M(\cdot)$ *the characteristic matrix as defined in* (1.2).

Proof. For more details on deriving this result, see Appendix C. ∎

Remark 5.2. When the shift σ is purely real, similar results can be obtained. More specifically, if ϕ is given by

$$\phi(\theta) = P_N^+(\theta) e^{\sigma\theta} + P_N^-(\theta) e^{-\sigma\theta}$$

with $P_N^+(\cdot)$ and $P_N^-(\cdot)$ arbitrary real vector-valued polynomials of degree N, then $\mathcal{R}_\sigma^{-1}\phi$ is given by

$$\big(\mathcal{R}_\sigma^{-1}\phi\big)(\theta) = Q_{N+1}^+(\theta) e^{\sigma\theta} + Q_{N+1}^-(\theta) e^{-\sigma\theta}$$

with $Q_{N+1}^+(\cdot)$ and $Q_{N+1}^-(\cdot)$ real vector-valued polynomials of degree $N+1$.

From a theoretical point of view, the result obtained above is appealing: similarly to the case $\sigma = 0$ one can operate on the coefficients of polynomials of growing degree. Furthermore, due to the lower triangular structure of the matrix at the left side in (5.5), the main computational cost of the extension step consists of solving a system in $M(\jmath\omega)$ which is of dimension $2n$ and whose factorisation needs to be computed only once. However, there are three difficulties that render this approach unsuited in practice.

1. Firstly, although the functions generated by Algorithm 3.1 are uniquely defined by the starting function ϕ_1 and can be uniquely decomposed as in (5.1) when such a structure is imposed on ϕ_1, the decomposition in terms of $e^{\jmath\omega\theta}$ and $e^{-\jmath\omega\theta}$ is not uniquely defined on X. For example, let $f \in X$ and consider the decomposition $f(\theta) = P(\theta)e^{\jmath\omega\theta} + \overline{P(\theta)}e^{-\jmath\omega\theta}$, then $R(\theta)e^{\jmath\omega\theta} + \overline{R(\theta)}e^{-\jmath\omega\theta}$ with $R(\theta) = P(\theta) + \sin(\omega\theta) + \jmath\cos(\omega\theta)$ is another decomposition for f. This non-uniqueness causes problems when working in finite precision. For example, the range of the polynomial $P_{N_i}(\cdot)$ can grow large, while the range of the overall function ϕ_i remains small. Another related numerical issue that we observed, the condition number of the matrix at the left-hand side of (5.5) grows large as the number of iterations increases.
2. Secondly, evaluating inner product (3.8) is now less trivial. An alternative would be: rather than using the coefficients of the Chebyshev expansion, use the coefficients of the polynomial (expressed in the scaled Chebyshev basis) that interpolates the function in a number of scaled Chebyshev points. If the number of interpolation points is sufficiently large, then the coefficients of this interpolating polynomial are a good approximation for the coefficients in the Chebyshev expansion [23]. The coefficients of this interpolating polynomial can be obtained using the `(i)fft`-transformation at a cost of

$\mathcal{O}\Big(2n\, N_{points} \log(N_{points})\Big)$ with N_{points} the number of discretisation points [1].

3. Finally, when using the representation (5.1), the simultaneous orthogonalisation with respect to the inner product (3.8) and bilinear form $\mathbf{B}(\cdot, J\cdot)$ typically destroys the structure of ϕ_{i+1}.

The three difficulties mentioned above render the representation used in Theorem 5.1 unattractive in practice. Next, we therefore introduce a different approach to implement Algorithm 3.1 for $\sigma \neq 0$ using finite-dimensional operations. Similarly as in Chebfun [23], we approximate the functions in X using polynomial approximations whose degree is adaptively chosen such that they match the original functions up to machine precision. More specifically, a function ψ can be approximated by the unique interpolating polynomial $\widehat{\psi}$ of degree N_d that interpolates ψ in the extreme points of the $N_d{}^{\text{th}}$ Chebyshev polynomial of the first kind rescaled to the interval $[-\tau_K, \tau_K]$:

(5.6)
$$\widehat{\psi}(\theta) = \sum_{l=0}^{N_d} c_l T_l\Big(\tfrac{\theta}{\tau_K}\Big) \text{ such that } \widehat{\psi}(\theta_l) = \psi(\theta_l) \text{ with } \theta_l = \tau_K \cos\big(\tfrac{l\pi}{N_d}\big) \text{ for } l = 0, \ldots, N_d,$$

with N_d chosen such that the approximation error is sufficiently small. As mentioned before, such an interpolating polynomial can be computed efficiently using the `(i)fft`-transformation.

Previously, by invoking Theorems 4.1 and 5.1, $\mathcal{R}_\sigma^{-1}\phi_i$ could be computed efficiently due to the chosen structure of ϕ_1 (which was preserved in ϕ_i). More specifically, for large n, the main computation cost of the extension step consisted of solving a system involving a particular instance of the characteristic matrix (1.2). However, for arbitrary (non-structured) functions ϕ_i computing $\mathcal{R}_\sigma^{-1}\phi_i$ involves solving an ordinary differential equation subjected to conditions (3.1)-(3.2). A finite-dimensional approximate solution for this differential equation can be obtained using spectral discretisation. This would however involve solving a system of dimension $2nN_d$ with N_d a sufficiently large number of discretisation points. To avoid having to solve a system with such dimensions, we will use the approach depicted in Figure 1, which is based on explicitly using the facotriazation $\mathcal{R}_\sigma^{-1} = (\mathcal{H} + \sigma\mathcal{I}_X)^{-1}(\mathcal{H} - \sigma\mathcal{I}_X)^{-1}$. First, $\phi_i(\cdot)$ is approximated by the function $\chi_i(\cdot)e^{\sigma\cdot}$ in which $\chi_i(\cdot)$ is the interpolating polynomial of form (5.6) of $\phi_i(\cdot)e^{-\sigma\cdot}$. Next one can solve $(\mathcal{H} - \sigma\mathcal{I}_X)^{-1}\chi_i(\cdot)e^{\sigma\cdot}$ analytically. More specifically, $(\mathcal{H} - \sigma\mathcal{I}_X)^{-1}\chi_i(\cdot)e^{\sigma\cdot}$ is equal to $\xi_i(\cdot)e^{\sigma\cdot}$ with $\xi_i(\cdot)$ a vector-valued polynomial whose degree is equal to one plus the degree of $\chi_i(\cdot)$. Using the notation

$$\chi_i(\theta) = \sum_{l=0}^{N_1} a_l T_l\left(\frac{\theta}{\tau_K}\right) \text{ and } \xi_i(\theta) = \sum_{l=0}^{N_1+1} b_l T_l\left(\frac{\theta}{\tau_K}\right)$$

the coefficients of $\xi_i(\cdot)$ are given by

$$\begin{bmatrix} b_1 & \ldots & b_{N_1+1} \end{bmatrix} = \begin{bmatrix} a_0 & \ldots & a_{N_1} \end{bmatrix} \tau_K \begin{bmatrix} 1 & & & \\ & \frac{1}{4} & & \\ \frac{-1}{2} & & \frac{1}{6} & \\ & \ddots & & \ddots \\ & & \frac{-1}{2(N_1-1)} & \frac{1}{2(N_1+1)} \end{bmatrix} \tag{5.7}$$

and

$$M(\sigma)\ b_0 = -\chi_i(0) + (H_0 - \sigma I_n) \sum_{l=1}^{N_1+1} b_l T_l(0) + \sum_{k=1}^{K} \left(H_k e^{\sigma\tau_k} \sum_{l=1}^{N_1+1} b_l T_l\left(\frac{\tau_k}{\tau_K}\right) + H_{-k} e^{-\sigma\tau_k} \sum_{l=1}^{N_1+1} b_l T_l\left(-\frac{\tau_k}{\tau_K}\right) \right). \tag{5.8}$$

with $M(\cdot)$ the characteristic matrix as defined in (1.2). As before, one now only has to solve a system of dimension $2n$. Notice that these expressions are similar to those in the extension step of the standard infinite Arnoldi method from [13]. Subsequently, we approximate $\xi_i(\cdot)e^{\sigma\cdot}$ by $\zeta_i(\cdot)e^{-\sigma\cdot}$ with $\zeta_i(\cdot)$ the interpolating polynomial of form (5.6) of $\xi_i(\cdot)e^{2\sigma\cdot}$. In step **(IV)** we compute $(\mathcal{H} + \sigma\mathcal{I}_X)^{-1}\zeta_i(\cdot)e^{-\sigma\cdot}$ for which we can again derive an analytical expression. More specifically, $(\mathcal{H} + \sigma\mathcal{I}_X)^{-1}\zeta_i(\cdot)e^{-\sigma\cdot}$ is equal to $\Upsilon_i(\cdot)e^{-\sigma\cdot}$ with $\Upsilon_i(\theta) = \sum_{l=0}^{N_2+1} d_l T_l\left(\frac{\theta}{\tau_K}\right)$ a vector-valued polynomial whose coefficients follow from those of $\zeta_i(\theta) = \sum_{l=0}^{N_2} c_l T_l\left(\frac{\theta}{\tau_K}\right)$ by the following relations

$$\begin{bmatrix} d_1 & \dots & d_{N_2+1} \end{bmatrix} = \begin{bmatrix} c_0 & \dots & c_{N_2} \end{bmatrix} \tau_K \begin{bmatrix} 1 & & & \\ & \frac{1}{4} & & \\ \frac{-1}{2} & & \frac{1}{6} & \\ & \ddots & & \ddots \\ & & \frac{-1}{2(N_{d,2}-1)} & \frac{1}{2(N_{d,2}+1)} \end{bmatrix} \tag{5.9}$$

and

$$M(-\sigma)\ d_0 = -\zeta_i(0) + (H_0 + \sigma I_n) \sum_{l=1}^{N_2+1} d_l T_l(0) + \sum_{k=1}^{K} \left(H_k e^{-\sigma\tau_k} \sum_{l=1}^{N_2+1} d_l T_l\left(\frac{\tau_k}{\tau_K}\right) + H_{-k} e^{\sigma\tau_k} \sum_{l=1}^{N_2+1} d_l T_l\left(-\frac{\tau_k}{\tau_K}\right) \right). \tag{5.10}$$

Finally, $\widehat{\varphi}_{i+1}$ is obtained by computing the interpolating polynomial of form (5.6) of $\Upsilon_i(\cdot)e^{-\sigma\cdot}$. If the interpolating polynomials in steps **(I)**, **(III)** and **(V)** are computed up to machine precision, the resulting $\widehat{\varphi}_{i+1}$ is an accurate approximation for $\varphi_{i+1} = \mathcal{R}_\sigma^{-1}\phi_i$.

Using this polynomial representation expressed in the Chebyshev basis, also the orthogonalisation step can be implemented with finite-dimensional operations. More specifically, by using the representation of Figure 1 for $\phi_1, \dots, \phi_i$ and $\hat{\varphi}_{i+1}$, we can use an orthogonalisation procedure similar to that in (4.4) to ensure orthogonality with respect to both the inner product (3.8) and bilinear form $\mathbf{B}(\cdot, J\cdot)$: instead of padding the vectors to length $2n(N_{i+1}+1)$, the vectors now must be padded to length $2n\Big(\max\{\max_{j=1,..,i}\{N_{\phi_j}\}, N_{\hat{\varphi}_{i+1}}\} + 1\Big)$.

A disadvantage of the approach introduced above is that, in contrast to Theorems 4.1 and 5.1, the degree of the polynomial $\hat{\varphi}_{i+1}$ is not known beforehand and therefore might grow large. However, the example in Subsection 6.2 shows that if σ is sufficiently small, the degree of these functions only grows slowly. For large σ this

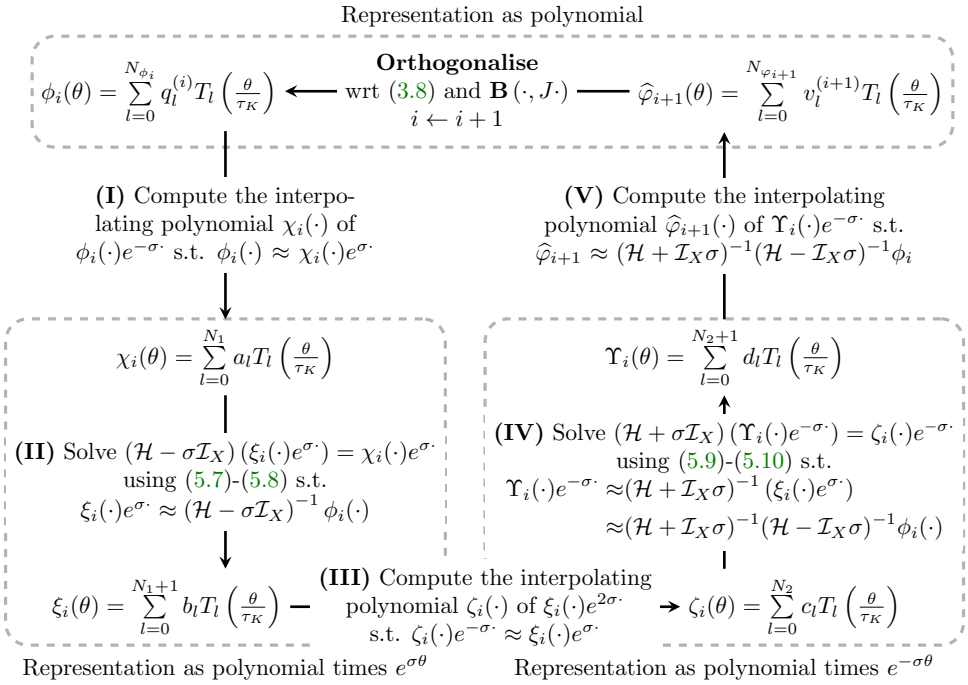


FIG. 1. *Structured representation of the used numerical implementation of Algorithm 3.1 for* $\sigma \neq 0$.

is often no longer the case as the sought for eigenfunctions are either highly oscillator (large imaginary component) or fast growing exponentials (large real component). A polynomial representation inherently requires a high degree to accurately approximate such functions.

6. Numerical illustration. In this section we consider two examples. The first example illustrates the importance of the chosen shift-invert transformation and of the explicit orthogonalisation of basis functions of the Krylov subspace with respect to $\mathbf{B}(\cdot, J\cdot)$ for computing purely imaginary eigenvalues of (1.1) in the presence of rounding errors. The second example shows the convergence behavior of the presented method for a large-scale example. Both experiments are carried out in MATLAB R2022a (version 9.12.0.1884302) on a Dell Latitude 7490 with a Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz eight core processor with 16GB of RAM running Ubuntu 20.04.4 LTS. The code for these experiments is available from https://twr.cs.kuleuven.be/research/software/delay-control/SPSIIA/index.html.

6.1. Example 1. Consider the following characteristic matrix of form (1.2):

$$\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 10 & 0.1 \\ c_0 & -10 \end{bmatrix} - \begin{bmatrix} a_1 & 0 \\ 0 & 0 \end{bmatrix} e^{-\lambda} - \begin{bmatrix} 0 & 0 \\ 0 & -a_1 \end{bmatrix} e^{\lambda}, \tag{6.1}$$

with $a_1 = (3\pi^2/4)/(20+\pi)$ and $c_0 = -1000 - 10a_1^2 - 10a_1\pi - \frac{5\pi^2}{2}$, which has purely imaginary eigenvalues at $\jmath\pi$ and $\frac{\jmath\pi}{2}$. Tables 1 and 2 compare three methods for computing the purely imaginary eigenvalues of the associated NEP using $\sigma = 0$ and $\sigma = \jmath\frac{3\pi}{4}$, respectively. These three methods are:

1. applying the infinite Arnoldi method to $\mathcal{H}^{-1}$ and $(\mathcal{H}-\sigma\mathcal{I}_X)^{-1}$, respectively, using a modification of the algorithm from [13] to deal with both positive and negative delays;
2. applying the infinite Arnoldi method to $\mathcal{R}_0^{-1}$ and $\mathcal{R}_\sigma^{-1}$ using the method presented in Section 4 and Section 5 (Figure 1), respectively, but without explicit orthogonalisation with respect to $\mathbf{B}(\cdot, J\cdot)$ (i.e., using (4.3) instead of (4.4));
3. applying the infinite Arnoldi method to $\mathcal{R}_0^{-1}$ and $\mathcal{R}_\sigma^{-1}$ using the method presented in Section 4 and Section 5 (Figure 1), respectively, with explicit orthogonalisation with respect to $\mathbf{B}(\cdot, J\cdot)$.

As initial function the constant function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top$ is used in all cases but one. When applying the infinite Arnoldi method to $(\mathcal{H}-\sigma\mathcal{I}_X)^{-1}$, the function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top e^{\sigma\theta}$ is employed, to be able to use the compact representation from [13] for the functions in the Krylov subspace.

For the first approach, we observe from Tables 1 and 2 that the obtained approximations have a significant non-zero real part, as this method does not explicitly take into account the symmetry of the spectrum with respect to the imaginary axis. For the second approach, we see that each purely imaginary eigenvalue of the NEP appears twice in the table. Recall that the eigenvalues $\pm\jmath\omega$ of $\mathcal{H}$ are mapped to the real eigenvalue $\mu = \frac{1}{-\omega^2-\sigma^2}$ of $\mathcal{R}_\sigma^{-1}$ with multiplicity 2. The eigenspace associated with μ is given by $\mathcal{L} = \operatorname{span}\{\varphi_+, \varphi_-\}$, with φ_+ and φ_- the eigenfunctions of $\mathcal{H}$ associated with $\jmath\omega$ and $-\jmath\omega$, respectively. Although in exact arithmetic we expect that only one linear independent component of $\mathcal{L}$ will be approximated by the Krylov subspace (a consequence of Theorem 3.5), this is typically no longer the case when working in finite precision as orthogonality with respect to $\mathbf{B}(\cdot, J\cdot)$ is lost due to rounding error. As a consequence, the Krylov subspace will eventually also approximate a second linear independent component in the eigenspace $\mathcal{L}$, meaning the double eigenvalue μ of $\mathcal{R}_\sigma^{-1}$ will be approximated by a pair of Ritz values. One either gets two nearby real Ritz values or a pair of complex conjugate Ritz values. In the latter case, after re-transformation the obtained approximations for purely imaginary eigenvalues of $\mathcal{H}$ thus have a small real part. When, however, explicit orthogonalisation against $\mathbf{B}(\cdot, J\cdot)$ is preformed (approach 3), the Krylov subspace satisfies the condition in Proposition 3.4 up to machine precision. As to be expected, the approximations for the purely imaginary eigenvalues are now purely imaginary.

Finally, from Table 2, we observe that when using the shift-invert transformation $(\mathcal{H}-\sigma\mathcal{I}_X)^{-1}$, the eigenvalues $\jmath\omega$ and $-\jmath\omega$ are approximated separately. In contrast, when using $\mathcal{R}_\sigma^{-1}$ this pair is really approximated as a pair.

TABLE 1

Obtained approximations after 21 iterations for the purely imaginary eigenvalues of (6.1) using approaches 1-3 for $\sigma = 0$ with the constant function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top$ as initial function.

Approach 1	Approach 2	Approach 3
$-2.525\times10^{-11} - \jmath 1.570796326838293$	$1.444\times10^{-12} - \jmath 1.570796326794669$	$\jmath 1.570796326750096$
$-2.525\times10^{-11} + \jmath 1.570796326838293$	$1.444\times10^{-12} + \jmath 1.570796326794669$	$-\jmath 1.570796326750096$
$-5.141\times10^{-9} - \jmath 3.141592653006948$	$-1.444\times10^{-12} - \jmath 1.570796326794669$	$\jmath 3.141592653831962$
$-5.141\times10^{-9} + \jmath 3.141592653006948$	$-1.444\times10^{-12} + \jmath 1.570796326794669$	$-\jmath 3.141592653831962$
	$-\jmath 3.141592653591730$	
	$\jmath 3.141592653591730$	
	$-\jmath 3.141592653286507$	
	$\jmath 3.141592653286507$	

TABLE 2

Obtained approximations after 21 iterations for the purely imaginary eigenvalues of (6.1) using approaches 1-3 for $\sigma = \jmath\frac{3\pi}{4}$. For approach 1 the initial function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top e^{\sigma\theta}$ is used, while for the other approaches the initial function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top$ is used.

Approach 1	Approach 2	Approach 3
$3.005 \times 10^{-13} + \jmath 1.570796326792780$	$8.260 \times 10^{-13} - \jmath 1.570796326794206$	$-\jmath 1.570796326657866$
$-3.238 \times 10^{-13} + \jmath 3.141592653590773$	$8.260 \times 10^{-13} + \jmath 1.570796326794206$	$\jmath 1.570796326657866$
$1.570 \times 10^{-6} - \jmath 1.570708037662637$	$-8.260 \times 10^{-13} - \jmath 1.570796326794206$	$-\jmath 3.141592653619107$
$-5.788 \times 10^{-4} - \jmath 3.140992633763653$	$-8.260 \times 10^{-12} + \jmath 1.570796326794206$	$\jmath 3.141592653619107$
	$-\jmath 3.141592653591559$	
	$\jmath 3.141592653591559$	
	$-\jmath 3.141592653589887$	
	$\jmath 3.141592653589887$	

6.2. Example 2. For the second example we consider the following dynamical system from [20], which describes a heated rod that is cooled using delayed feedback. The evolution of the temperature in the rod, v, is governed by the partial differential equation

$$\frac{\partial v(x,t)}{\partial t} = \frac{\partial^2 v(x,t)}{\partial x^2} + a_0(x)v(x,t) + a_1(x)v(\pi - x, t-1) \text{ for } x \in [0,\pi],$$

with $a_0(x) = -2\sin(x)$, $a_1(x) = 2\sin(x)$ and $v(0,t) = v(\pi,t) = 0$. Discretising this partial differential equation in the space coordinate x results in a system of delay-differential equations of dimension n. To obtain a dynamical system of form (1.5), we define a performance output matrix $C = \frac{1}{n}\begin{bmatrix}1 & \dots & 1\end{bmatrix}$ (which gives the average temperature of the rod) and a performance input matrix $B = C^\top$. Choosing $n = 1000$ and plotting the magnitude of the transfer matrix in function of ω, one observes that the magnitude of the transfer function is equal to 0.00018 for ω approximately equal to 2.009437, 3.790888 and 5.571120. It now follows from the discussion in the introduction that for $\gamma = 0.00018$, the NEP associated with (1.6) must have purely imaginary eigenvalues around $\jmath 2.009437$, $\jmath 3.790888$ and $\jmath 5.571120$.

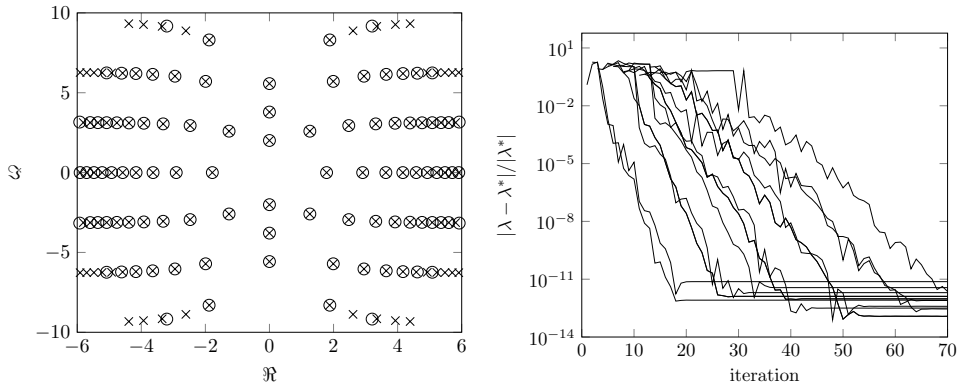
To verify this, we first use the method presented in Section 4 to compute the eigenvalues of (1.6) near the shift $\sigma = 0$. Figure 2 shows the obtained approximations for the eigenvalues in the region $[-6,6] \times \jmath[-10,10]$ after 70 iterations and their convergence behavior. One observes that the approximations converge quickly to the eigenvalues of (1.6) close to the origin.

Next, we apply the method from Section 5, with implementation sketched in Figure 1, for $\sigma = \jmath 4.5$. Figure 3 shows the eigenvalues and the obtained approximations near this shift after 70 iterations. We have again fast convergence to the eigenvalues near the shift.

Finally, recall that the degree of the polynomial $\hat{\varphi}_{i+1}$, defined in Figure 1, is not known beforehand. Figure 4 therefore shows the evolution of the degree of this polynomial with respect to the iteration number. We observe that the degree of these polynomials grows slowly with respect to the number of iterations.

7. Conclusions and outlook. In this work we presented an iterative method to approximate the eigenvalues of NEPs of form (1.1) closest to a given shift σ while preserving the symmetries of the spectrum. The presented work can thus be seen as a generalization of the results from [16, 15] to a class of NEPs.

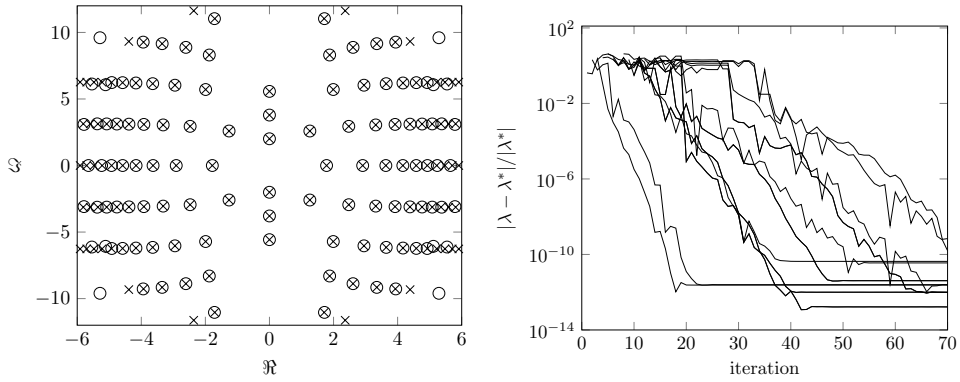
To derive this method, the equivalence between the considered NEP and a linear but infinite-dimensional eigenvalue problem was used. Based on this equivalence, we



(a) Eigenvalues (crosses) and their approximations (circles) near $\sigma = 0$.

(b) Relative forward error

FIG. 2. *Approximated eigenvalues near the shift $\sigma = 0$ after 70 iterations and their convergence behavior for Example 2 obtained using the method from Section 4.*



(a) Eigenvalues (crosses) and their approximations (circles) near $\sigma = \jmath 4.5$.

(b) Relative forward error

FIG. 3. *Approximated eigenvalues near the shift $\sigma = \jmath 4.5$ after 70 iterations and their convergence behavior for Example 2 obtained using the method from Section 5, as depicted in Figure 1.*

introduced a shift-invert transformation that preserves the Hamiltonian structure of the spectrum. Next, the ideas behind the infinite Arnoldi method from [13], which operates on functions rather than on vectors, were applied to this transformed eigenvalue problem to construct a Krylov subspace. It was then shown that this subspace is orthogonal with respect to the bilinear functional $\mathbf{B}(\cdot, J\cdot)$. This result was subsequently used to demonstrate that simple purely imaginary eigenvalues of (1.1) close to the shift are generally approximated by purely imaginary eigenvalues. Although this method was initially defined on function spaces, Sections 4 and 5 showed how it can be implemented using finite-dimensional linear algebra operations. The performance of these numerical algorithms was finally verified using two numerical experiments in Section 6.

To conclude this paper we give some directions for future research. Firstly, as mentioned in the introduction, the presented algorithm can be used as a building block for algorithms that compute the $\mathcal{H}_\infty$-norm of time-delay systems. Secondly, a

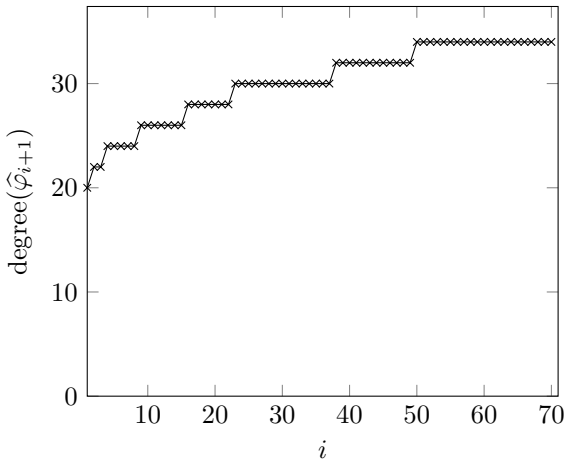


FIG. 4. *Evolution of the degree of $\widehat{\varphi}_{i+1}$, as defined in Figure 1, as function of the iteration number i.*

more extensive study on the effect of the chosen initial function and the chosen inner product on the convergence behavior of the method is necessary. Thirdly, for large σ the representation in (5.1) has as advantage that it might require polynomials of a lower degree to approximate the eigenfunctions of nearby eigenvalues in comparison to purely polynomial approximations. The functions $e^{\jmath\omega\theta}$ and $e^{-\jmath\omega\theta}$ act in this case as carrier functions. However, as mentioned before, it is not yet clear how this representation can be used in practice because of problems with numerical stability. Finally, for the infinite Arnoldi method in [13], the convergence behavior of the method can be related to the approximation error of a Padé-approximation of the DEVP with growing degree on the one hand and the convergence behavior of the finite-dimensional Arnoldi method on the other hand. Such a connection is yet to be established for the method presented here.

Appendix A. Proof that $\mathbf{B}(\cdot, J\cdot)$ is anti-symmetric.

LEMMA A.1. *The bilinear form $\mathbf{B}(\cdot, J\cdot)$ is anti-symmetric, this means that the equality $\mathbf{B}(\varphi, J\psi) = -\mathbf{B}(\psi, J\varphi)$ holds.*

Proof.

$$\begin{aligned}
\mathbf{B}(\varphi, J\psi) &= -\Big[\psi(0)^\top J\varphi(0) + \sum_{k=1}^{K}\Big(\int_0^{\tau_k}\psi(\theta)^\top JH_{-k}\varphi(\theta-\tau_k)\,\mathrm{d}\theta - \int_0^{\tau_k}\psi(\theta-\tau_k)^\top JH_k\varphi(\theta)\,\mathrm{d}\theta\Big)\Big]\\
&= -\Big[\varphi(0)^\top J^\top\psi(0) + \sum_{k=1}^{K}\Big(\int_0^{\tau_k}\varphi(\theta-\tau_k)^\top(JH_{-k})^\top\psi(\theta)\,\mathrm{d}\theta - \int_0^{\tau_k}\varphi(\theta)^\top(JH_k)^\top\psi(\theta-\tau_k)\,\mathrm{d}\theta\Big)\Big]\\
&= \Big[\varphi(0)^\top J\psi(0) + \sum_{k=1}^{K}\Big(\int_0^{\tau_k}\varphi(\theta)^\top JH_{-k}\psi(\theta-\tau_k)\,\mathrm{d}\theta - \int_0^{\tau_k}\varphi(\theta-\tau_k)^\top JH_k\psi(\theta)\,\mathrm{d}\theta\Big)\Big]\\
&= -\mathbf{B}(\psi, J\varphi)
\end{aligned}$$

∎

Appendix B. Derivation of the extension step for $\sigma = 0$. Let $\phi_i(\theta)$ and $\varphi_{i+1}(\theta)$ be as defined in Theorem 4.1, then the equality $\mathcal{R}_0\varphi_{i+1} = \phi_i$ becomes

$$\Big(\mathcal{R}_0\varphi_{i+1}(\theta) = \Big)\ \sum_{l=2}^{N_i+2}\frac{v_l^{(i+1)}}{(\tau_K)^2}T_l''\left(\frac{\theta}{\tau_K}\right) = \sum_{l=0}^{N_i}q_l^{(i)}T_l\left(\frac{\theta}{\tau_K}\right)\Big(=\phi_i(\theta)\Big).$$

Expression (4.1) now follows by noting that

$$
\begin{aligned}
T_0(t) =& T_2''(t)/4, \\
T_1(t) =& T_3''(t)/24, \\
T_2(t) =& T_4''(t)/48 - T_2''(t)/6 \text{ and} \\
T_l(t) =& \frac{T_{l+2}''(t)}{4(l+1)(l+2)} - \frac{T_l''(t)}{2(l+1)(l-1)} + \frac{T_{l-2}''(t)}{4(l-1)(l-2)} \text{ for } l \geq 3.
\end{aligned} \tag{B.1}
$$

The expressions for $v_1^{(i+1)}$ follows directly from (3.2) and using this result $v_0^{(i+1)}$ can be computed from (3.1).

Appendix C. Derivation of the extension step for $\sigma = \jmath\omega$. Let ϕ_i and φ_{i+1} be as defined in (5.3) and (5.4), then the equality $\mathcal{R}_\sigma \varphi_{i+1} = \phi_i$ becomes

$$
\begin{aligned}
&\left[\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1} v_l^{(i+1)} T_l'\left(\frac{\theta}{\tau_K}\right) + \sum_{l=2}^{N_i+1} \frac{v_l^{(i+1)}}{\tau_K^2} T_l''\left(\frac{\theta}{\tau_K}\right)\right] e^{\jmath\omega\theta} + \\
&\qquad \left[-\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1} \overline{v_l^{(i+1)}} T_l'\left(\frac{\theta}{\tau_K}\right) + \sum_{l=2}^{N_i+1} \frac{\overline{v_l^{(i+1)}}}{\tau_K^2} T_l''\left(\frac{\theta}{\tau_K}\right)\right] e^{-\jmath\omega\theta} \\
&\qquad\qquad = \sum_{l=0}^{N_i} q_l^{(i)} T_l\left(\frac{\theta}{\tau_K}\right) e^{\jmath\omega\theta} + \sum_{l=0}^{N_i} \overline{q_l^{(i)}} T_l\left(\frac{\theta}{\tau_K}\right) e^{-\jmath\omega\theta}.
\end{aligned}
$$

Matching the terms associated with $e^{\jmath\omega\theta}$ and $e^{-\jmath\omega\theta}$, gives the equality

$$
\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1} v_l^{(i+1)} T_l'\left(\frac{\theta}{\tau_K}\right) + \sum_{l=2}^{N_i+1} \frac{v_l^{(i+1)}}{\tau_K^2} T_l''\left(\frac{\theta}{\tau_K}\right) = q_l^{(i)} T_l\left(\frac{\theta}{\tau_K}\right).
$$

Using (B.1) in combination with

$$
\begin{aligned}
T_1'(t) &= T_2''(t)/4 \text{ and} \\
T_l'(t) &= \frac{T_{l+1}''(t)}{2(l+1)} - \frac{T_{l-1}''(t)}{2(l-1)} \text{ for } l \geq 2
\end{aligned}
$$

results in (5.5). The expression for $v_0^{(i+1)}$ follows again from (3.1) and (3.2).

Acknowledgments. The first author would like to thank P. Schwerdtner, M. Voigt and V.Mehrmann for the fruitful discussions.
This work was supported by the project C14/17/072 of the KU Leuven Research Council and the project G092721N of the Research Foundation-Flanders (FWO - Vlaanderen).

REFERENCES

[1] N. Ahmed and P. Fisher, *Study of algorithmic properties of Chebyshev coefficients*, International Journal of Computer Mathematics, pp. 307–317, https://doi.org/10.1080/00207167008803043.

[2] P. Benner, D. Kressner, and V. Mehrmann, *Skew-Hamiltonian and Hamiltonian Eigenvalue Problems: Theory, Algorithms and Applications*, in Proceedings of the Conference on Applied Mathematics and Scientific Computing, Springer, Dordrecht, the Netherlands, pp. 3–39, https://doi.org/10.1007/1-4020-3197-1_1.

[3] S. Boyd and V. Balakrishnan, *A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its* L_∞*-norm*, Systems & Control Letters, pp. 1–7, https://doi.org/10.1016/0167-6911(90)90037-U.

[4] N. Bruinsma and M. Steinbuch, *A fast algorithm to compute the* H_∞*-norm of a transfer function matrix*, Systems & Control Letters, pp. 287–293, https://doi.org/10.1016/0167-6911(90)90049-Z.

[5] K. Gu, V. L. Kharitonov, and J. Chen, *Stability of Time-Delay Systems*, Birkhäuser, Boston, MA, https://doi.org/10.1007/978-1-4612-0039-0.

[6] S. Gumussoy and W. Michiels, *Fixed-Order H-Infinity Control for Interconnected Systems Using Delay Differential Algebraic Equations*, SIAM Journal on Control and Optimization, pp. 2212–2238, https://doi.org/10.1137/100816444.

[7] S. Güttel and F. Tisseur, *The nonlinear eigenvalue problem*, Acta Numerica, 26 (2017), pp. 1–94, https://doi.org/10.1017/S0962492917000034.

[8] S. Güttel, R. Van Beeumen, K. Meerbergen, and W. Michiels, *NLEIGS: A class of fully rational Krylov methods for nonlinear eigenvalue problems*, SIAM Journal on Scientific Computing, 36 (2014), pp. A2842–A2864, https://doi.org/10.1137/130935045.

[9] J. Hale, *Theory of functional differential equations*, vol. 3 of Applied Mathematical Sciences, Springer, New York, NY, https://doi.org/10.1007/978-1-4612-9892-2.

[10] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Applied Mathematical Sciences, Springer, New York, NY, https://doi.org/10.1007/978-1-4612-4342-7.

[11] D. Hinrichsen and A. J. Pritchard, *Mathematical Systems Theory I*, Texts in Applied Mathematics, Springer, Berlin, Germany, https://doi.org/10.1007/b137541.

[12] E. Jarlebring, M. Bennedich, G. Mele, E. Ringh, and P. Upadhyaya, *NEP-PACK: A Julia package for nonlinear eigenproblems - v0.2*, arXiv, (2018), https://arxiv.org/abs/1811.09592.

[13] E. Jarlebring, K. Meerbergen, and W. Michiels, *A Krylov Method for the Delay Eigenvalue Problem*, SIAM Journal on Scientific Computing, pp. 3278–3300, https://doi.org/10.1137/10078270X.

[14] E. Jarlebring, W. Michiels, and K. Meerbergen, *A linear eigenvalue algorithm for the nonlinear eigenvalue problem*, Numerische Mathematik, 122 (2012), pp. 169–195, https://doi.org/10.1007/s00211-012-0453-0.

[15] V. Mehrmann, C. Schröder, and V. Simoncini, *An implicitly-restarted Krylov subspace method for real symmetric/skew-symmetric eigenproblems*, Linear Algebra and Its Applications, pp. 4070–4087, https://doi.org/10.1016/j.laa.2009.11.009.

[16] V. Mehrmann and D. Watkins, *Structure-Preserving Methods for Computing Eigenpairs of Large Sparse Skew-Hamiltonian/Hamiltonian Pencils*, SIAM Journal on Scientific Computing, pp. 1905–1925, https://doi.org/10.1137/S1064827500366434.

[17] V. Mehrmann and D. Watkins, *Polynomial eigenvalue problems with Hamiltonian structure*, Electronic Transactions on Numerical Analysis, 13 (2002), pp. 106–118.

[18] V. L. Mehrmann, *The autonomous linear quadratic control problem : theory and numerical solution*, Springer, Berlin, Germany, 1991, https://doi.org/10.1007/BFb0039443.

[19] W. Michiels and S. Gumussoy, *Characterization and Computation of* $\mathcal{H}_\infty$ *Norms for Time-Delay Systems*, SIAM Journal on Matrix Analysis and Applications, pp. 2093–2115, https://doi.org/10.1137/090758751.

[20] W. Michiels, E. Jarlebring, and K. Meerbergen, *Krylov-Based Model Order Reduction of Time-delay Systems*, SIAM Journal on Matrix Analysis and Applications, pp. 1399–1421, https://doi.org/10.1137/100797436.

[21] W. Michiels and S.-I. Niculescu, *Stability, Control, and Computation for Time-Delay Systems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, https://doi.org/10.1137/1.9781611973631.

[22] Y. Saad, *Iterative Methods for Sparse Linear Systems*, Society for Industrial and Applied Mathematics, https://doi.org/10.1137/1.9780898718003.

[23] L. N. Trefethen, *Approximation Theory and Approximation Practice*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013, https://doi.org/10.1137/1.9781611975949.

[24] C. Van Loan, *A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix*, Linear Algebra and its Applications, pp. 233–251, https://doi.org/10.1016/0024-3795(84)90034-X.

[25] H. Voss, *Nonlinear Eigenvalue Problems*, in Handbook of Linear Algebra, L. Hogben, ed., Chapman and Hall/CRC, Boca Raton, FL, 2nd ed., 2014, pp. 115:1–115:24, https://doi.org/10.1201/b16113-70.

[26] K. Zhou and J. C. Doyle, *Essentials of robust control*, Prentice Hall, Upper Saddle River, NJ, 1998.