A structure preserving shift-invert infinite Arnoldi algorithm for a class of delay eigenvalue problems with Hamiltonian symmetry

Pieter Appeltans and Wim Michiels
Department of Computer Science, KU Leuven, Leuven, Belgium
{pieter,wim}.{appeltans,michiels}@kuleuven.be

October 14, 2021

Abstract

In this work we consider a class of non-linear 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. More precisely, we present a method to iteratively approximate the eigenvalues of such non-linear eigenvalue problems closest to a given purely real or imaginary shift, while preserving the symmetries of the spectrum. To this end the presented method exploits the equivalence between the considered non-linear 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 in 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 non-linear 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 this numerical algorithm will subsequently be 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 demonstrated its applicability to a large scale problems.

1 Introduction

In this manuscript, we consider non-linear eigenvalue problems (NLEVPs) of the form

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

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 form:

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

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. The matrix H_0 is Hamiltonian meaning that

$$\left(JH_0\right)^\top = JH_0 \tag{3}$$

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

$$\left(JH_{-k}\right)^\top = JH_k \text{ for } k = 1, \ldots, K, \tag{4}$$

with the matrix J defined as

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

The goal of this paper is to develop a method to accurately compute the eigenvalues of (1) that lie close to the imaginary axis when the dimension of the characteristic matrix is large.

A motivation for considering NLEVPs with a characteristic matrix of form (2), stems from a popular approach to compute the $\mathcal{H}_\infty$-norm, which is an important performance measure in the robust control framework as it can be used to quantify both the input-to-output noise suppression and the distance to instability of the system [10, 25], of dynamical time-delay systems. More specifically, consider the following state-space time-delay system

$$\begin{cases} \dot{x}(t) &= A_0x(t) + \sum_{k=1}^K A_kx(t-\tau_k) + Bw(t) \\ z(t) &= Cx(t) \end{cases} \tag{5}$$

with $x(t) \in \mathbb{R}^n$ the state, $w(t) \in \mathbb{R}^m$ the performance input, $z(t) \in \mathbb{R}^p$ the performance output, $0 < \tau_1 < ... < \tau_K < \infty$ discrete delays and A_0, $\ldots$, A_K, B and C real-valued matrices of appropriate dimensions. If system (5) is exponentially stable, its $\mathcal{H}_\infty$-norm is defined as

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

with $\jmath$ the imaginary unit and $T(\cdot)$ the transfer matrix of the system that describes the system's input-output map in the frequency domain:

$$T(\jmath\omega) := C\left(\jmath\omega I - A_0 - \sum_{k=1}^K A_ke^{-\jmath\omega\tau_k}\right)^{-1} B.$$

To compute this $\mathcal{H}_\infty$-norm for time-delay systems, the methods presented in [6, 18] use a level set approach, which can be seen as an extension of the well-known Boyd-Balakrishnan-Bruinsma-Steinbuch algorithm [3, 4] to time-delay systems. 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. To verify whether this inequality holds in the case of time-delay systems of form (5), the following equivalence from [18, Lemma 2.1] is 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 a solution of the NLEVP associated with the following characteristic matrix,

$$\lambda\begin{bmatrix} I_n & 0 \\ 0 & I_n \end{bmatrix} - \begin{bmatrix} A_0 & \gamma^{-1}BB^\top \\ -\gamma^{-1}C^\top C & -A_0^\top \end{bmatrix} - \sum_{k=1}^K \left(\begin{bmatrix} A_k & 0 \\ 0 & 0 \end{bmatrix} e^{-\lambda\tau_k} + \begin{bmatrix} 0 & 0 \\ 0 & -A_k^\top \end{bmatrix} e^{\lambda\tau_k}\right). \tag{6}$$

Verifying whether $\|T(\cdot)\|_{\mathcal{H}_\infty} \geq \gamma$ is thus equivalent with checking whether the NLEVP associated with (6) has purely imaginary solutions. For this application, it is thus important to be able to accurately compute purely imaginary eigenvalues of the NLEVP associated with (6). Notice that this characteristic matrix fits the structure given in (2).

Next, we will see that the spectrum, i.e., the set of eigenvalues, of the considered NLEVP has some interesting features. Firstly, notice that the considered NLEVP bears some similarities

with a retarded delay eigenvalue problem (RDEVP) [5, 9, 20], but in contrast has both positive and negative delays. It is therefore not surprising that, similar to RDEVPs, non-trivial (at least one of the matrices H_k for $k \neq 0$ is non-zero) NLEVPs of the considered class have infinitely many eigenvalues. However, in contrast to RDEVPs, there does not exist a right half plane that contains only finitely many eigenvalues. Secondly, due to the considered structure of the characteristic matrix and Assumption 1, the considered class of NLEVPs can be seen as an extension of the linear eigenvalue problem

$$(I_{2n}\lambda - H)\, v = 0, \tag{7}$$

with $H \in \mathbb{R}^{2n\times 2n}$ a Hamiltonian matrix, as studied in, among others, [2, 17, 23]. 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 $-\lambda$) [15]. In the next section we will see that the spectrum of the considered NLEVP possesses the same symmetries.

As mentioned before, we are interested in the case for which characteristic matrix (2) is large. For linear, finite dimensional eigenvalue problems, Krylov subspace methods, such as the (shift-invert) Arnoldi method, are well established [21] to iteratively approximate the eigenvalues closest to a given shift when the dimension of the eigenvalue problem is large. Recently, these methods have been generalized to NLEVPs, such as DEVPs, see for example [7, 11, 12, 13, 24]. In the case of DEVPs, the method presented in [11] can be interpreted in two ways: firstly, it can be understood as applying the Arnoldi method to a sufficiently large linearisation of the original DEVP with vectors of increasing length and secondly as applying the Arnoldi method to an associated linear but infinite dimensional operator resulting in a Krylov subspace that is spanned by functions (more specifically polynomials) instead of vectors. This last interpretation gave rise to its name, the infinite Arnoldi method. 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 [17] which preserves the Hamiltonian structure. This approach was subsequently generalized to both matrix pencils and polynomial eigenvalue problems in [14] and [16], respectively. The goal of this paper is to extend this method to the considered class of NLEVP.

The remainder of this work is structured as follows. Section 2 recalls some preliminary results: first some properties of the spectrum of (1) are highlighted, next the most important components of the structure preserving shift-invert Arnoldi method for linear Hamiltonian eigenvalue problems of finite dimension from [17] are reviewed and finally the equivalence of NLEVP (1) with the eigenvalue problem associated with a linear operator acting on an infinite dimensional space is demonstrated. In Section 3, the structure preserving shift-invert Arnoldi method from [17] is generalized to NLEVP (1) using the aforementioned equivalence with a linear, but infinite dimensional eigenvalue problem. This section also presents an algorithmic implementation of the structure preserving shift-invert infinite Arnoldi method that is defined on the infinite dimensional function space associated with the operator. Subsequently Sections 4 and 5 discus how this method can be implemented using finite dimensional operations for $\sigma = 0$ and $\sigma \neq 0$, respectively, with σ the shift employed in the (shift-invert) Arnoldi method. Next, Section 6 applies the resulting numerical implementations on two examples. 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 systems. Finally, Section 7 draws some concluding remarks.

2 Preliminary results

In this section we will recall some important preliminary results. First some properties of NLEVP (1) will be examined. Next, we will briefly review the structure preserving shift-invert Arnoldi

method for linear Hamiltonian eigenvalue problems from [17]. Finally the equivalence between NLEVP associated with (2) and two infinite-dimensional eigenvalue problems, one related to the left eigenspace and one to the right eigenspace, is demonstrated.

2.1 Important properties of the considered NLEVP

Throughout the paper a nonzero column vector $w \in \mathbb{C}^{2n}$ is a left eigenvector of characteristic matrix (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 notation used here is however common for Hamiltonian eigenvalue problems and simplifies the notation in the remainder of this text.

Firstly, we will show that the spectrum of (1) admits the same symmetry as the spectrum of a real-valued Hamiltonian matrix, i.e., symmetry with respect to both the real and imaginary axis. Symmetry with respect to the real axis follows from the fact that all matrices in (2) are real-valued, while symmetry with respect to the origin then follows from the following proposition.

Proposition 1. If λ is an eigenvalue of (2) with associated right eigenvector v_+ and left eigenvector w_+, then $-\lambda$ is also an eigenvalue of (2) and the corresponding right and left eigenvectors are Jw_+ and Jv_+, respectively.

Proof. The pair $(-\lambda, Jw_+)$ is an eigenvalue/right eigenvector-pair of the NLEVP associated with (2) if (and only if)

$$\left(-\lambda I_{2n} - H_0 - \sum_{k=1}^{K} \left(H_k e^{-\lambda\tau_k} + H_{-k} e^{\lambda\tau_k}\right)\right) Jw_+ = 0.$$

By taking the transpose of the left-hand side expression and some algebraic manipulations we arrive at

$$\begin{aligned}
&w_+^\top J^\top \left(-\lambda I_{2n} - H_0^\top - \sum_{k=1}^{K} \left(H_k^\top e^{-\lambda\tau_k} + H_{-k}^\top e^{\lambda\tau_k}\right)\right) \\
&= w_+^\top \left(-\lambda J^\top - J^\top H_0^\top - \sum_{k=1}^{K} \left(J^\top H_k^\top e^{-\lambda\tau_k} + J^\top H_{-k}^\top e^{\lambda\tau_k}\right)\right) \\
&= w_+^\top \left(\lambda J - H_0 J - \sum_{k=1}^{K} \left(H_{-k} J e^{-\lambda\tau_k} + H_k J e^{\lambda\tau_k}\right)\right) \\
&= w_+^\top \left(\lambda I_{2n} - H_0 - \sum_{k=1}^{K} \left(H_{-k} e^{-\lambda\tau_k} + H_k e^{\lambda\tau_k}\right)\right) J,
\end{aligned}$$

in which we used $J^\top = -J$ and conditions (3) and (4). Because (λ, w_+) is an eigenvalue/left eigenvector pair of the NLEVP associated with (2), the last of these expressions is equal to zero and hence Jw_+ is a right eigenvector associated with $-\lambda$. Using a similar approach, it can be shown that Jv_+ is a left eigenvector associated with $-\lambda$. □

Secondly, as mentioned before, the considered NLEVP typically 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. For any $c > 0$, NLEVP (1) 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 [18, Corollary 2.6]. □

2.2 Structure preserving shift-invert Arnoldi method for linear Hamiltonian eigenvalue problems

In this subsection we review how the method presented in [17] can be used to compute the eigenvalues of (7) closest to a given purely real or imaginary shift while preserving the Hamiltonian structure of the spectrum. To this end, we will first consider the traditional shift-invert Arnoldi method. In this approach the eigenvalues of H closest to a given shift σ are approximated, by applying the Arnoldi method to the shifted and inverted matrix $(H - \sigma I_{2n})^{-1}$. Note that the eigenvalues of $(H - \sigma I_{2n})^{-1}$ are equal to $\left\{\frac{1}{\lambda_i - \sigma}\right\}_{i=1}^{2n}$, with $\{\lambda_i\}_{i=1}^{2n}$ the eigenvalues of H, meaning that the eigenvalues of H closest to σ are mapped to the eigenvalues of $(H - \sigma I_{2n})^{-1}$ with the largest magnitude. As a consequence, the Arnoldi method is likely to converge first to these eigenvalues. However, this mapping destroys the symmetries of the spectrum. As a consequence, when working in finite precision, the computed approximations for purely imaginary eigenvalues of H typically have a small but non-zero real part due to rounding errors (see also Section 6.1). In applications for which the detection of purely imaginary eigenvalues is important, such as the application mentioned in the introduction, additional processing to determine whether an eigenvalue is purely imaginary, is therefore necessary.

To avoid this additional processing, the structure preserving shift-invert Arnoldi method from [17] is preferred. For a purely real or imaginary shift σ, the Arnoldi method is now applied to the matrix

$$R_\sigma^{-1} := (H + \sigma I_{2n})^{-1} (H - \sigma I_{2n})^{-1}. \tag{8}$$

The eigenvalues of this new matrix are $\left\{\frac{1}{\lambda_i^2 - \sigma^2}\right\}_{i=1}^{2n}$ with $\{\lambda_i\}_{i=1}^{2n}$ the eigenvalues of H. Thus, as for the traditional shift-invert method, the eigenvalues of H closest to the shift σ are likely to be approximated first by the Arnoldi procedure. As we will see below, the advantage of this transformation is that the matrix R_σ^{-1} is both real and skew-Hamiltonian[1]. Furthermore, notice that each eigenvalue of R_σ^{-1} has even multiplicity (which is to be expected for a skew-Hamiltonian matrix) as both the eigenvalues λ and $-\lambda$ of H are mapped to the same eigenvalue of R_σ^{-1}. 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 projected eigenvalue problem obtained by applying the Arnoldi procedure to R_σ^{-1}. More specifically, consider the following Krylov subspace generate by R_σ^{-1},

$$\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right) = \mathrm{span}\left\{q_1, R_\sigma^{-1} q_1, R_\sigma^{-2} q_1, \ldots, R_\sigma^{-(m-1)} q_1\right\} \tag{9}$$

with $q_1 \in \mathbb{R}^{2n}$ an arbitrary real-valued starting vector. It can be shown that, due to the fact that R_σ^{-1} is skew-Hamiltonian, 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 $x^\top J y = 0$ holds, see [17, Proposition 3.3]. This has the following important consequence.

Proposition 3. Let q_1 be an arbitrary real-valued vector of length $2n$ and let $\lambda \neq 0$ be a simple eigenvalue of (7) with corresponding right eigenvector v_+ and let v_- be a right eigenvector associated with $-\lambda$, then the dimension of the intersection of $\mathrm{span}\{v_+, v_-\}$ 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 [14, Lemma 2.3], which is repeated here to ease the derivations for the non-linear case. More specifically, it follows from Property 1 (by choosing H_k equal to zero for $k \neq 0$) that if v_+ is a right eigenvector of (7) associated with λ then $w_- := J v_+$ is a left eigenvector of (7) associated with $-\lambda$. This implies that $v_-^\top J v_+ = v_-^\top w_- \neq 0$, because otherwise $-\lambda$ would not be a simple eigenvalue. Thus if the intersection of $\mathrm{span}\{v_+, v_-\}$ and $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$ has dimension 2, this would imply that both v_+ and v_- are in $\mathrm{K}_m\left(R_\sigma^{-1}, q_1\right)$, which is impossible as the Krylov subspace is J-neutral. □

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

This proposition has an important effect on the obtained approximations for the eigenvalues of R_σ^{-1}. For sake of simplicity, assume that all eigenvalues of H are simple and different from zero. Then all eigenvalues of R_σ^{-1} have multiplicity two and the right eigenspace corresponding to the eigenvalue $\frac{1}{(\pm\lambda)^2+\sigma^2}$ 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 $\mathrm{K}_m\left({R_\sigma}^{-1}, q_1\right)$. The Arnoldi recurrence relation can now be written as

$$R_\sigma^{-1} Q_m = Q_m \Psi_m + \Psi_{[m+1,m]} q_{m+1} e_m^\top \tag{10}$$

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 $e_m \in \mathbb{R}^{2n}$ the mth 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 $\Sigma_m = \mathrm{diag}(s_1, \dots, s_m)$ 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 (10) 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 the space span by v_+ and v_- that lie in the column space of Q_m what would contradict Proposition 3.

Next we will see that the result above implies that the obtained approximation for a pair of simple purely imaginary eigenvalues, $\pm\jmath\omega$, of H, is purely imaginary. Firstly, note that the eigenvalues $\pm\jmath\omega$ are mapped together to the real eigenvalue $\frac{1}{-\omega^2-\sigma^2}$ of R_σ^{-1}. 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 eigenvalue of R_σ^{-1} will be well approximated by a real and simple eigenvalue of Ψ_m. As most methods for computing eigenvalues of real matrices preserve real and simple eigenvalues even in the presence of rounding error, the approximations for $\pm\jmath\omega$ obtained via the transformation $\pm\sqrt{\frac{1}{s}+\sigma^2}$ with s the corresponding eigenvalue of Ψ_m computed in finite precision, are strictly imaginary.

Remark 1. *Note that the Arnoldi procedure described in* (10) *breaks down for $m = n$ (recall that the size of the eigenvalue problem is $2n$). This can be understood as follows: due to the J-neutrality of the constructed Krylov subspace, the vectors $v_1, \dots, v_m, Jv_1, \dots, Jv_m$ must form an orthonormal basis meaning that it is impossible to further extend the Krylov subspace for $m > n$.*

Note however that to assure that the eigenvalues of R_σ^{-1} only appear once in the projected eigenvalue problem associated with Ψ_m, 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 quickly lost when working in finite precision. Therefore, it is suggested in [17, Section 5] to use the Arnoldi recurrence relation,

$$\Psi_{[m+1,m]} q_{m+1} = R_\sigma^{-1} q_m - Q_m \Psi_{[:,m]} - JQ_m \Upsilon_{[:,m]} \tag{11}$$

with $\Psi_{[:,m]} = Q_m^\top R_\sigma^{-1} q_m$, $\Upsilon_{[:,m]} = (JQ_m)^\top R_\sigma^{-1} q_m$ (recall that this term is zero when working exact arithmetic) and $\Psi_{[m+1,m]}$ a normalisation factor such that $\|q_{m+1}\|_2 = 1$. Orghogonalization of the vector with which the Krylov subspace is extended against both Q_m and JQ_m assures that the constructed Krylov subspace remains J-neutral.

Remark 2. *Recall that σ was assumed either purely real or purely imaginary in the derivations above. 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 [14, Equation 3.2].*

2.3 Equivalent eigenvalue problem on infinite dimensional space

In this subsection we examine the relation between NLEVP (1) and a linear but infinite-dimensional eigenvalue problem. To this end, consider the space 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

$$\mathcal{H}\varphi(\theta) := \varphi'(\theta) \text{ for } \theta \in [-\tau_K, \tau_K], \tag{12}$$

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

$$\varphi'(0) = H_0\varphi(0) + \sum_{k=1}^{K} \Big(H_{-k}\varphi(-\tau_k) + H_k\varphi(\tau_k)\Big), \tag{13}$$

or in other words

$$D(\mathcal{H}) := \{\varphi \in X : \varphi' \in X \;\&\; \varphi \text{ satisfies (13)}\}.$$

It can be shown that the operator $\mathcal{H}$ only features a point spectrum. Furthermore, 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

$$\mathcal{H}\varphi = \lambda\varphi. \tag{14}$$

This function φ is called an eigenfunction of $\mathcal{H}$ associated with the eigenvalue λ. From (12), it is clear that these eigenfunctions must have the form $ve^{\lambda\cdot}$ with $v \in \mathbb{C}^{2n}$. By plugging this result into (13) it follows that there exists a correspondence between an eigenvalue/right eigenvector pair (λ, v) of the considered NLEVP and an eigenvalue/eigenfunction pair (λ, φ) of (14). This relation is stated more rigorously in the following proposition.

Proposition 4. It holds that

1. if (λ, v) is an eigenvalue/right eigenvector pair of the NLEVP associated with (2), then $(\lambda, ve^{\lambda\cdot})$ is eigenvalue/eigenfunction pair of (14), and
2. if (λ, φ) is an eigenvalue/eigenfunction pair of (14), then the eigenfunction φ is of the form $ve^{\lambda\cdot}$ with (λ,v) an eigenvalue/right eigenvector-pair of the NLEVP associated with (2).

Proof. This proposition follows from a similar argument as in [18, Proposition 2.2]. □

As seen in Proposition 4, the right eigenvectors of the considered NLEVP are connected with the eigenfunctions of the operator $\mathcal{H}$. Now we will look for an operator whose eigenfunctions have a similar connection with the left eigenvectors of the NLEVP associated with (2). 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

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

The infinite dimensional eigenvalue problem associated with this operator

$$\mathcal{G}\psi = \lambda\psi$$

has the following relation with the eigenvalue/left eigenvectors-pairs of the NLEVP associated with (2).

Proposition 5. It holds that,

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

Proof. As before, the assertions follow from a similar argument as in [18, Proposition 2.2]. □

Combined with Proposition 1, this result implies that the following relation between the eigenvalues and eigenfunctions of $\mathcal{H}$ and those of $\mathcal{G}$ holds.

Corollary 1. If $\big(\lambda, \varphi\big)$ is an eigenvalue/eigenfunction pair of $\mathcal{H}$, then $\big(-\lambda, \psi\big)$, with $\psi(\theta) = J\varphi(\theta)$ for $\theta \in [-\tau_K, \tau_K]$, is eigenvalue/eigenfunction pair of $\mathcal{G}$ and visa versa.

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

$$\mathbf{B}(\varphi, \psi) := \psi(0)^\top \varphi(0) + \sum_{k=1}^{K} \left(\int_0^{\tau_k} \psi(\theta)^\top H_{-k} \varphi(\theta - \tau_k) \, \mathrm{d}\theta - \int_0^{\tau_k} \psi(\theta - \tau_k)^\top H_k \varphi(\theta) \, \mathrm{d}\theta \right). \quad (15)$$

Notice that this bilinear form does not define an inner product as it is neither Hermetian symmetric nor is it positive definite. However, this bilinear form does induce two important relations between operators $\mathcal{H}$ and $\mathcal{G}$. Firstly, borrowing terminology from [8, Chapter 7], $\mathcal{G}$ can be seen as the *formal adjoint* of $\mathcal{H}$ with respect to bilinear form (15), since the following result holds.

Proposition 6. 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

$$\begin{aligned}
\mathbf{B}(\varphi, \mathcal{G}\psi) =& -\psi'(0)^\top \varphi(0) - \sum_{k=1}^{K} \left(\int_0^{\tau_k} \psi'(\theta)^\top H_{-k} \varphi(\theta - \tau_k) \, \mathrm{d}\theta - \int_0^{\tau_k} \psi'(\theta - \tau_k)^\top H_k \varphi(\theta) \, \mathrm{d}\theta \right) \\
=& \left(H_0^\top \psi(0) + \sum_{k=1}^{K} \left(H_k^\top \psi(-\tau_k) + H_{-k}^\top \psi(\tau_k) \right) \right)^\top \varphi(0) - \sum_{k=1}^{K} \Bigg(\Big[\psi(\theta)^\top H_{-k} \varphi(\theta - \tau_k) \Big]_0^{\tau_k} \\
& - \int_0^{\tau_k} \psi(\theta)^\top H_{-k} \varphi'(\theta - \tau_k) \, \mathrm{d}\theta - \Big[\psi(\theta - \tau_k)^\top H_k \varphi(\theta) \Big]_0^{\tau_k} + \int_0^{\tau_k} \psi(\theta - \tau_k)^\top H_k \varphi'(\theta) \, \mathrm{d}\theta \Bigg) \\
=& \psi(0)^\top \underbrace{\left(H_0 \varphi(0) + \sum_{k=1}^{K} \left(H_{-k} \varphi(-\tau_k) + H_k \varphi(\tau_k) \right) \right)}_{\varphi'(0)} + \sum_{k=1}^{K} \Bigg(\int_0^{\tau_k} \psi(\theta)^\top H_{-k} \varphi'(\theta - \tau_k) \, \mathrm{d}\theta \\
& - \int_0^{\tau_k} \psi(\theta - \tau_k)^\top H_k \varphi'(\theta) \, \mathrm{d}\theta \Bigg) \\
=& \mathbf{B}(\mathcal{H}\varphi, \psi)
\end{aligned}$$

□

Secondly, under bilinear form (15) the eigenfunctions of $\mathcal{G}$ are complementary to the eigenfunctions of $\mathcal{H}$, as spelled out in the following proposition.

Proposition 7. 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 $\mathbf{B}\left(\mathcal{H}\varphi_\lambda, \psi_\mu\right) = \mathbf{B}\left(\varphi_\lambda, \mathcal{G}\psi_\mu\right)$. Using the properties of a bilinear form one finds that

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

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

3 Structure preserving shift-invert Arnoldi method for Hamiltonian delay eigenvalue problems

In the previous section we saw that the eigenvalues of NLEVP (1) correspond to those of the linear but infinite dimensional operator $\mathcal{H}$. To compute some of the eigenvalues of this operator, a generalization of the infinite Arnoldi method for DEVPs, proposed in [11], can be used. As mentioned in the introduction, this method can be interpreted as applying the Arnoldi procedure to an operator instead of to a matrix, constructing in the process a Krylov subspace that is spanned by functions rather than by vectors. Here we want to apply similar ideas to compute the eigenvalues of $\mathcal{H}$ in the neighborhood of some purely real or imaginary shift σ, while preserving the special Hamiltonian-like structure of the spectrum. To this end, as was the case in Section 2.2, we first have to derive an adequate shift-invert transformation. Inspired by (8), lets us consider the following linear but infinite dimensional operator:

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

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

$$\mathcal{R}_\sigma\varphi(\theta) = \varphi''(\theta) - \sigma^2\varphi(\theta) \text{ for } \theta \in [-\tau_K, \tau_K] \text{ and } \varphi \in D(\mathcal{R}_\sigma), \tag{16}$$

with $D(\mathcal{R}_\sigma)$, the domain of $\mathcal{R}_\sigma$, consisting of the functions φ which are twice continuously differentiable and which fulfill the following two conditions

$$\varphi'(0) = H_0\varphi(0) + \sum_{k=1}^{K}\Big(H_{-k}\varphi(-\tau_k) + H_k\varphi(\tau_k)\Big) \text{ and} \tag{17}$$

$$\varphi''(0) = H_0\varphi'(0) + \sum_{k=1}^{K}\Big(H_{-k}\varphi'(-\tau_k) + H_k\varphi'(\tau_k)\Big), \tag{18}$$

or in other words

$$D\left(\mathcal{R}_\sigma\right) = \left\{\varphi \in X : \varphi' \in X \ ,\ \varphi'' \in X \ \&\ \varphi \text{ satisfies (17) and (18)}\right\}.$$

If σ is not an eigenvalue of $\mathcal{H}$, the operator $\mathcal{R}_\sigma$ can be inverted. More precisely, if $\sigma \neq 0$ is not an eigenvalue of $\mathcal{H}$ then for each $\phi \in X$ there exists a unique function $\varphi := \mathcal{R}_\sigma^{-1}\phi$ with

$$\varphi(\theta) := \left(\mathcal{R}_\sigma^{-1}\phi\right)(\theta) = \left(\int_0^\theta \frac{\phi(\eta)}{2\sigma}e^{-\sigma\eta}d\eta + \mathrm{C}_\sigma[\phi]\right)e^{\sigma\theta} + \left(-\int_0^\theta \frac{\phi(\eta)}{2\sigma}e^{\sigma\eta}d\eta + \mathrm{C}_{-\sigma}[\phi]\right)e^{-\sigma\theta}, \tag{19}$$

in which the integration constants $\mathrm{C}_\sigma[\phi]$ and $\mathrm{C}_{-\sigma}[\phi] \in \mathbb{C}^{2n}$ follow from conditions (17) and (18):

$$2\sigma M(\sigma)\ \mathrm{C}_\sigma[\phi] = -\phi(0) + \sum_{k=1}^{K}\left[H_k\int_0^{\tau_k}\phi(\eta)e^{-\sigma(\eta-\tau_k)}\,\mathrm{d}\eta - H_{-k}\int_0^{\tau_k}\phi(\eta-\tau_k)e^{-\sigma\eta}\,\mathrm{d}\eta\right] \tag{20}$$

$$2\sigma M(-\sigma)\ \mathrm{C}_{-\sigma}[\phi] = \quad \phi(0) - \sum_{k=1}^{K}\left[H_k \int_0^{\tau_k} \phi(\eta) e^{\sigma(\eta-\tau_k)}\ \mathrm{d}\eta - H_{-k}\int_0^{\tau_k}\phi(\eta-\tau_k)e^{\sigma\eta}\ \mathrm{d}\eta\right], \quad (21)$$

with $M(\cdot)$ the characteristic matrix as introduced in (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. If $\sigma = 0$ is not an eigenvalue of $\mathcal{H}$, then for each $\phi \in X$ the expression for $\varphi := \mathcal{R}_0^{-1}\phi$ becomes

$$\varphi(\theta) := \mathcal{R}_0^{-1}\phi(\theta) = \int_0^\theta \int_0^{\eta_2} \phi(\eta_1)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2 + \mathrm{C}_1[\phi]\theta + \mathrm{C}_0[\phi], \quad (22)$$

in which $\mathrm{C}_1[\phi]$ and $\mathrm{C}_0[\phi]$ again follow from (17) and (18). For more details, see Appendix A.

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

Proposition 8. If $\phi \in X$ is a real 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 (19), (20) and (21) by noting that $\mathrm{C}_\sigma[\phi]$ and $\mathrm{C}_{-\sigma}[\phi]$ are real if σ is purely 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 (22) and the expressions for $C_1[\phi]$ and $C_0[\phi]$ given in Appendix A. □

Lemma 1. For σ not an eigenvalue of $\mathcal{H}$, 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 bilinear form $\mathbf{B}(\cdot,\cdot)$ given in (15), and using (19) we find

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

Firstly, observe that

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

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:

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

In a similar fashion we obtain:

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

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

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

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

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

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

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

Next, we introduce some properties of the eigenvalue problem associated with this transformed operator:

$$\mathcal{R}_\sigma^{-1}\phi=\mu\phi. \tag{23}$$

Proposition 9. For σ not an eigenvalue of $\mathcal{H}$, infinite dimensional eigenvalue problem (23) 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_-, 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 1.

Thus if $\lambda \neq 0$, the pair $(\lambda,-\lambda)$ (which are both eigenvalues of $\mathcal{H}$) 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 associated with $\mathcal{R}_\sigma^{-1}$:

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

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

Proposition 10. 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 (15).

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 1 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 B), 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 1. *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 $\mathcal{H}$ with associated eigenfuction φ_+ and let φ_- be the eigenfunction associated with $-\lambda$, then the dimension of the intersection of* $\mathrm{span}\{\varphi_+,\varphi_-\}$ *and* $\mathrm{K}_m(\mathcal{R}_\sigma^{-1},\phi_1)$ *is at most 1.*

Proof. Recall from Proposition 4 that $\varphi_+(\theta) = v_+e^{\lambda\theta}$ and $\varphi_-(\theta) = v_-e^{-\lambda\theta}$. Furthermore, from Proposition 1 it follows that there exists a vector w_+ which is a left eigenvector of (2) associated with λ, such that $v_- = Jw_+$. Combing these properties with the definition of the bilinear form in (15), we find that

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

in which the last inequality follows from the fact that λ is a simple eigenvalue. Thus if the intersection of $\mathrm{span}\{\phi_+,\phi_-\}$ 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 10. □

Corollary 2. If σ is not an eigenvalue of $\mathcal{H}$, $\phi_1 \in X$ is a real-valued function and $\lambda \neq 0$ is a simple eigenvalue of $\mathcal{H}$, then 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 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 ith 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 9.

Each iteration in Algorithm 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 infinite Arnoldi method, one now has to as assure that ϕ_{i+1} is *orthogonal* against $\phi_1, \ldots, \phi_i$ with respect to $\mathbf{B}(\cdot, J\cdot)$, to avoid that the same eigenvalue of $\mathcal{R}_\sigma^{-1}$ appears twice in the projected eigenvalue problem. Although this orthogonality constraint holds by construction in exact arithmetic, it is no longer the case when working in finite precision, as we shall illustrate in Section 6.1.

To preform the orthogonalisation step, an appropriate inner product on X needs to be defined. Inspired by [11], we define the following inner product based on the scaled Chebyshev expansion of the functions in X. 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 [22]:

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

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

$$\langle \psi, \varphi \rangle_X := \sum_{l=0}^{\infty} c_l^H d_l. \tag{24}$$

Furthermore, it can be shown, see [11, Equation 4.10], that this inner product is equivalent with the following expression

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

with

$$I[f] = \int_{-\tau_K}^{\tau_K} \frac{f}{\sqrt{1 - (\theta/\tau_K)^2}} \, \mathrm{d}\theta.$$

As in the final dimensional case, Theorem 1 has an important consequence for the obtained approximations of purely imaginary eigenvalues. More specifically, assume that all eigenvalues of $\mathcal{H}$ are simple, then, as long as the Krylov subspace satisfies Proposition 10, for each eigenvalue of $\mathcal{R}_\sigma^{-1}$ only a single component of the associated two-dimensional eigenspace can appear in the Krylov subspace and hence each eigenvalue of $\mathcal{R}_\sigma^{-1}$ can only appear once in the projected eigenvalue problem associated with the first m rows of Ψ. Furthermore, since ϕ_1 is chosen real-valued, Proposition 8 and the definition of the inner product in (24), 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. As in Section 2.2, this has

as a consequence that if the Krylov space is sufficiently large, the obtained approximations for purely imaginary eigenvalues of $\mathcal{H}$ are purely imaginary.

Algorithm 1 High-level description of the infinite dimensional Arnoldi method for eigenvalue problem (1).

Choose a **real-valued** initial function ϕ_0 with $\|\phi_0\|_X = 1$.
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 $\Psi_{[i,j]} = \langle \phi_j, \varphi_{i+1} \rangle_X$ for $j = 1, \ldots, i$ and $\Psi_{[i+1,i]} = \|\varphi_{i+1} - \Psi_{[1,i]}\phi_1 - \cdots - \Psi_{[i,i]}\phi_i\|_X$.
Assure 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): $\pm\sqrt{\frac{1}{\mu} + \sigma^2}$

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 assure that the built up Krylov subspace satisfies Proposition 10, 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 1 appears. To this end, observe from (19) 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 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 (24), it makes sense to work with scaled Chebyshev functions. The extension step in Algorithm 1 can now be computed using straightforward linear algebra operations, as demonstrated in the following theorem.

Theorem 2. *Let*

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

with $q_l^{(i)} \in \mathbb{R}^{2n}$*, 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*

$$\begin{bmatrix} v_2^{(i+1)} & \cdots & v_{N_i+2}^{(i+1)} \end{bmatrix} = \tau_K^2 \begin{bmatrix} q_0^{(i)} & \cdots & q_{N_i}^{(i)} \end{bmatrix} \begin{bmatrix} \frac{1}{4} & & & & & & \\ & \frac{1}{24} & & & & & \\ \frac{-1}{6} & & \frac{1}{48} & & & & \\ & \frac{-1}{16} & & \frac{1}{80} & & & \\ \frac{1}{24} & & \frac{-1}{30} & & \frac{1}{120} & & \\ & \ddots & & \ddots & & \ddots & \\ & \frac{1}{4(N_i-1)(N_i-2)} & & \frac{-1}{2(N_i-1)(N_i+1)} & & \frac{1}{4(N_i+1)(N_i+2)} \end{bmatrix}, \quad (25)$$

while $v_1^{(i+1)}$ *and* $v_0^{(i+1)}$ *follow from* (17) *and* (18)*:*

$$M(0)\ v_1^{(i+1)} = \sum_{j=2}^{N_i+2} \left(I_{2n} \frac{T_j''(0)}{\tau_K} - H_0 T_j'(0) - \sum_{k=1}^{K} \left(H_k T_j' \left(\frac{\tau_k}{\tau_K} \right) + H_{-k} T_j' \left(\frac{-\tau_k}{\tau_K} \right) \right) \right) v_j^{(i+1)}$$

$$M(0)\ v_0^{(i+1)} = \sum_{j=1}^{N_i+2} \left(I_{2n} \frac{T_j'(0)}{\tau_K} - H_0 T_j(0) - \sum_{k=1}^{K} \left(H_k T_j \left(\frac{\tau_k}{\tau_K} \right) + H_{-k} T_j \left(\frac{-\tau_k}{\tau_K} \right) \right) \right) v_j^{(i+1)},$$

with $M(\cdot)$ *the characteristic matrix as defined in* (2)*.*

Proof. We refer to Appendix C 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, the orthogonalisation step can be carried out using standard linear algebra operations. More specifically, as a consequence of Theorem 2, the functions encountered in the ith orthogonalisation step of Algorithm 1 can be written as

$$\phi_j(\theta) = \sum_{l=0}^{N_j} q_l^{(j)} T_l \left(\frac{\theta}{\tau_K} \right) \text{ for } j = 1, \ldots, i+1 \text{ and } \varphi_{i+1}(\theta) = \sum_{l=0}^{N_{i+1}} v_l^{(i+1)} T_l \left(\frac{\theta}{\tau_K} \right), \quad (26)$$

with $q_l^{(j)}$ and $v_l^{(i+1)} \in \mathbb{R}^{2n}$. 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

$$\Psi_{[i+1,i]} q^{(i+1,i+1)} = v^{(i+1)} - Q_{(i,i+1)} Q_{(i,i+1)}^{\top} v^{(i+1)} \quad (27)$$

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 [11], 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 10. 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 assures 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 (26) the bilinear form $\mathbf{B}(\cdot, J\cdot)$ can be evaluated using matrix-vector operators:

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

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

$$-T_{l_1}(0)T_{l_2}(0),\ -\int_0^{\tau_k} T_{l_1}\left(\frac{\theta}{\tau_K}\right) T_{l_2}\left(\frac{\theta-\tau_k}{\tau_K}\right) \mathrm{d}\theta \text{ and } \int_0^{\tau_k} T_{l_1}\left(\frac{\theta-\tau_k}{\tau_K}\right) T_{l_2}\left(\frac{\theta}{\tau_K}\right) \mathrm{d}\theta,$$

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 ith 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 $q^{(j,i+1),T}S_{N_{i+1}}q^{(j,i+1)} = 0$ 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 (27):

$$\begin{aligned}\Psi_{[i+1,i]}q^{(i+1,i+1)} &= v^{(i+1)} - Q_{(i,i+1)}Q^\top_{(i,i+1)}v^{(i+1)} \\ &\quad - S_{N_{i+1}}Q_{(i,i+1)}\left(\left(S_{N_{i+1}}Q_{(i,i+1)}\right)^\top S_{N_{i+1}}Q_{(i,i+1)}\right)^{-1}\left(S_{N_{i+1}}Q_{(i,i+1)}\right)^\top v^{(i+1)}, \quad (28)\end{aligned}$$

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 10 up to machine precision, even in the presence of rounding errors.

Remark 3. *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 σ is purely real, as will be explained in Remark 4.

Similar to the case $\sigma = 0$ one can derive a compact representation for the functions and operations in Algorithm 1 by choosing an appropriate structure for the initial function ϕ_1. More specifically, if

$$\phi(\theta) = P_N(\theta)e^{\jmath\omega\theta} + \overline{P_N(\theta)}e^{-\jmath\omega\theta} \text{ for } \theta \in [-\tau_K, \tau_K], \quad (29)$$

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

$$\left(\mathcal{R}_\sigma^{-1}\phi\right)(\theta) = Q_{N+1}(\theta)e^{\jmath\omega\theta} + \overline{Q_{N+1}(\theta)}e^{-\jmath\omega\theta} \text{ for } \theta \in [-\tau_K, \tau_K], \quad (30)$$

in which $Q_{N+1}(\cdot)$ is a complex vector-valued polynomial of degree $N+1$. Using this representation the extension step in Algorithm 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 3. *If ϕ_i is given by*

$$\phi_i(\theta) := \left(\sum_{l=0}^{N_i} q_l^{(i)}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{\jmath\omega\theta} + \left(\sum_{l=0}^{N_i} \overline{q_l^{(i)}}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{-\jmath\omega\theta} \text{ for} \tag{31}$$

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

$$\varphi_{i+1}(\theta) := \left(\mathcal{R}_\sigma^{-1}\phi_i\right)(\theta) = \left(\sum_{l=0}^{N_i+1} v_l^{(i+1)}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{\jmath\omega\theta} + \left(\sum_{l=0}^{N_i+1} \overline{v_l^{(i+1)}}\, T_l\left(\frac{\theta}{\tau_K}\right)\right) e^{-\jmath\omega\theta} \tag{32}$$

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

$$\begin{aligned}
&\begin{bmatrix} v_1^{(i+1)} & \ldots & v_{N_i+1}^{(i+1)} \end{bmatrix} \left(2\jmath\omega\tau_K \begin{bmatrix} \frac{1}{4} & & & & \\ & \frac{1}{6} & & & \\ \frac{-1}{4} & & \frac{1}{8} & & \\ & \ddots & & \ddots & \\ & & \frac{-1}{2N_i} & & \frac{1}{2(N_i+2)} \end{bmatrix} + \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} \right) \\
&= \begin{bmatrix} q_0^{(i)} & \ldots & q_{N_i}^{(i)} \end{bmatrix} \tau_K^2 \begin{bmatrix} \frac{1}{4} & & & & & \\ & \frac{1}{24} & & & & \\ \frac{-1}{6} & & \frac{1}{48} & & & \\ & \frac{-1}{16} & & \frac{1}{80} & & \\ \frac{1}{24} & & \frac{-1}{30} & & \frac{1}{120} & \\ & \ddots & & \ddots & & \ddots \\ & & \frac{1}{4(N_i-1)(N_i-2)} & & \frac{-1}{2(N_i-1)(N_i+1)} & \frac{1}{4(N_i+1)(N_i+2)} \end{bmatrix}
\end{aligned} \tag{33}$$

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

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

with $M(\cdot)$ the characteristic matrix as defined in (2).

Proof. For more details on deriving this result, see Appendix D. □

Remark 4. *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

$$\left(\mathcal{R}_\sigma^{-1}\phi\right)(\theta) = Q_N^+(\theta) e^{\sigma\theta} + Q_N^-(\theta) e^{-\sigma\theta}$$

with $Q_N^+(\cdot)$ and $Q_N^-(\cdot)$ 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 (33), 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 1 are uniquely defined by the starting function ϕ_1 and can be uniquely decomposed as in (29) 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 function $P_{N_i}(\cdot)$ can grow large, while the overall function ϕ_i remains small. Another related numerical issue that can arise is that the condition number of the matrix at the left-hand side of (33) grows large as the number of iterations increases.

2. Secondly, evaluating inner product (24) is now less trivial. One may therefore opt for an approximation instead: rather than using the coefficients of the Chebyshev expansion, one can use the coefficients of the polynomial (expressed in 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 coefficient of this interpolating polynomial are a good approximation for the coefficients in the Chebyshev expansion [22]. 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 representation (29), the simultaneous orthogonalisation with respect to inner product (24) 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 3 unattractive in practice. Next, we therefore introduce a different approach to implement Algorithm 1 for $\sigma \neq 0$ using finite dimensional operations. This procedure uses polynomial approximations whose degree is adaptively chosen such that these approximations match the functions in Algorithm 1 to machine precision. To obtain these approximating polynomials we use an approach similar to the one used in Chebfun [22]. More specifically, a function ψ can be approximated by the unique 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 $[-\tau_K, \tau_K]$:

$$\widehat{\psi}(\theta) = \sum_{l=0}^{N_d} c_l T_l\left(\frac{\theta}{\tau_K}\right) \text{ such that } \widehat{\psi}(\theta_l) = \psi(\theta_l) \text{ with } \theta_l = \tau_K \cos\left(\frac{l\pi}{N_d}\right) \text{ for } l = 0, \ldots, N_d, \quad (34)$$

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 2 and 3, $\mathcal{R}_\sigma^{-1}\phi_i$ 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 characteristic matrix (2). However, for arbitrary (non-structured) functions ϕ_i computing $\mathcal{R}_\sigma^{-1}\phi_i$ involves solving ordinary differential equation (16) subjected to conditions (17)-(18). 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 to preform the extension step, 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 (34) 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

$$\begin{bmatrix} b_1 & \dots & b_{N_1+1} \end{bmatrix} = \begin{bmatrix} a_0 & \dots & a_{N_1} \end{bmatrix} \tau_K \begin{bmatrix} 1 & & & \\ & \frac{1}{4} & & \\ \frac{-1}{2} & & \frac{1}{6} & \\ & \ddots & & \ddots \\ & & \frac{-1}{2(N_1-1)} & \frac{1}{2(N_1+1)} \end{bmatrix} \tag{35}$$

and

$$M(\sigma)\ b_0 = -\chi_i(0) + (H_0 - \sigma I_n) \sum_{j=1}^{N_d+1} b_j T_j(0) + \sum_{k=1}^{K} \left(H_k e^{\sigma\tau_k} \sum_{j=1}^{N_d+1} b_j T_j\left(\frac{\tau_k}{\tau_K}\right) + H_{-k} e^{-\sigma\tau_k} \sum_{j=1}^{N_d+1} b_j T_j\left(-\frac{\tau_k}{\tau_K}\right) \right). \tag{36}$$

with $M(\cdot)$ the characteristic matrix as defined in (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 [11]. 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 (34) 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 now 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

$$\begin{bmatrix} d_1 & \dots & d_{N_2+1} \end{bmatrix} = \begin{bmatrix} c_0 & \dots & c_{N_2} \end{bmatrix} \tau_K \begin{bmatrix} 1 & & & \\ & \frac{1}{4} & & \\ \frac{-1}{2} & & \frac{1}{6} & \\ & \ddots & & \ddots \\ & & \frac{-1}{2(N_{d,2}-1)} & \frac{1}{2(N_{d,2}+1)} \end{bmatrix} \tag{37}$$

and

$$M(-\sigma)\ d_0 = -\zeta_i(0) + (H_0 + \sigma I_n) \sum_{j=1}^{N_2+1} d_j T_j(0) + \sum_{k=1}^{K} \left(H_k e^{-\sigma\tau_k} \sum_{j=1}^{N_2+1} d_j T_j\left(\frac{\tau_k}{\tau_K}\right) + H_{-k} e^{\sigma\tau_k} \sum_{j=1}^{N_2+1} d_j T_j\left(-\frac{\tau_k}{\tau_K}\right) \right). \tag{38}$$

Finally, $\widehat{\varphi}_{i+1}$ is obtained by computing the interpolating polynomial of form (34) 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 now 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 (28) to assure orthogonality with respect to both inner product (24) 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 2 and 3, the degree of the polynomial $\hat{\varphi}_{i+1}$ is not known beforehand and therefore might grow large. However,

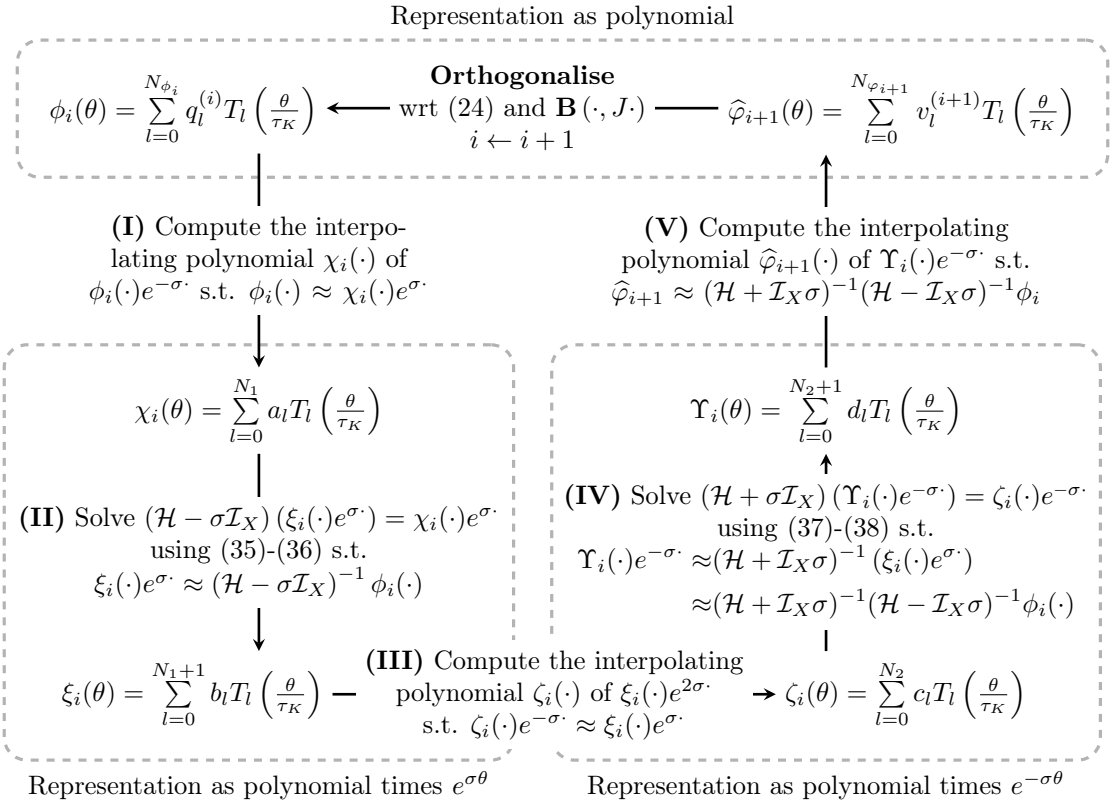


Figure 1: Structured representation of the used numerical implementation of Algorithm 1 for $\sigma \neq 0$.

the example in Section 6.2 shows that if σ is sufficiently small, the degree of these functions only grows slowly. For large σ this 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, thus 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 the Krylov subspace with respect to $\mathbf{B}(\cdot, J\cdot)$ in preserving purely imaginary eigenvalues of (1) in the presence of rounding errors. The second example shows the convergence behavior of the presented method for a large scale example.

6.1 Example 1

Consider the following characteristic matrix of form (2):

$$\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 10 & 0.1 \\ c_0 & -10 \end{bmatrix} - \begin{bmatrix} a_1 & 0 \\ 0 & 0 \end{bmatrix} e^{-\lambda} - \begin{bmatrix} 0 & 0 \\ 0 & -a_1 \end{bmatrix} e^{\lambda}, \tag{39}$$

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 characteristic roots at $\jmath\pi$ and $\frac{\jmath\pi}{2}$. Tables 1 and 2 compare three methods for approximating the

purely imaginary eigenvalues of the associated NLEVP using $\sigma = 0$ and $\sigma = \jmath\frac{3\pi}{4}$, respectively. These three methods are:

1. applying the infinite Arnoldi method to $\mathcal{H}^{-1}/\left(\mathcal{H} - \sigma\mathcal{I}_X\right)^{-1}$ using a modification of the algorithm from [11] to deal with both positive and negative delays;

2. applying the infinite Arnoldi method to $\mathcal{R}_0^{-1}/\mathcal{R}_\sigma^{-1}$ using the method presented in Section 4/Section 5 (Figure 1) but without explicit orthogonalisation with respect to $\mathbf{B}\left(\cdot, J\cdot\right)$ (i.e., using (27) instead of (28));

3. applying the infinite Arnoldi method to $\mathcal{R}_0^{-1}/\mathcal{R}_\sigma^{-1}$ using the method presented in Section 4/Section 5 (Figure 1) with explicit orthogonalisation with respect to $\mathbf{B}\left(\cdot, J\cdot\right)$.

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 $\left(\mathcal{H} - \sigma\mathcal{I}_X\right)^{-1}$, the function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top e^{\sigma\theta}$ is used.

For the first approach, we observe 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 NLEVP appears twice. This observation can be understood as follows. Recall that the eigenvalues of $\mathcal{R}_0^{-1}/\mathcal{R}_\sigma^{-1}$ always have even multiplicity. Although in exact arithmetic each eigenvalue of $\mathcal{R}_0^{-1}/\mathcal{R}_\sigma^{-1}$ appears only once in the projected eigenvalue problem as a consequence of Theorem 1, it is no longer the case when working in finite precision as orthogonality with respect to $\mathbf{B}\left(\cdot, J\cdot\right)$ is typically lost due to rounding error. As a consequence, the multiple eigenvalues of $\mathcal{R}_0^{-1}/\mathcal{R}_\sigma^{-1}$ appear twice in the projected eigenvalue problem. Numerical methods for computing eigenvalues typically do not preserve such multiple eigenvalue. Therefore, if a negative purely real eigenvalue of $\mathcal{R}_0^{-1}$, which corresponds to a pair of purely imaginary eigenvalues of $\mathcal{H}$, has multiplicity two in the projected eigenvalue problem the obtained approximations are either a pair of complex conjugate eigenvalues with a small imaginary part or two nearby purely real eigenvalues. In the former case, the obtained approximations for the purely imaginary eigenvalues of the NLEVP, have a non-zero real part. For $\mathcal{R}_\sigma^{-1}$ something similar happens. When, however, explicit orthogonalisation against $\mathbf{B}\left(\cdot, J\cdot\right)$ is preformed (approach 3), the Krylov subspace satisfies the condition in Proposition 10 up to machine precision. As to be expected, the approximations for the purely imaginary eigenvalues of (1) now only appear once in the projected eigenvalue problem and are, after the transformation $\sqrt{\frac{1}{\mu} + \sigma^2}$, purely imaginary.

From Table 2, we also observe that when using $\left(\mathcal{H} - \sigma\mathcal{I}_X\right)^{-1}$ for shifts close to origin, 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 20 iterations for the purely imaginary characteristic roots of (39) using approaches 1-3 with the constant function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top$ as initial function.

Approach 1	Approach 2	Approach 3
$2.481 \times 10^{-11} - \jmath 1.570796326781284$	$1.455 \times 10^{-12} - \jmath 1.570796326794346$	$\jmath 1.570796326748563$
$2.481 \times 10^{-11} + \jmath 1.570796326781284$	$1.455 \times 10^{-12} + \jmath 1.570796326794346$	$-\jmath 1.570796326748563$
$-4.614 \times 10^{-9} - \jmath 3.141592614193304$	$-1.455 \times 10^{-12} - \jmath 1.570796326794346$	$\jmath 3.141592653771110$
$-4.614 \times 10^{-9} + \jmath 3.141592614193304$	$-1.455 \times 10^{-12} + \jmath 1.570796326794346$	$-\jmath 3.141592653771110$
	$-\jmath 3.141592653579488$	
	$\jmath 3.141592653579488$	
	$-\jmath 3.141592653504050$	
	$\jmath 3.141592653504050$	

Table 2: Obtained approximations after 20 iterations for the purely imaginary characteristic roots of (39) using approaches 1-3 for $\sigma = \jmath\frac{3\pi}{4}$. For the approach based on $(\mathcal{H} - \sigma I_X)^{-1}$ the initial function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top e^{\sigma\theta}$ is used while for the approaches based on $\mathcal{R}_\sigma^{-1}$ the initial function $\begin{bmatrix}0.6 & 0.8\end{bmatrix}^\top$ is used.

Approach 1	Approach 2	Approach 3
$-8.913 \times 10^{-13} + \jmath 1.570796326793909$	$1.009 \times 10^{-12} - \jmath 1.570796326794204$	$-\jmath 1.570796326803690$
$2.749 \times 10^{-13} + \jmath 3.141592653590288$	$1.009 \times 10^{-12} + \jmath 1.570796326794204$	$\jmath 1.570796326803690$
$4.206 \times 10^{-6} - \jmath 1.570784493538230$	$-1.009 \times 10^{-12} - \jmath 1.570796326794204$	$-\jmath 3.141592653701270$
$-1.023 \times 10^{-3} - \jmath 3.145057545042714$	$-1.009 \times 10^{-12} + \jmath 1.570796326794204$	$\jmath 3.141592653701270$
	$-\jmath 3.141592653591553$	
	$\jmath 3.141592653591553$	
	$-\jmath 3.141592653589905$	
	$\jmath 3.141592653589905$	

6.2 Example 2

For the second example we consider the following dynamical system from [19], which describes a heated rod which 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 (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 norm of the transfer matrix in function of ω, one observes that this norm 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 $\gamma = 0.00018$, NLEVP (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 (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 close to the shift converge quickly to the eigenvalues of (6).

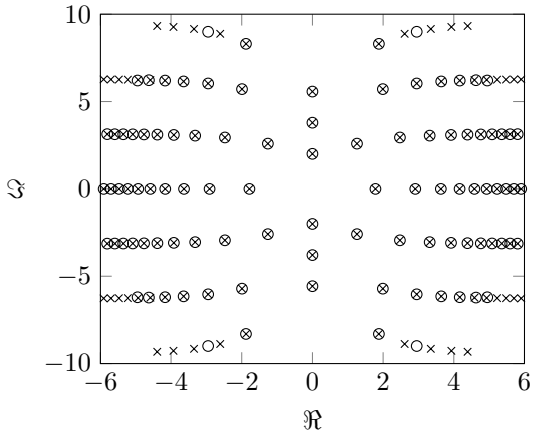
Next, we apply the method from Section 5, with implementation sketched in Figure 1, for $\sigma = \jmath 4.5$. Figure 3 shows the eigenvalues and the obtained approximations near this shift after 70 iterations. We have again fast convergence to the eigenvalues near the shift.

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 slow with respect to the number of iterations.

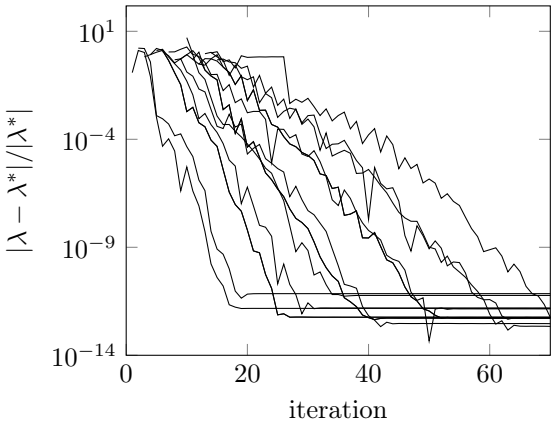
7 Conclusions and outlook

In this work we presented an iterative method to approximate the eigenvalues of NLEVPs of form (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 [14, 17] to a class of NLEVPs.

To derive this method, the equivalence between the considered NLEVP and a linear but infinite-dimensional eigenvalue problem was used. Based on this equivalence, we introduced a shift-invert transformation that preserves the Hamiltonian structure of the spectrum. Next, the ideas behind the infinite Arnoldi method from [11], which operates on functions rather than on

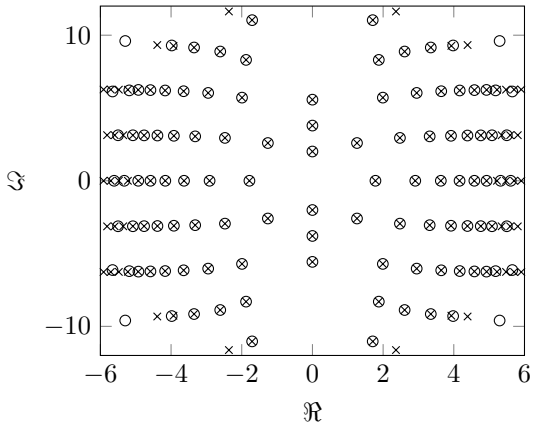


(a) Eigenvalues (crosses) and their approximations (circles) near $\sigma = 0$.

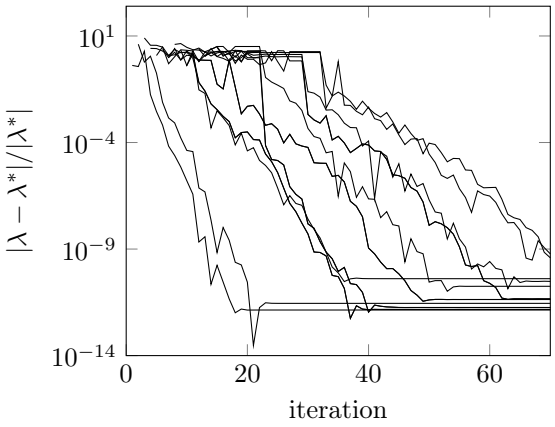


(b) Relative forward error

Figure 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

Figure 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.

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) close to the shift are 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. The code for these experiments is available from `https://twr.cs.kuleuven.be/research/software/delay-control/SPSIIA/index.html`.

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 more extensive study on the effect

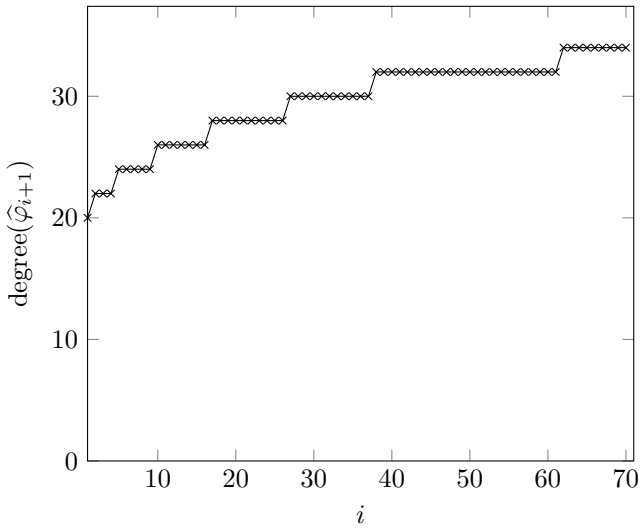


Figure 4: Evolution of the degree of $\widehat{\varphi}_{i+1}$ as defined in Figure 1, as function of the iteration number i.

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 (29) 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 [11], the convergence behavior of the method can be related to the approximation error of a Padé-approximation of the NLEVP with growing degree. Such a connection is yet to be established for the method presented here.

Acknowledgements

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". In: *International Journal of Computer Mathematics* 2.1-4 (1968), pp. 307–317.

[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*. Berlin, Heidelberg: Springer-Verlag, pp. 3–39.

[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". In: *Systems & Control Letters* 15.1 (1990), pp. 1–7.

[4] N. Bruinsma and M. Steinbuch. "A fast algorithm to compute the H∞-norm of a transfer function matrix". In: *Systems & Control Letters* 14.4 (1990), pp. 287–293.

[5] K. Gu, V. L. Kharitonov, and J. Chen. *Stability of Time-Delay Systems*. Boston, MA: Birkhäuser, 2003.

[6] S. Gumussoy and W. Michiels. "Fixed-Order H-Infinity Control for Interconnected Systems Using Delay Differential Algebraic Equations". In: *SIAM Journal on Control and Optimization* 49.5 (2011), pp. 2212–2238.

[7] S. Güttel and F. Tisseur. "The nonlinear eigenvalue problem". In: *Acta Numerica* 26 (2017), pp. 1–94.

[8] J. Hale. *Theory of functional differential equations*. Vol. 3. Applied Mathematical Sciences. New York, NY: Springer, 1977.

[9] J. K. Hale and S. M. V. Lunel. *Introduction to Functional Differential Equations*. Vol. 99. Applied Mathematical Sciences. New York, NY: Springer, 1993.

[10] D. Hinrichsen and A. J. Pritchard. *Mathematical Systems Theory I*. Vol. 48. Texts in Applied Mathematics. Berlin, Heidelberg: Springer, 2005.

[11] E. Jarlebring, K. Meerbergen, and W. Michiels. "A Krylov Method for the Delay Eigenvalue Problem". In: *SIAM Journal on Scientific Computing* 32.6 (2010), pp. 3278–3300.

[12] E. Jarlebring, W. Michiels, and K. Meerbergen. "A linear eigenvalue algorithm for the nonlinear eigenvalue problem". In: *Numerische Mathematik* 122.1 (2012), pp. 169–195.

[13] E. Jarlebring et al. "NEP-PACK: A Julia package for nonlinear eigenproblems - v0.2". In: *arXiv[1811.09592]* (2018).

[14] V. Mehrmann, C. Schröder, and V. Simoncini. "An implicitly-restarted Krylov subspace method for real symmetric/skew-symmetric eigenproblems". In: *Linear Algebra and Its Applications* 436.10 (2012), pp. 4070–4087.

[15] V. L. Mehrmann. *The autonomous linear quadratic control problem : theory and numerical solution*. Berlin, Heidelberg: Springer, 1991.

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

[17] V. Mehrmann and D. Watkins. "Structure-Preserving Methods for Computing Eigenpairs of Large Sparse Skew-Hamiltonian/Hamiltonian Pencils". In: *SIAM Journal on Scientific Computing* 22.6 (2001), pp. 1905–1925.

[18] W. Michiels and S. Gumussoy. "Characterization and Computation of $\mathcal{H}_\infty$ Norms for Time-Delay Systems". In: *SIAM Journal on Matrix Analysis and Applications* 31.4 (2010), pp. 2093–2115.

[19] W. Michiels, E. Jarlebring, and K. Meerbergen. "Krylov-Based Model Order Reduction of Time-delay Systems". In: *SIAM Journal on Matrix Analysis and Applications* 32.4 (2011), pp. 1399–1421.

[20] W. Michiels and S.-I. Niculescu. *Stability, Control, and Computation for Time-Delay Systems*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2014.

[21] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, 2003.

[22] L. N. Trefethen. *Approximation Theory and Approximation Practice*. Society for Industrial and Applied Mathematics, 2013.

[23] C. Van Loan. "A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix". In: *Linear Algebra and its Applications* 61 (1984), pp. 233–251.

[24] H. Voss. "Nonlinear Eigenvalue Problems". In: *Handbook of Linear Algebra*. Ed. by L. Hogben. Boca Raton, FL: Chapman and Hall/CRC, 2013. Chap. 60, pp. 1063–1086.

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

Appendix

A Expression for $\mathcal{R}_\sigma^{-1}$

In this section we derive an expression for the inverse of $\mathcal{R}_\sigma$ for σ not an eigenvalue of $\mathcal{H}$. More specifically, the operator $\mathcal{R}_\sigma : D(\mathcal{R}_\sigma) \subseteq X \mapsto X$ is invertible if for every function $\phi \in X$ there exists a unique function $\varphi \in D(\mathcal{R}_\sigma)$ such that $\phi = \mathcal{R}_\sigma \varphi$. It follows from the definition of $\mathcal{R}_\sigma$ and its domain that this function φ must satisfy the following four conditions: it must be twice continuously differentiable, it must satisfy the differential equation

$$\varphi''(\theta) - \sigma^2 \varphi(\theta) = \phi(\theta) \text{ for } \theta \in [-\tau_K, \tau_K] \tag{40}$$

and it must fulfill conditions (17) and (18). Because any function $\phi \in X$ is continuous, differential equation (40) always has a solution, defined up to some integration constants. This solution can be found using the variation of parameters approach. For $\sigma \neq 0$ one finds:

$$\varphi(\theta) = \left(\mathrm{C}_\sigma[\phi] + \frac{1}{2\sigma} \int_0^\theta \phi(\eta) e^{-\sigma\eta} \,\mathrm{d}\eta \right) e^{\sigma\theta} + \left(\mathrm{C}_{-\sigma}[\phi] - \frac{1}{2\sigma} \int_0^\theta \phi(\eta) e^{\sigma\eta} \,\mathrm{d}\eta \right) e^{-\sigma\theta},$$

in which the constants $\mathrm{C}_\sigma[\phi]$ and $\mathrm{C}_{-\sigma}[\phi]$ are uniquely defined by conditions (17) and (18) when σ is not an eigenvalue of $\mathcal{H}$. An expression for these integration constants can be obtained by plugging the above solution into conditions (17) and (18):

$$\begin{aligned}
&\sigma \left(\mathrm{C}_\sigma[\phi] - \mathrm{C}_{-\sigma}[\phi] \right) = H_0 \left(\mathrm{C}_\sigma[\phi] + \mathrm{C}_{-\sigma}[\phi] \right) \times \\
&\quad + \sum_{k=1}^{K} H_k \left(\left(\mathrm{C}_\sigma[\phi] + \frac{1}{2\sigma} \int_0^{\tau_k} \phi(\eta) e^{-\sigma\eta} \,\mathrm{d}\eta \right) e^{\sigma\tau_k} + \left(C_{-\sigma}[\phi] - \frac{1}{2\sigma} \int_0^{\tau_k} \phi(\eta) e^{\sigma\eta} \,\mathrm{d}\eta \right) e^{-\sigma\tau_k} \right) \\
&\quad + H_{-k} \left(\left(C_\sigma[\phi] + \frac{1}{2\sigma} \int_0^{-\tau_k} \phi(\eta) e^{-\sigma\eta} \,\mathrm{d}\eta \right) e^{-\sigma\tau_k} + \left(C_{-\sigma}[\phi] - \frac{1}{2\sigma} \int_0^{-\tau_k} \phi(\eta) e^{\sigma\eta} \,\mathrm{d}\eta \right) e^{\sigma\tau_k} \right)
\end{aligned}$$

$$\begin{aligned}
&\sigma^2 \left(\mathrm{C}_\sigma[\phi] + \mathrm{C}_{-\sigma}[\phi] \right) + \phi(0) = \sigma H_0 \left(\mathrm{C}_\sigma[\phi] - \mathrm{C}_{-\sigma}[\phi] \right) \times \\
&\quad + \sum_{k=1}^{K} \sigma H_k \left(\left(\mathrm{C}_\sigma[\phi] + \frac{1}{2\sigma} \int_0^{\tau_k} \phi(\eta) e^{-\sigma\eta} \,\mathrm{d}\eta \right) e^{\sigma\tau_k} - \left(C_{-\sigma}[\phi] - \frac{1}{2\sigma} \int_0^{\tau_k} \phi(\eta) e^{\sigma\eta} \,\mathrm{d}\eta \right) e^{-\sigma\tau_k} \right) \\
&\quad + \sigma H_{-k} \left(\left(C_\sigma[\phi] + \frac{1}{2\sigma} \int_0^{-\tau_k} \phi(\eta) e^{-\sigma\eta} \,\mathrm{d}\eta \right) e^{-\sigma\tau_k} - \left(C_{-\sigma}[\phi] - \frac{1}{2\sigma} \int_0^{-\tau_k} \phi(\eta) e^{\sigma\eta} \,\mathrm{d}\eta \right) e^{\sigma\tau_k} \right).
\end{aligned}$$

Expressions (20) and (21) now follow by dividing the second equation by σ and taking the sum and difference with the first equation, respectively.

For $\sigma = 0$, one finds

$$\varphi(\theta) = \int_0^\theta \int_0^{\eta_2} \phi(\eta_1) \,\mathrm{d}\eta_1 \,\mathrm{d}\eta_2 + \mathrm{C}_1[\phi]\theta + \mathrm{C}_0[\phi]. \tag{41}$$

Plugging this solution into conditions (17) and (18) one finds the following expressions for $\mathrm{C}_1[\phi]$ and $\mathrm{C}_0[\phi]$:

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

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

Lemma 2. The bilinear form $\mathbf{B}(\cdot, J\cdot)$ is anti-symmetric, meaning that the equality $\mathbf{B}(\varphi, J\psi) = -\mathbf{B}(\psi, J\varphi)$ holds.

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

C Derivation of the extension step for $\sigma = 0$

Let $\phi_i(\theta)$ and $\varphi_{i+1}(\theta)$ be as defined in Theorem 2, 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 (25) now follows by noting that

$$\begin{aligned}
T_0(t) =& T_2''(t)/4, \\
T_1(t) =& T_3''(t)/24 \text{ and} \\
T_l(t) =& \frac{T_{l+2}''(t)}{4(l+1)(l+2)} - \frac{T_l''(t)}{2(l+1)(l-1)} + \frac{T_{l-2}''(t)}{4(l-1)(l-2)} \text{ for } l \geq 2.
\end{aligned} \tag{42}$$

The expressions for $v_1^{(i+1)}$ follows directly from (18) and using this result $v_0^{(i+1)}$ can be computed from (17).

D Derivation of the extension step for $\sigma = \jmath\omega$

Let ϕ_i and φ_{i+1} as defined in (31), then the equality $\mathcal{R}_\sigma\varphi_{i+1} = \phi_i$ becomes

$$\begin{aligned}
&\left[\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1} v_l^{(i+1)}T_l'\left(\frac{t}{\tau_K}\right) + \sum_{l=2}^{N_i+1}\frac{v_l^{(i+1)}}{\tau_K^2}T_l''\left(\frac{t}{\tau_K}\right)\right]e^{\jmath\omega\theta} + \\
&\qquad\left[-\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1}\overline{v_l^{(i+1)}}T_l'\left(\frac{t}{\tau_K}\right) + \sum_{l=2}^{N_i+1}\frac{\overline{v_l^{(i+1)}}}{\tau_K^2}T_l''\left(\frac{t}{\tau_K}\right)\right]e^{-\jmath\omega\theta} \\
&\qquad\qquad = \sum_{l=0}^{N_i} q_l^{(i)}T_l\left(\frac{\theta}{\tau_K}\right)e^{\jmath\omega\theta} + \sum_{l=0}^{N_i}\overline{q_l^{(i)}}T_l\left(\frac{\theta}{\tau_K}\right)e^{-\jmath\omega\theta}.
\end{aligned}$$

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

$$\frac{2\jmath\omega}{\tau_K}\sum_{l=1}^{N_i+1} v_l^{(i+1)}T_l'\left(\frac{t}{\tau_K}\right) + \sum_{l=2}^{N_i+1}\frac{v_l^{(i+1)}}{\tau_K^2}T_l''\left(\frac{t}{\tau_K}\right) = q_l^{(i)}T_l\left(\frac{\theta}{\tau_K}\right).$$

Using (42) in combination with

$$T_1'(t) = T_2''(t)/4$$
$$T_l'(t) = \frac{T_{l+1}''(t)}{2(l+1)} - \frac{T_{l-1}''(t)}{2(l-1)}$$

results in (33). The expression for $v_0^{(i+1)}$ follows again from (17) and (18).