Scientific
Research
Publishing

A Block-Preconditioned Inexact Linear Solver for Computing the Complex Eigenpairs of a Large Sparse Matrix

Richard Olatokunbo Akinola[1*], Stephen Yakubu Kutchin[1], Ayodeji Sunday Ayodele[1], Kingsley Obiajulu Muka[2]

[1]Department of Mathematics, Faculty of Natural Sciences, University of Jos, Jos, Nigeria
[2]Department of Mathematics, University of Benin, Benin, Nigeria
Email: *roakinola@gmail.com, *roakinola@yahoo.com

Abstract

In computing eigenpairs of the matrix pencil, one obtains a linear system of equations. In this work, we show how a block triadiagonal preconditioner in GMRES can be used to solve a large sparse unsymmetric system of equations inexactly using a fixed and decreasing tolerance. While the fixed tolerance solver converged superlinearly to the eigenvalue of interest, the decreasing one converged quadratically. This surpasses an earlier result which converged harmonically.

Keywords

Preconditioner, Superlinear Convergence, Quadratic Convergence, Eigenvalues

1. Introduction

Let A be a large sparse, real n by n nonsymmetric matrix and $B \in \mathbb{R}^{n \times n}$ a symmetric positive definite matrix. In this paper, we consider the problem of computing the eigenpair (z, λ) from the following generalized complex eigenvalue problem

$$Az = \lambda Bz, \ z \in \mathbb{C}^n, \ z \neq \mathbf{0}, \tag{1}$$

where $\lambda \in \mathbb{C}$ is the eigenvalue of the pencil (A, B) and z its corresponding complex eigenvector. We assume that the eigenpair of interest (z, λ) is algebraically simple, with ψ^H the corresponding left eigenvector such that [1]

$$\psi^H Bz \neq 0. \tag{2}$$

By adding the normalization

$$z^H \boldsymbol{B} z = 1, \tag{3}$$

to (1), the combined system of equations can be expressed in the form $\boldsymbol{F}(z) = \boldsymbol{0}$ as

$$\boldsymbol{F}(z) = \begin{bmatrix} (\boldsymbol{A} - \lambda \boldsymbol{B}) z \\ -\dfrac{1}{2} z^H \boldsymbol{B} z + \dfrac{1}{2} \end{bmatrix} = \boldsymbol{0}. \tag{4}$$

Note that $z^H \boldsymbol{B} z$ is real since $\boldsymbol{B}$ is symmetric and positive definite. This results in solving a system of nonlinear n complex and one real equation for the $n+1$ complex unknowns $\boldsymbol{v} = [z, \lambda]^{\mathrm{T}}$. The reason we cannot use Newton's method to solve (4) is because $\bar{z}$ in the normalization $z^H \boldsymbol{B} z = \bar{z}^{\mathrm{T}} \boldsymbol{B} z$ is not differentiable.

Recall that for a real eigenpair (z, λ), (4) is $(n+1)$ real equations for $(n+1)$ real unknowns and Newton's method for solving (4) involves the solution of the $(n+1)$ square linear systems

$$\begin{bmatrix} \boldsymbol{A} - \lambda^{(k)} \boldsymbol{B} & -\boldsymbol{B} z^{(k)} \\ -\left(\boldsymbol{B} z^{(k)}\right)^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \Delta z^{(k)} \\ \Delta \lambda^{(k)} \end{bmatrix} = - \begin{bmatrix} \left(\boldsymbol{A} - \lambda^{(k)} \boldsymbol{B}\right) z^{(k)} \\ -\dfrac{1}{2} z^{(k)\mathrm{T}} \boldsymbol{B} z^{(k)} + \dfrac{1}{2} \end{bmatrix}, \tag{5}$$

for the $(n+1)$ real unknowns $\Delta \boldsymbol{v}^{(k)} = \left[\Delta z^{(k)\mathrm{T}}, \Delta \lambda^{(k)} \right]$, and updating $\boldsymbol{v}^{(k+1)} = \boldsymbol{v}^{(k)} + \Delta \boldsymbol{v}^{(k)}$. Secondly, if instead of the normalization (3), we add $c^H z = 1$, where c is a fixed complex vector (see, for example, [2]), (1) and $c^H z = 1$ provide $(n+1)$ complex equations for $(n+1)$ complex unknowns, and the Jacobian of this new system is

$$\begin{bmatrix} (\boldsymbol{A} - \lambda \boldsymbol{B}) & -\boldsymbol{B} z \\ c^H & 0 \end{bmatrix}.$$

The above Jacobian is square and can be easily shown to be nonsingular, using the ABCD Lemma [3] if the eigenvalue of interest is algebraically simple and $c^H z \neq 0$. Thirdly, if (z, λ) is complex, then, as stated earlier, we have n complex and one real equation. Also, if z solves (4), then so does $z \mathrm{e}^{i\theta}$ for any θ, such that $0 \leq \theta \leq 2\pi$.

Our approach for analyzing the solution of (4) for $\boldsymbol{v}$ begins by splitting the eigenpair (z, λ) into their real and imaginary parts: $z = z_1 + i z_2$, $\lambda = \alpha + i\beta$ where $z_1, z_2 \in \mathbb{R}^n$, and $\alpha, \beta \in \mathbb{R}$. After expanding (4), we will obtain a real $2n+1$ under-determined system of nonlinear equations in $2n+2$ real unknowns $\boldsymbol{v} = [z_1, z_2, \alpha, \beta]^{\mathrm{T}}$, and it is natural to use the Gauss-Newton method (see, for example, Deuflhard ([4], pp. 222-223)) to obtain a solution (see also, [5] [6] [7] [8]). By linearizing the under-determined system of nonlinear equations, we obtain an under-determined system of linear equations involving the Jacobian. This paper is structured as follows: in Section 2, we show that the Jacobian has a unique nullvector at the root. This is then followed in Section 3, we present two orthogonality results and **Algorithm 1** is given. In Section 4, we present an inexact inverse iteration algorithm with preconditioning for solving the large system of equations encountered.

Algorithm 1. Computing the complex eigenvalues of the pencil (A, B).

Input: A, B, J , $w^{(0)} = \left[z_1^{(0)}, z_2^{(0)} \right]^T, \alpha^{(0)}, \beta^{(0)}$, k_{max} ,

1) for $k = 0, 1, 2, \cdots$ until convergence do

2) Solve $Mu = B_1 w$ for u :

- Find the lu factorization of M *i.e.*, $[L, U] = lu(M)$.
- Solve the lower triangular system $Lc = B_1 w$ for c .
- Solve the Upper triangular system $Uu = c$ for u .

3) Compute $n_\alpha = w^T B_1 J u$ and $n_\beta = w^T B_1 u$

4) Form the matrix

$$T = \begin{bmatrix} n_\beta & -n_\alpha \\ n_\alpha & n_\beta \end{bmatrix}.$$

5) Find the LU factorization of T , that is $[L, U] = lu(T)$

6) Solve $Lc = \begin{bmatrix} \dfrac{1}{2} \left(w^T B_1 w + 1 \right) \\ \dfrac{n_w^T w}{\left(1 + \|u\|^2 \right)} \end{bmatrix}$ for c .

7) Solve $U\omega = c$ for $\omega = \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix}$.

8) Form the vector $w^+ = \Delta\alpha u - \Delta\beta J u$.

9) Update

$$\begin{bmatrix} \alpha^+ \\ \beta^+ \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix}.$$

10) **end for**

Output: $v^{(k_{max})} = \left[w^{(k_{max})}, \alpha^{(k_{max})}, \beta^{(k_{max})} \right]$.

The main mathematical tools used in this paper are the LU factorization, inexact inverse iteration and preconditioned Generalized Minimal Residual (GMRES) [9]. The main reason for using inexact inverse iteration is due to the fact that as mentioned in an earlier paper, we do not solve $Mu = B_1 w$, in practice but we will use it to show that the solution is possible. Theorem 2.1 shows that the Jacobian has a single nullvector at the root, while Theorem 3.1 gives an important orthogonality result. **Algorithms 1-3** are presented. We remark that in the limit, the approximate nullvector converges to the exact. A numerical example is given which supports the validity of the algorithms presented though, as usual, relies on good initial guesses to the desired eigenpair. The classical inverse iteration for the matrix pencil converges slowly for some eigenvalue problems while we present algorithms that converge quadratically. Throughout this paper, unless otherwise stated all norms are the 2-norm.

The following result helps to enforce the validity of the results in this paper.

Lemma 1.1: [10] *Let $F_w(w)$ be of full rank. If $F_w(w)\Delta w = F(w)$, is an under-determined linear system of equations, then its least squares solution $\Delta w = -F_w(w)^T \left[F_w(w) F_w(w)^T \right]^{-1} F(w)$, is orthogonal to the nullspace of $F_w(w)$.*

Algorithm 2. Inexact inverse iteration algorithm.

Input: A, B , $w^{(0)} = \left[z_1^{(0)}, z_2^{(0)}, \alpha^{(0)}, \beta^{(0)} \right]^{\mathrm{T}}$, $k_{\max}$.

   for $k = 0, 1, 2, \cdots$, until convergence

      Choose $\tau^{(k)}$.

      Solve $M^{(k)} u^{(k)} = B_1 w^{(k)}$ inexactly,

$$\left\| \mathcal{P}^{-1} M^{(k)} u^{(k)} - \mathcal{P}^{-1} B_1 w^{(k)} \right\| \le \tau^{(k)}.$$

   end for

Output: $u^{(k_{\max})}$.

Algorithm 3. Complex eigenvalues of the pencil (A, B) using Inexact Inverse Iteration with preconditioning.

Input: A, B, J , $w^{(0)} = \left[z_1^{(0)} ; z_2^{(0)} \right], \alpha^{(0)}, \beta^{(0)}$, $k_{\max}$, tol

   for $k = 0, 1, 2, \cdots$ until convergence do

      Solve $M^{(k)} u^{(k)} = B_1 w^{(k)}$ inexactly using **Algorithm 2.**

      Compute $n_\alpha = w^{\mathrm{T}} B_1 J u$ and $n_\beta = w^{\mathrm{T}} B_1 u$

      Solve

$$\begin{bmatrix} n_\beta & -n_\alpha \\ n_\alpha & n_\beta \end{bmatrix} \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(w^{\mathrm{T}} B_1 w + 1 \right) \\ 0 \end{bmatrix}, \quad \text{for } \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix}.$$

      Form the vector $w^+ = \Delta\alpha u - \Delta\beta J u$.

      Update

$$\begin{bmatrix} \alpha^+ \\ \beta^+ \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix}.$$

   end for

Output: $v^{(k_{\max})} = \left[w^{(k_{\max})}, \alpha^{(k_{\max})}, \beta^{(k_{\max})} \right]$.

In the next section, we will express both z and λ as $\lambda = \alpha + i\beta$ and $z = z_1 + iz_2$, convert the nonlinear system (4) to a real under-determined system of nonlinear equations and prove some important results.

2. Computation of Complex Eigenpairs by Solving an Under-Determined System of Nonlinear Equations

In this section, we will expand the system of nonlinear n complex and one real equations in $n+1$ complex unknowns (4) by writing z and λ as $z = z_1 + iz_2$ and $\lambda = \alpha + i\beta$, respectively. The reason for having an under-determined system of equations instead of a square system of equations is because, expanding $z^H B z = 1$ gives only one real equation, since B is symmetric positive definite, while $(A - \lambda B) z = 0$ results in $2n$ real equations. This results in a $2n+1$ real under-determined system of nonlinear equations in $2n + 2$ real unknowns. This will then be followed by presenting the under-determined system of nonlinear equations and explicit expression for its Jacobian. Furthermore, we will show in the main result of this section-Theorem 2.1 that if the eigenvalue of interest in (A, B) is algebraically simple, then the Jacobian has linearly independent rows. We will find the right nullvector of the

Jacobian at the root and proof that it is unique.

If we let $z = z_1 + iz_2$ and $\lambda = \alpha + i\beta$, then the square nonlinear system of Equations (4) can be written as

$$
\begin{aligned}
(A - \lambda B)z &= \left[A - (\alpha + i\beta)B\right](z_1 + iz_2) \\
&= (A - \alpha B)z_1 + \beta B z_2 + i\left[(A - \alpha B)z_2 - \beta B z_1\right],
\end{aligned}
\tag{6}
$$

and

$$
z^H B z = z_1^{\mathrm{T}} B z_1 + z_2^{\mathrm{T}} B z_2.
\tag{7}
$$

Hence,

$$
-\frac{1}{2}z^H B z + \frac{1}{2} = -\frac{1}{2}\left(z_1^{\mathrm{T}} B z_1 + z_2^{\mathrm{T}} B z_2\right) + \frac{1}{2} = 0.
$$

Since $(A - \lambda B)z = 0$, we equate the real and imaginary parts of (6) to zero and obtain the $2n$ real equations $(A - \alpha B)z_1 + \beta B z_2 = 0$ and $(A - \alpha B)z_2 - \beta B z_1 = 0$. This means, $F(v)$ consists of the $2n$ real equations arising from (6) and one real equation $-\frac{1}{2}\left(z_1^{\mathrm{T}} B z_1 + z_2^{\mathrm{T}} B z_2\right) + \frac{1}{2} = 0$;

$$
F(v) = \begin{bmatrix} (A - \alpha B)z_1 + \beta B z_2 \\ -\beta B z_1 + (A - \alpha B)z_2 \\ -\frac{1}{2}\left(z_1^{\mathrm{T}} B z_1 + z_2^{\mathrm{T}} B z_2\right) + \frac{1}{2} \end{bmatrix} = 0,
\tag{8}
$$

where $F : \mathbb{R}^{(2n+2)} \to \mathbb{R}^{(2n+1)}$. The Jacobian $F_v(v)$ of $F(v)$ which has the following explicit expression

$$
F_v(v) = \begin{bmatrix} (A - \alpha B) & \beta B & -B z_1 & B z_2 \\ -\beta B & (A - \alpha B) & -B z_2 & -B z_1 \\ -(B z_1)^{\mathrm{T}} & -(B z_2)^{\mathrm{T}} & 0 & 0 \end{bmatrix},
\tag{9}
$$

is a $2n + 1$ by $2n + 2$ real matrix. From the Jacobian (9) above, we define the real $2n$ by $2n$ matrix M as

$$
M = \begin{bmatrix} (A - \alpha B) & \beta B \\ -\beta B & (A - \alpha B) \end{bmatrix}.
\tag{10}
$$

Also, we form the $2n$ by 2 real matrix

$$
N = \begin{bmatrix} -B z_1 & B z_2 \\ -B z_2 & -B z_1 \end{bmatrix} = \begin{bmatrix} -B_1 w & B_1 w_1 \end{bmatrix},
\tag{11}
$$

consisting of the product of $B_1 = \begin{bmatrix} B & O \\ O & B \end{bmatrix}$ and the matrix of right nullvectors of M at the root, where

$$
w = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad w_1 = \begin{bmatrix} z_2 \\ -z_1 \end{bmatrix},
\tag{12}
$$

and O is the n by n zero matrix. The Jacobian (9) can be rewritten in the following partitioned form

$$F_v(v) = \begin{bmatrix} M & -B_1 w & B_1 w_1 \\ -(B_1 w)^{\mathrm{T}} & 0 & 0 \end{bmatrix} = \begin{bmatrix} M & N \\ -(B_1 w)^{\mathrm{T}} & 0^{\mathrm{T}} \end{bmatrix}, \qquad (13)$$

with M, N are as defined in (10) and (11) respectively. Note that because at the root,

$$\begin{bmatrix} (A - \alpha B) & \beta B \\ -\beta B & (A - \alpha B) \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} (A - \alpha B)z_1 + \beta B z_2 \\ (A - \alpha B)z_2 - \beta B z_1 \end{bmatrix} = 0,$$

this implies that $\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$ or its nonzero scalar multiple is a right nullvector of M. In the same vein, we find

$$\begin{bmatrix} (A - \alpha B) & \beta B \\ -\beta B & (A - \alpha B) \end{bmatrix}\begin{bmatrix} z_2 \\ -z_1 \end{bmatrix} = \begin{bmatrix} (A - \alpha B)z_2 - \beta B z_1 \\ -\{(A - \alpha B)z_1 + \beta B z_2\} \end{bmatrix} = 0,$$

and $\begin{bmatrix} z_2 \\ -z_1 \end{bmatrix}$ or its nonzero scalar multiple is also a right nullvector of M at the root. Since the eigenvalue λ of (A, B) is algebraically simple by assumption, then by (2), we need to give explicit expressions for the left nullvector of $(A - \lambda B)$ in order to prove that the Jacobian has full row rank at the root. Observe that if we define $\psi = \psi_1 + i\psi_2$, where $\psi_1, \psi_2 \in \mathbb{R}^n \setminus \{0\}$ for all $\psi \in \mathcal{N}(A - \lambda B)^H \setminus \{0\}$, then this implies

$$\psi^H(A - \lambda B) = (\psi_1^{\mathrm{T}} - i\psi_2^{\mathrm{T}})[(A - \alpha B) - i\beta B]$$
$$= \psi_1^{\mathrm{T}}(A - \alpha B) - \beta\psi_2^{\mathrm{T}}B - i[\beta\psi_1^{\mathrm{T}}B + \psi_2^{\mathrm{T}}(A - \alpha B)] = 0^{\mathrm{T}}.$$

Hence, $\psi_1^{\mathrm{T}}(A - \alpha B) - \beta\psi_2^{\mathrm{T}}B = 0^{\mathrm{T}}$ and $\beta\psi_1^{\mathrm{T}}B + \psi_2^{\mathrm{T}}(A - \alpha B) = 0^{\mathrm{T}}$. The implication of this is that

$$\begin{bmatrix} \psi_1^{\mathrm{T}} & \psi_2^{\mathrm{T}} \end{bmatrix}M = \begin{bmatrix} \psi_1^{\mathrm{T}} & \psi_2^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} (A - \alpha B) & \beta B \\ -\beta B & (A - \alpha B) \end{bmatrix}$$
$$= \begin{bmatrix} \psi_1^{T}(A - \alpha B) - \beta\psi_2^{\mathrm{T}}B & \beta\psi_1^{\mathrm{T}}B + \psi_2^{\mathrm{T}}(A - \alpha B) \end{bmatrix} = 0^{\mathrm{T}}.$$

which means, $\begin{bmatrix} \psi_1^{\mathrm{T}}, \psi_2^{\mathrm{T}} \end{bmatrix}$ or its nonzero scalar multiple is a left nullvector of M. Similarly,

$$\begin{bmatrix} \psi_2^{\mathrm{T}} & -\psi_1^{\mathrm{T}} \end{bmatrix}M = \begin{bmatrix} \psi_2^{\mathrm{T}} & -\psi_1^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} (A - \alpha B) & \beta B \\ -\beta B & (A - \alpha B) \end{bmatrix}$$
$$= \begin{bmatrix} \beta\psi_1^{\mathrm{T}}B + \psi_2^{\mathrm{T}}(A - \alpha B) & -\{\psi_1^{\mathrm{T}}(A - \alpha B) - \beta\psi_2^{\mathrm{T}}B\} \end{bmatrix} = 0^{\mathrm{T}},$$

and it shows that $\begin{bmatrix} \psi_2^{\mathrm{T}}, -\psi_1^{\mathrm{T}} \end{bmatrix}$ is also a left nullvector of M.

So we form the matrix C consisting of the 2-dimensional left nullvectors of M at the root (in practice C is not computed), as

$$C = \begin{bmatrix} \psi_1 & \psi_2 \\ \psi_2 & -\psi_1 \end{bmatrix}. \qquad (14)$$

Now, observe that the condition (2), implies

$$\psi^H Bz = \begin{bmatrix} \psi_1^{\mathrm{T}}Bz_1 + \psi_2^{\mathrm{T}}Bz_2 \end{bmatrix} + i\begin{bmatrix} \psi_1^{\mathrm{T}}Bz_2 - \psi_2^{\mathrm{T}}Bz_1 \end{bmatrix} \neq 0.$$

Therefore, at the root, either $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_1 + \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_2 \neq 0$ or $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_2 - \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_1 = 0$; $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_1 + \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_2 = 0$ or $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_2 - \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_1 \neq 0$ or both $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_1 + \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_2$ and $\boldsymbol{\psi}_1^{\mathrm{T}}\boldsymbol{B}z_2 - \boldsymbol{\psi}_2^{\mathrm{T}}\boldsymbol{B}z_1$ are nonzero. It excludes the possibility that they are both zero.

Before we continue with the rest of the analysis, we pause a little to present the main result of this section which shows that the Jacobian (9) has a one dimensional nullvector at the root.

Theorem 2.1 *Assume that the eigenpairs* (z, λ) *of the pencil* $(\boldsymbol{A}, \boldsymbol{B})$ *are algebraically simple. If* z_1 *and* z_2 *are nonzero vectors, then* $\boldsymbol{\phi} = \tau\left[z_2^{\mathrm{T}}, -z_1^{\mathrm{T}}, 0, 0\right]$, $\tau \in \mathbb{R}$ *is the unique nonzero nullvector of* $\boldsymbol{F}_v(\boldsymbol{v})$ *at the root.*

Proof: See [5].

After linearizing $\boldsymbol{F}(\boldsymbol{v}) = \boldsymbol{0}$, we have the following under-determined linear system of equations

$$\boldsymbol{F}_v\left(\boldsymbol{v}^{(k)}\right)\Delta\boldsymbol{v}^{(k)} = -\boldsymbol{F}\left(\boldsymbol{v}^{(k)}\right). \tag{15}$$

Let $\boldsymbol{n}^{(k)}$ be the exact nullvector of the Jacobian $\boldsymbol{F}_v\left(\boldsymbol{v}^{(k)}\right)$. By adding the extra condition $\boldsymbol{n}^{\mathrm{T}}\Delta\boldsymbol{v}^{(k)} = 0$, which stems from Lemma 1.1 to the under-determined linear system of Equations (15), we obtain the following square linear system of equations

$$\begin{bmatrix} \boldsymbol{F}_v\left(\boldsymbol{v}^{(k)}\right) \\ \boldsymbol{n}^{(k)\mathrm{T}} \end{bmatrix}\Delta\boldsymbol{v}^{(k)} = -\begin{bmatrix} \boldsymbol{F}\left(\boldsymbol{v}^{(k)}\right) \\ 0 \end{bmatrix}. \tag{16}$$

3. Square System of Equations for the Numerical Computation of the Complex Eigenvalues of a Matrix

In the preceding section, we saw that by adding an extra equation to the under-determined linear system of equations 15, we obtained a square system of equations (16). However, in practice we would never compute $\boldsymbol{n}^{(k)}$, but Theorem 2.1 guarantees the existence of a unique nullvector $\boldsymbol{\phi}^{(k)}$ at the root. This is the motivation for the discussion in this section. In this section, we will use $\boldsymbol{\phi}^{(k)}$ defined by $\boldsymbol{\phi}^{(k)} = \left[z_2^{(k)}, -z_1^{(k)}, 0, 0\right]$ as an approximation to $\boldsymbol{n}$ in (16) and show that the solution obtained by solving (15) is equivalent to the solution obtained by solving

$$\begin{bmatrix} \boldsymbol{F}_v\left(\boldsymbol{v}^{(k)}\right) \\ \boldsymbol{\phi}^{(k)\mathrm{T}} \end{bmatrix}\Delta\boldsymbol{v}^{(k)} = -\begin{bmatrix} \boldsymbol{F}\left(\boldsymbol{v}^{(k)}\right) \\ 0 \end{bmatrix}, \tag{17}$$

in the absence of round off errors. To do this, we will show that $\boldsymbol{\phi}^{(k)\mathrm{T}}\Delta\boldsymbol{v}^{(k)} = 0$ for each *k*, and this is presented in the main result of this section: Theorem 3.1. **Algorithm 1** is given for computing an algebraically simple eigenpair of the pencil $(\boldsymbol{A}, \boldsymbol{B})$. Note that since $\boldsymbol{M}$ has been shown to be singular at the root in section 2, this section is anchored on the assumption that when $\boldsymbol{v}$ is not at the root, $\boldsymbol{M}$ is nonsingular.

First, we define the $2n$ by $2n$ matrix $\boldsymbol{J}$ as (see, also [11])

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \tag{18}$$

and

$$Jw = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} z_2 \\ -z_1 \end{bmatrix}. \tag{19}$$

The matrix J satisfies the following properties:

1) $J^{\mathrm{T}} = -J$.

2) $J^{\mathrm{T}}J = I_{2n}$, where I_{2n} is the $2n$ by $2n$ identity matrix.

3) $J^2 = -I_{2n}$.

4) The matrix J commutes with M, *i.e.,* $JM = MJ$.

5) For $w \in \mathbb{R}^{2n}$, $w^{\mathrm{T}}B_1 Jw = w^{\mathrm{T}}JB_1 w = 0$.

6) Let u be an unknown vector that solves $Mu = B_1 w$. By premultiplying both sides by J we obtain $JMu = JB_1 w$ and hence $MJu = JB_1 w$ because of the commutativity of M and J. Therefore, if $Mu = B_1 w$, then Ju solves $M(Ju) = JB_1 w$.

We begin by writing the linear system of Equations (15) explicitly. For ease of notation, we shall drop the superscripts and define $w^+ = w + \Delta w$ where $w^+ = w^{(k+1)}$, and replace $w^{(k)}$ and $\left[\Delta z_1^{(k)\mathrm{T}}, \Delta z_2^{(k)\mathrm{T}} \right]$ with w and Δw respectively. This means that (15) can now be rewritten as:

$$\begin{bmatrix} M & -B_1 w & B_1 Jw \\ -(B_1 w)^{\mathrm{T}} & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta w \\ \Delta \alpha \\ \Delta \beta \end{bmatrix} = - \begin{bmatrix} Mw \\ -\dfrac{1}{2}w^{\mathrm{T}}B_1 w + \dfrac{1}{2} \end{bmatrix}, \tag{20}$$

which is equivalent to the following system of equations

$$M\Delta w - \Delta \alpha B_1 w + \Delta \beta B_1 Jw = -Mw$$

$$-w^{\mathrm{T}}B_1 \Delta w = \frac{1}{2}w^{\mathrm{T}}B_1 w - \frac{1}{2}.$$

After rearrangement, the first n-equation reduces to

$$Mw^+ - \Delta \alpha B_1 w + \Delta \beta B_1 Jw = \mathbf{0}. \tag{21}$$

By multiplying both sides of the $(n+1)th$ equation by 2, we obtain:

$$2w^{\mathrm{T}}B_1 \Delta w + w^{\mathrm{T}}B_1 w = 1.$$

This in turn reduces to

$$w^{\mathrm{T}}B_1 (w + 2\Delta w) = 1. \tag{22}$$

Observe that since $w^+ = w + \Delta w$, $2\Delta w = 2w^+ - 2w$ and $w + 2\Delta w = 2w^+ - w$. Now,
$w^{\mathrm{T}}B_1 (w + 2\Delta w) = w^{\mathrm{T}}B_1 (2w^+ - w) = 2w^{\mathrm{T}}B_1 w^+ - w^{\mathrm{T}}B_1 w$. Consequently,

$$w^{\mathrm{T}}B_1 w^+ = \frac{1}{2}(w^{\mathrm{T}}B_1 w + 1). \tag{23}$$

The combined set of Equations (21) and (23), which is the simplified form of (20), can be expressed as:

$$\begin{bmatrix} M & -B_1 w & B_1 Jw \\ (B_1 w)^{\mathrm{T}} & 0 & 0 \end{bmatrix} \begin{bmatrix} w^+ \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2}(w^{\mathrm{T}} B_1 w + 1) \end{bmatrix}. \tag{24}$$

We assume that the $2n$ by $2n$ matrix M is nonsingular except at the root. This is what forms the basis for the following discussion. That is to say, we want to show that when not at the root, $\phi^{(k)^{\mathrm{T}}} \Delta v^{(k)} = 0$.

First of all, let the exact nullvector n of

$$F_v(v) = \begin{bmatrix} M & -B_1 w & B_1 Jw \\ -(B_1 w)^{\mathrm{T}} & 0 & 0 \end{bmatrix},$$

be defined as $n = \left[n_w^{\mathrm{T}}, n_\alpha, n_\beta \right]$, where $n_w \in \mathbb{R}^{2n}$, n_α, n_β are real scalars, Jw and M are defined respectively by (19) and (10). Hence,

$$\begin{bmatrix} M & -B_1 w & B_1 Jw \\ -(B_1 w)^{\mathrm{T}} & 0 & 0 \end{bmatrix} \begin{bmatrix} n_w \\ n_\alpha \\ n_\beta \end{bmatrix} = 0,$$

then after expanding the matrix-vector multiplication, we obtain

$$M n_w - n_\alpha B_1 w + n_\beta (B_1 Jw) = 0 \tag{25}$$

$$w^{\mathrm{T}} B_1 n_w = 0. \tag{26}$$

We make distinctly clear at this juncture, that the nullvector $n = \left[n_w^{\mathrm{T}}, n_\alpha, n_\beta \right]$ is not exactly the same as $\phi = \left[(Jw)^{\mathrm{T}}, 0, 0 \right]$ because, the later has the form of the exact nullvector at the root, but is evaluated at the kth iterate while the former is the nullvector even when not at the root.

Another way of writing (24) is as follows

$$M w^+ = \Delta\alpha B_1 w - \Delta\beta B_1 Jw. \tag{27}$$

This means that we could solve (24) by solving

$$M u = B_1 w, \tag{28}$$

for u. After which the solution of (27) is given by

$$w^+ = \Delta\alpha u - \Delta\beta Ju. \tag{29}$$

With this expression for w^+, it can be observed that

$$M w^+ = \Delta\alpha M u - \Delta\beta M Ju = \Delta\alpha B_1 w - \Delta\beta J M u$$
$$= \Delta\alpha B_1 w - \Delta\beta J B_1 w = \Delta\alpha B_1 w - \Delta\beta B_1 Jw.$$

Which means that w^+ is well defined. Furthermore, from (25)

$$M n_w = n_\alpha B_1 w - n_\beta (B_1 Jw),$$

using the fact that J commutes with B_1 and (28) gives

$$n_w = n_\alpha u - n_\beta Ju. \tag{30}$$

Since w is B_1-orthogonal to n_w by virtue of Equation (26), taking the B_1-inner product of both sides of the above with w yields

$$w^{\mathrm{T}} B_1 n_w = n_\alpha w^{\mathrm{T}} B_1 u - n_\beta w^{\mathrm{T}} B_1 Ju = 0.$$

From which we deduce

$$\boldsymbol{n}_\alpha = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{J}\boldsymbol{u} \text{ and } \boldsymbol{n}_\beta = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{u}. \tag{31}$$

Consider the problem of solving the under-determined linear system of Equations (20) for the $2n+2$ real unknowns $\Delta\boldsymbol{v} = \left[\Delta\boldsymbol{w}^{\mathrm{T}}, \Delta\alpha, \Delta\beta\right]$. It was stated in Lemma 1.1 that the minimum norm solution to an under-determined linear system of equations is orthogonal to the nullspace. It is an application of this result that yields the following important relationship:

$$0 = \boldsymbol{n}^{\mathrm{T}}\Delta\boldsymbol{v} = \boldsymbol{n}_w^{\mathrm{T}}\Delta\boldsymbol{w} + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta. \tag{32}$$

If we add the nullvector $\boldsymbol{n}$ to the last row of (24), then

$$\begin{bmatrix} \boldsymbol{M} & -\boldsymbol{B}_1\boldsymbol{w} & \boldsymbol{B}_1\boldsymbol{Jw} \\ (\boldsymbol{B}_1\boldsymbol{w})^{\mathrm{T}} & 0 & 0 \\ \boldsymbol{n}_w^T & \boldsymbol{n}_\alpha & \boldsymbol{n}_\beta \end{bmatrix}\begin{bmatrix} \boldsymbol{w}^+ \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \frac{1}{2}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}+1\right) \\ \boldsymbol{n}_w^{\mathrm{T}}\boldsymbol{w} \end{bmatrix}. \tag{33}$$

By expanding the second to the last row, $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}^+ = \frac{1}{2}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}+1\right)$. But from (29), $\boldsymbol{w}^+ = \Delta\alpha\boldsymbol{u} - \Delta\beta\boldsymbol{Ju}$. This implies that, by taking the inner product of both sides with $\boldsymbol{w}$, yields

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}^+ = \Delta\alpha\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{u}\right) - \Delta\beta\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{Ju}\right) = \frac{1}{2}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}+1\right).$$

Using the definition (31) for $\boldsymbol{n}_\alpha$ and $\boldsymbol{n}_\beta$, we obtain

$$\boldsymbol{n}_\beta\Delta\alpha - \boldsymbol{n}_\alpha\Delta\beta = \frac{1}{2}\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{B}_1\boldsymbol{w}+1\right), \tag{34}$$

where the unknown quantities $\Delta\alpha$ and $\Delta\beta$ are to be determined, so we need an extra equation to be able to do so. The last row of the matrix-vector multiplication (*cf.* (33)) above comes from (32) since

$$\boldsymbol{n}_w^{\mathrm{T}}\boldsymbol{w}^+ + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta$$
$$= \boldsymbol{n}_w^{\mathrm{T}}\left(\boldsymbol{w}+\Delta\boldsymbol{w}\right) + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta$$
$$= \boldsymbol{n}_w^{\mathrm{T}}\boldsymbol{w} + \underbrace{\left(\boldsymbol{n}_w^{\mathrm{T}}\Delta\boldsymbol{w} + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta\right)}_{=0} = \boldsymbol{n}_w^{\mathrm{T}}\boldsymbol{w}.$$

If we substitute the expression (30) for $\boldsymbol{n}_w$ and (29) for $\boldsymbol{w}^+$ into the left hand side, then one obtains

$$\left[\boldsymbol{n}_\alpha\boldsymbol{u}^{\mathrm{T}} - \boldsymbol{n}_\beta\left(\boldsymbol{Ju}\right)^{\mathrm{T}}\right]\left[\Delta\alpha\boldsymbol{u} - \Delta\beta\boldsymbol{Ju}\right] + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta = \boldsymbol{n}_w^{\mathrm{T}}\boldsymbol{w}.$$

Furthermore, by expanding the first term on the left hand side, using the properties of $\boldsymbol{J}$, then

$$\left[\boldsymbol{n}_\alpha\boldsymbol{u}^{\mathrm{T}} - \boldsymbol{n}_\beta\left(\boldsymbol{Ju}\right)^{\mathrm{T}}\right]\left(\Delta\alpha\boldsymbol{u} - \Delta\beta\boldsymbol{Ju}\right)$$
$$= \boldsymbol{n}_\alpha\Delta\alpha\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u} + \boldsymbol{n}_\beta\Delta\beta\boldsymbol{u}^{\mathrm{T}}\boldsymbol{J}^{\mathrm{T}}\boldsymbol{Ju}$$
$$= \boldsymbol{n}_\alpha\Delta\alpha\|\boldsymbol{u}\|^2 + \boldsymbol{n}_\beta\Delta\beta\|\boldsymbol{u}\|^2.$$

Consequently,

$$n_\alpha \Delta\alpha \|\boldsymbol{u}\|^2 + n_\beta \Delta\beta \|\boldsymbol{u}\|^2 + n_\alpha \Delta\alpha + n_\beta \Delta\beta$$

$$= \left(1 + \|\boldsymbol{u}\|^2\right)\left(n_\alpha \Delta\alpha + n_\beta \Delta\beta\right) = \boldsymbol{n}_w^{\mathrm{T}} \boldsymbol{w}.$$

Observe that because $\boldsymbol{u}$ is real, $\left(1 + \|\boldsymbol{u}\|^2\right)$ is nonzero. Accordingly, after dividing both sides by $\left(1 + \|\boldsymbol{u}\|^2\right)$

$$n_\alpha \Delta\alpha + n_\beta \Delta\beta = \frac{\boldsymbol{n}_w^{\mathrm{T}} \boldsymbol{w}}{\left(1 + \|\boldsymbol{u}\|^2\right)}. \tag{35}$$

We combine the two equations (34) and (35) below

$$\begin{bmatrix} \boldsymbol{n}_\beta & -\boldsymbol{n}_\alpha \\ \boldsymbol{n}_\alpha & \boldsymbol{n}_\beta \end{bmatrix} \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \dfrac{1}{2}\left(\boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1 \boldsymbol{w} + 1\right) \\ \dfrac{\boldsymbol{n}_w^{\mathrm{T}} \boldsymbol{w}}{\left(1 + \|\boldsymbol{u}\|^2\right)} \end{bmatrix}, \tag{36}$$

to compute $\Delta\alpha$ and $\Delta\beta$ simultaneously. The matrix on the left hand side is always nonsingular except at the root (in which case all entries are zero), this is because, its determinant is $n_\alpha^2 + n_\beta^2$. Equation (35) can now be applied to simplify

$$\begin{aligned} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1 \boldsymbol{J} \boldsymbol{w}^+ &= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1 \boldsymbol{J}\left(\Delta\alpha \boldsymbol{u} - \Delta\beta \boldsymbol{J} \boldsymbol{u}\right) \\ &= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1\left(\Delta\alpha \boldsymbol{J} \boldsymbol{u} + \Delta\beta \boldsymbol{u}\right) \\ &= \Delta\alpha \left(\boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1 \boldsymbol{J} \boldsymbol{u}\right) + \Delta\beta \left(\boldsymbol{w}^{\mathrm{T}} \boldsymbol{B}_1 \boldsymbol{u}\right) \\ &= \boldsymbol{n}_\alpha \Delta\alpha + \boldsymbol{n}_\beta \Delta\beta = 0. \end{aligned} \tag{37}$$

Notice that we have used the property $\boldsymbol{J}^2 = -\boldsymbol{I}_{2n}$ to arrive at the second step above and the definition (29) for $\boldsymbol{w}^+$.

Next, we want to establish the orthogonality of $\boldsymbol{\phi}$ and $\Delta\boldsymbol{v}$ in the next key result. Before we do that, notice from Theorem (2.1) that $\boldsymbol{\phi}$, at the root, is a scalar multiple of $\left[\boldsymbol{z}_2^{\mathrm{T}}, -\boldsymbol{z}_1^{\mathrm{T}}, 0, 0\right]$ and by the definition of $\boldsymbol{J}$ in (18), we can also write $\boldsymbol{\phi} = \left[\left(\boldsymbol{J}\boldsymbol{w}\right)^{\mathrm{T}}, 0, 0\right]$, with $\boldsymbol{w} = \left[\boldsymbol{z}_1^{\mathrm{T}}, \boldsymbol{z}_2^{\mathrm{T}}\right]$. This result holds when $\boldsymbol{v} = \boldsymbol{v}^{(k)}$ is at the root or not, but because the analysis used to establish the orthogonality is based on the assumption that $\boldsymbol{M}$ is nonsinsingular when not at the root. As a result of this, after presenting Algorithm (1) (to follow shortly), we then show that the same result holds when $\boldsymbol{M}$ is singular at the root.

Theorem 3.1 *Let $\boldsymbol{\phi}$ be an approximation to the exact nullvector $\boldsymbol{n}$ of $\boldsymbol{F}_v\left(\boldsymbol{v}\right)$. Then, $\boldsymbol{\phi}$ is orthogonal to $\Delta\boldsymbol{v}$.*

Proof: To proof this, recall that $\boldsymbol{v} = \left[\boldsymbol{w}^{\mathrm{T}}, \alpha, \beta\right]$, $\boldsymbol{v}^+ = \left[\boldsymbol{w}^{+\mathrm{T}}, \alpha^+, \beta^+\right]$ and $\boldsymbol{\phi} = \left[\left(\boldsymbol{J}\boldsymbol{w}\right)^{\mathrm{T}}, 0, 0\right]$. This implies

$$\begin{aligned} \boldsymbol{\phi}^{\mathrm{T}} \Delta\boldsymbol{v} &= \boldsymbol{\phi}^{\mathrm{T}}\left(\boldsymbol{v}^+ - \boldsymbol{v}\right) = \left(\boldsymbol{J}\boldsymbol{w}\right)^{\mathrm{T}}\left(\boldsymbol{w}^+ - \boldsymbol{w}\right) = \boldsymbol{w}^{\mathrm{T}} \boldsymbol{J}^{\mathrm{T}} \boldsymbol{w}^+ - \boldsymbol{w}^{\mathrm{T}} \boldsymbol{J}^{\mathrm{T}} \boldsymbol{w} \\ &= \boldsymbol{w}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{w} - \boldsymbol{w}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{w}^+ = -\boldsymbol{w}^{\mathrm{T}} \boldsymbol{J} \boldsymbol{w}^+ = 0, \end{aligned} \tag{38}$$

showing that $\boldsymbol{\phi}^{\mathrm{T}} \Delta\boldsymbol{v} = 0$. In arriving at the last step above, we have used the properties of $\boldsymbol{J}$ and a special case of (37) where $\boldsymbol{B}_1 = \boldsymbol{I}_{2n}$.

We present Algorithm (1), which involves the solution of two linear systems.

The first is the $2n$ by $2n$ linear system of equations in (28), while the second is the 2 by 2 linear system (36).

Stop **Algorithm 1** as soon as

$$\left\|\Delta v^{(k)}\right\| \le tol,$$

where $\Delta v = \left[w^{+} - w, \omega\right]$. The above analysis shows that $\phi^{\mathrm{T}}\Delta v = 0$ when $v^{(k)}$ is not at the root. Next, we want to show that the same result holds at the root.

In a manner analogous to the proof of Lemma (2.1), we postmultiply both sides of the above system of equation by C^{T} where C is the $2n$ by 2 real matrix defined by (14), consisting of the left nullvectors of M. If M and N are as defined respectively in (10) and (11), then, this is the same as

$$C^{\mathrm{T}}Mw^{+} + C^{\mathrm{T}}N \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \mathbf{0}.$$

But by the definition of C, the first term on the left hand side of the equation above is zero, since $C^{\mathrm{T}}M = \mathbf{0}^{\mathrm{T}}$. It can be recalled from the proof of Lemma 2.1 that the 2 by 2 real matrix $H = C^{\mathrm{T}}N$ is nonsingular at the root. This implies

$$C^{\mathrm{T}}N \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = H \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} -\left(\psi_1^{\mathrm{T}}Bz_1 + \psi_2^{\mathrm{T}}Bz_2\right) & \psi_1^{\mathrm{T}}Bz_2 - \psi_2^{\mathrm{T}}Bz_1 \\ \psi_1^{\mathrm{T}}Bz_2 - \psi_2^{\mathrm{T}}Bz_1 & \psi_1^{\mathrm{T}}Bz_1 + \psi_2^{\mathrm{T}}Bz_2 \end{bmatrix} \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \mathbf{0}.$$

Accordingly, $\Delta\alpha = \Delta\beta = 0$ because of the nonsingularity of H. Therefore,

$$Mw^{+} = \mathbf{0}. \tag{39}$$

From the property of M at the root, it has two nonzero nullvectors and hence singular. The implication of this fact and the above is that, w^{+} is in the nullspace of M. But, we have already established that the nullspace of M consists of $w^{\mathrm{T}} = \left[z_1^{\mathrm{T}}, z_2^{\mathrm{T}}\right]$ and $w_1 = Jw = \left[z_2^{\mathrm{T}}, -z_1^{\mathrm{T}}\right]$. Hence, we can write

$$w^{+} = \mu w + \tau w_1.$$

Now, from the last equation in (24),

$$\frac{1}{2}\left(w^{\mathrm{T}}B_1 w + 1\right) = w^{\mathrm{T}}B_1 w^{+} = \mu w^{\mathrm{T}}B_1 w + \tau w^{\mathrm{T}}B_1 w_1$$

$$= \mu w^{\mathrm{T}}B_1 w + \tau\left(z_1^{\mathrm{T}}Bz_2 - z_2^{\mathrm{T}}Bz_1\right) = \mu w^{\mathrm{T}}B_1 w.$$

Consequently,

$$\mu = \frac{w^{\mathrm{T}}B_1 w + 1}{2w^{\mathrm{T}}B_1 w}.$$

But at the root, $w^{\mathrm{T}}B_1 w = z_1^{\mathrm{T}}Bz_1 + z_2^{\mathrm{T}}Bz_2 = 1$. Therefore, $\mu = 1$, $w^{+} = w$, $z_1^{+} = z_1$ and $z_2^{+} = z_2$. This will now be used to deduce the following corollary at the root.

Corollary 3.1 *Let* $\phi = \left[\left(Jw\right)^{\mathrm{T}}, 0, 0\right]$. *Let* $v^{+} = \left[w^{+}, \alpha^{+}, \beta^{+}\right]^{\mathrm{T}}$. *Then,* ϕ *is orthogonal to* Δv *at the root.*

Proof: This follows from the second to the last line of the proof of Theorem 3.1 (*cf.*, Equation (38)) where $w^{+} = w$. Hence,

$$\boldsymbol{\phi}^{\mathrm{T}}\Delta\boldsymbol{v} = -\boldsymbol{w}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{w}^{+} = -\boldsymbol{w}^{\mathrm{T}}\boldsymbol{J}\boldsymbol{w} = 0.$$

4. Inexact Inverse Iteration with Preconditioning for Solving (28)

In Section 2, we found two nonzero nullvectors for $\boldsymbol{M}$ at the root. As a result of this property of $\boldsymbol{M}$ at the root, in this section, we will describe an inexact inverse iteration technique for solving the large sparse system of Equations (28) in step 2 of Algorithm 1 and present Algorithm 2 and Algorithm 3. Result of a numerical experiment is given which supports the theory in Section 5.

We give the following version of inexact inverse iteration in Algorithm 2. We will use a fixed tolerance. Note that because of the special nature of $\boldsymbol{M}$ at the root, the choice of a preconditioner is crucial for convergence to the desired accuracy to be achieved. The inexact linear solver that we use is the preconditioned GMRES [9] where we define the following block tridiagonal preconditioner,

$$\mathcal{P} \approx \begin{bmatrix} \boldsymbol{A} - \alpha\boldsymbol{B} & \beta\boldsymbol{B} \\ & \boldsymbol{A} - \alpha\boldsymbol{B} \end{bmatrix}. \tag{40}$$

Next, we present Algorithm 3, which is the inexact inverse iteration equivalence of Algorithm 1. The stopping criterion for the outer iteration in Algorithm 3 depends on the norm of the eigenvalue residuals, that is

$$\left\| \boldsymbol{r}_1^{(k)} \right\| = \left\| \left(\boldsymbol{A} - \alpha^{(k)} \right) \boldsymbol{z}_1^{(k)} + \beta^{(k)} \boldsymbol{B} \boldsymbol{z}_2^{(k)} \right\| \le tol,$$

and

$$\left\| \boldsymbol{r}_2^{(k)} \right\| = \left\| \left(\boldsymbol{A} - \alpha^{(k)} \right) \boldsymbol{z}_2^{(k)} - \beta^{(k)} \boldsymbol{B} \boldsymbol{z}_1^{(k)} \right\| \le tol,$$

and

$$\left\| \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} \right\| \le tol \quad \text{and} \quad \left\| \boldsymbol{w}^{+} - \boldsymbol{w} \right\| \le tol.$$

5. Numerical Experiments

As mentioned earlier, the sparse linear system of equations in step 2 of Algorithm 1 is solved with an LU type factorization of $\boldsymbol{M}$ which is expensive and besides the L and U factors may have more nonzero elements than $\boldsymbol{M}$. In this section, our main goal is to use preconditioned GMRES with the block triangular preconditioner $\mathcal{P}$ in (40) to solve the system $\boldsymbol{M}\boldsymbol{u} = \boldsymbol{B}_1\boldsymbol{w}$ inexactly. We do this by considering a single numerical experiment with a fixed and a decreasing tolerance. Results are presented by means of tables and figures.

Example 5.1

Consider the 200 by 200 matrix $\boldsymbol{A}$ bwm200.mtx from the matrix market library [12]. It is the discretised Jacobian of the Brusselator wave model for a chemical reaction. The resulting eigenvalue problem with $\boldsymbol{B} = \boldsymbol{I}$ was also studied in [13] and we are interested in finding the rightmost eigenvalue of $\boldsymbol{A}$

which is closest to the imaginary axis and its corresponding eigenvector.

In this example, we take $\alpha^{(0)} = 0.0, \beta^{(0)} = 2.5$ in line with [13] and took $z_1^{(0)} = 1/2\|\mathbf{1}\|$ and $z_2^{(0)} = 1/\|\mathbf{1}\|$, where $\mathbf{1}$ is the vector of all ones.

In Table 1 and Table 2, we present results for the computation of the eigenpairs (z, λ) for the matrix in Example 5.1, and stop once the norms of $[\Delta\alpha, \Delta\beta]^{\mathrm{T}}$ and the eigenvalue residuals $r_1^{(k)}$ and $r_2^{(k)}$ are smaller than the outer tolerance $\tau_{\text{outer}} = 2.5 \times 10^{-14}$. *f* represents the number of inner iterations used by preconditioned GMRES to satisfy the fixed inner tolerance $\tau = 0.6$ in Table 1, while *d* in Table 2 represents number of inner iterations used by preconditioned GMRES to satisfy the decreasing inner tolerance $\tau_{\text{inner}} = \min\left\{0.6, 0.6\left\|r_1^{(k)}\right\|\right\}$. Quadratic convergence to $\lambda = 1.81999\mathrm{e} - 05 + 2.13950i$ is easily observed from the second up to the seventh iterates in columns five and six of Table 2. However, this quadratic convergence is lost in the last iterate. This could be due to the fact that we are solving a nearly singular system as we approach the root. As shown in columns five and six of Table 1, as we approach the root, more number of inner iterations

Table 1. Table showing convergence to the eigenvalue $\lambda = 1.81999\mathrm{e} - 05 + 2.13950i$ with a fixed inner tolerance for Example 5.1. The last column shows the number of inner iterations it took to satisfy the fixed inner tolerance $\tau = 0.6$.

k	α	β	$\|w^+ - w\|$	$\left\|\begin{bmatrix}\Delta\alpha\\\Delta\beta\end{bmatrix}\right\|$	$\|\Delta v\|$	f
0	0.00000e+00	2.50000	3.9e+00	7.3e−01	4.0e+00	2
1	1.81138e−01	1.78811	1.9e+00	2.3e−01	1.9e+00	2
2	1.14991e−01	2.01153	8.4e−01	1.5e−01	8.6e−01	2
3	9.16745e−02	2.15831	2.7e−01	1.0e−01	2.9e−01	2
4	7.62531e−02	2.26056	4.7e−02	1.4e−01	1.5e−01	3
5	3.87835e−03	2.14905	2.7e−03	1.4e−02	1.4e−02	2
6	4.00647e−03	2.13521	3.0e−03	7.8e−03	8.4e−03	4
7	4.36285e−04	2.14167	2.8e−03	2.7e−03	3.9e−03	2
8	4.37573e−04	2.13896	9.5e−04	1.2e−03	1.5e−03	2
9	4.36516e−04	2.14015	2.5e−04	7.9e−04	8.3e−04	5
10	1.83638e−05	2.13950	1.2e−07	3.8e−07	4.0e−07	2
11	1.83638e−05	2.13950	9.3e−08	2.9e−07	3.0e−07	8
12	1.82110e−05	2.13950	7.5e−08	6.0e−08	9.6e−08	2
13	1.82110e−05	2.13950	1.1e−08	1.8e−08	2.2e−08	11
14	1.81999e−05	2.13950	6.0e−11	5.3e−12	6.0e−11	15
15	1.81999e−05	2.13950	1.4e−12	4.5e−12	4.7e−12	2
16	1.81999e−05	2.13950	1.2e−12	3.8e−12	4.0e−12	14
17	1.81999e−05	2.13950	1.1e−12	8.3e−13	1.3e−12	2
18	1.81999e−05	2.13950	4.7e−14	2.1e−14	5.1e−14	18

Table 2. Table showing quadratic convergence to the eigenvalue $\lambda = 1.81999e - 05 + 2.13950i$ with a decreasing tolerance for Example 5.1. The last column shows the number of inner iterations it took to satisfy the decreasing inner tolerance $\tau = \min\left\{0.6, 0.6\left\|r_1^{(k)}\right\|\right\}$.

k	α	β	$\left\|w^+ - w\right\|$	$\left\|\begin{bmatrix}\Delta\alpha\\\Delta\beta\end{bmatrix}\right\|$	$\left\|\Delta v\right\|$	d
0	0.00000e+00	2.50000	3.9e+00	7.3e−01	4.0e+00	2
1	1.81138e−01	1.78811	1.9e+00	2.3e−01	1.9e+00	2
2	1.14991e−01	2.01153	8.4e−01	1.5e−01	8.6e−01	2
3	9.16745e−02	2.15831	2.7e−01	8.0e−02	2.8e−01	3
4	1.19965e−02	2.15159	3.8e−02	1.7e−02	4.1e−02	5
5	3.56419e−04	2.13989	7.2e−04	5.4e−04	9.0e−04	7
6	1.79381e−05	2.13950	2.6e−07	3.8e−07	4.6e−07	13
7	1.81999e−05	2.13950	4.0e−14	4.5e−14	6.0e−14	22
8	1.81999e−05	2.13950	2.1e−14	5.2e−14	5.6e−14	22
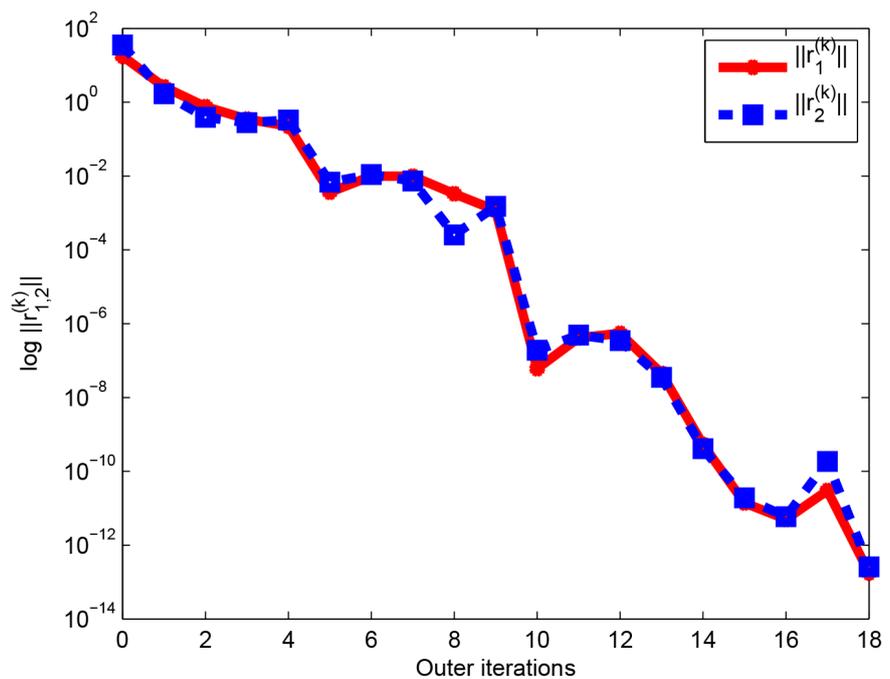
Figure 1. Convergence history of the eigenvalue residuals on Example 5.1 with a fixed tolerance $\tau_{inner} = 0.6$.

were needed to satisfy the decreasing inner tolerance as against the fixed tolerance. From **Table 1**, we observed superlinear convergence in columns five and six.

Figure 1 and **Figure 2** shows a plot of the norm of the eigenvalue residuals against the outer iterations with fixed and decreasing inner tolerances respectively. While the norm of the eigenvalue residuals decayed almost
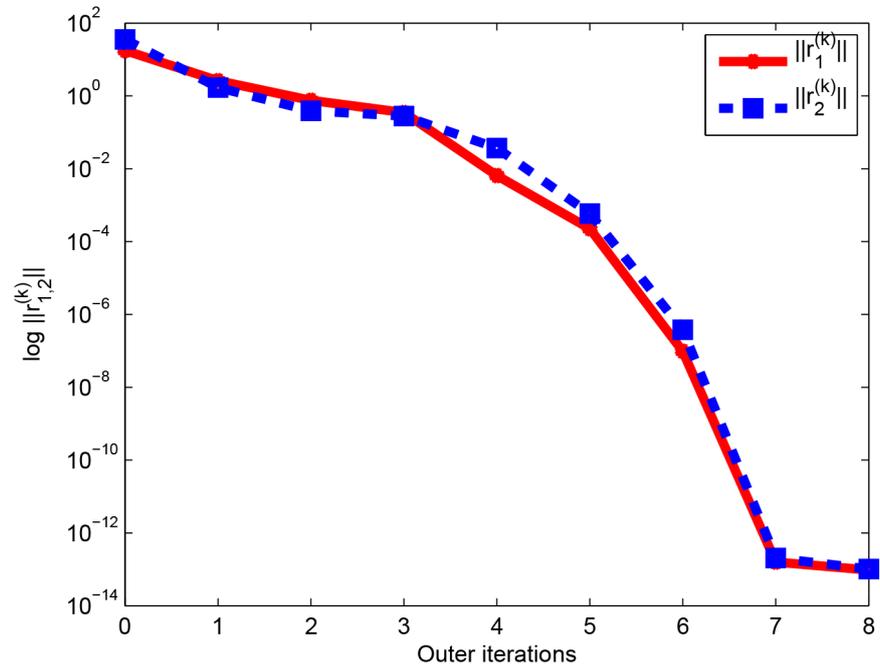
Figure 2. Convergence history of the eigenvalue residuals on Example 5.1 with a decreasing tolerance $\tau_{\text{inner}} = \min\left\{0.6, 0.6\left\|r_1^{(k)}\right\|\right\}$.

superlinearly in **Figure 1**, we observed a quadratic decrease in the norm of the eigenvalue residuals in **Figure 2**. It is quite surprising to see that **Algorithm 3** works because M is singular at the root, which means we solved a singular system at the root.

6. Conclusions

While Ruhe ([2] Section 3) used the normalisation $c^H z = 1$ and solved the resulting $(n+1)$ by $(n+1)$ nonlinear system of equations to obtain $[z, \lambda]^{\mathrm{T}}$. We have been able to show that, with the addition of the non differentiable normalisation $z^H z = 1$, it is still possible to convert the resulting system of under-determined nonlinear equations into a square one.

Nevertheless, in this work, we obtained **Algorithm 1** which consists of a combination of a $2n$-by-$2n$ system of equations (which is the same as those in [13]) and a 2-by-2 system. A numerical example show that using an LU-solver on the one hand and preconditioned GMRES as an inexact solver on the other hand to solve the large sparse $2n$-by-$2n$ system of Equations (28), followed by the 2 by 2 system in each case, give similar results in the limit. By and large, the algorithms presented in this paper relies on good initial guesses to the desired eigenpair of interest.

Acknowledgements

References

[1] Stewart, G.W. (2001) Matrix Algorithms. Vol. II, Eigensystems, SIAM.

[2] Ruhe, A. (1973) Algorithms for the Nonlinear Eigenvalue Problem. *SIAM Journal on Matrix Analysis and Applications*, **10**, 674-689. https://doi.org/10.1137/0710059

[3] Keller, H.B. (1977) Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems. In: Rabinowitz, P., Ed., *Applications of Bifurcation Theory*, Academic Press, New York, 359-384.

[4] Deuflhard, P. (2004) Newton Methods for Nonlinear Problems. Springer, 174-175.

[5] Akinola, R.O. (2014) Theoretical Expression for the Nullvector of the Jacobian: Inverse Iteration with a Complex Shift. *International Journal of Innovation in Science and Mathematics*, **2**, 367-371.

[6] Akinola, R.O. and Spence, A. (2014) Two-Norm Normalization for the Matrix Pencil: Inverse Iteration with a Complex Shift. *International Journal of Innovation in Science and Mathematics*, **2**, 435-439.

[7] Akinola, R.O. and Spence, A. (2015) Numerical Computation of the Complex Eigenvalues of a Matrix by solving a Square System of Equations. *Journal of Natural Sciences Research*, **5**, 144-157.

[8] Akinola, R.O. (2015) Computing the Complex Eigenpair of a Large Sparse Matrix in Complex Arithmetic. *International Journal of Pure & Engineering Mathematics* (*IJPEM*), **3**, 137-158.

[9] Saad, Y. and Schultz, M.H. (1986) GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM Journal on Scientific and Statistical Computing*, **7**, 856-869. https://doi.org/10.1137/0907058

[10] Boyd, S. (2008/2009) Lecture 8: Least Norm Solutions of Under-Determined Equations, EE263 Autumn.

[11] Freitag, M.A. and Spence, A. (2009) The Calculation of the Distance to Instability by the Computation of a Jordan Block, Linear and Nonlinear Eigen Problems for PDEs, 274-275.

[12] Biosvert, B., Pozo, R., Remington, K., Miller, B. and Lipman, R. Matrix Market. http://math.nist.gov/MatrixMarket/

[13] Parlett, B.N. and Saad, Y. (1987) Complex Shift and Invert Strategies for Real Matrices. *Linear Algebra and Its Applications*, 575-595.