Journal of Natural Sciences Research
ISSN 2224-3186 (Paper) ISSN 2225-0921 (Online)
Vol.5, No.5, 2015

Numerical Computation of the Complex Eigenvalues of a Matrix by solving a Square System of Equations

R. O. AKINOLA* and A. SPENCE†

February 10, 2015

Abstract

It is well known that if the largest or smallest eigenvalue of a matrix has been computed by some numerical algorithms and one is interested in computing the corresponding eigenvector, one method that is known to give such good approximations to the eigenvector is inverse iteration with a shift. For complex eigenpairs, instead of using Ruhe's normalization, we show that the natural two norm normalization for the matrix pencil, yields an underdetermined system of equation and by adding an extra equation, the augmented system becomes square which can be solved by LU factorization at a cheaper rate and quadratic convergence is guaranteed. While the underdetermined system of equations can be solved using QR factorization as shown in an earlier work by the same authors, converting it to a square system of equations has the added advantage that besides using LU factorization, it can be solved by several approaches including iterative methods. We show both theoretically and numerically that both algorithms are equivalent in the absence of roundoff errors.

1 Introduction

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

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

where $\lambda \in \mathbb{C}$ is the eigenvalue of the pencil $(\mathbf{A}, \mathbf{B})$ and $\mathbf{z}$ its corresponding complex eigenvector. We assume that the eigenpair of interest $(\mathbf{z}, \lambda)$ is algebraically simple, so that $\boldsymbol{\psi}^H$ the corresponding left eigenvector is such that [1, p. 136]

$$\boldsymbol{\psi}^H\mathbf{Bz} \neq 0.$$

By adding the normalisation

$$\mathbf{z}^H\mathbf{Bz} = 1, \tag{2}$$

*Department of Mathematics, University of Jos, P. M. B. 2084, Nigeria. (roakinola@yahoo.com).
†Department of Mathematical Sciences, University of Bath, BA2 7AY, UK.

to (1) and with $\mathbf{v} = [\mathbf{z}^T, \lambda]$, the combined system of equations can be expressed in the form $\mathbf{F}(\mathbf{v}) = \mathbf{0}$ as

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B})\mathbf{z} \\ -\frac{1}{2}\mathbf{z}^H\mathbf{B}\mathbf{z} + \frac{1}{2} \end{bmatrix} = \mathbf{0}. \tag{3}$$

Note that $\mathbf{z}^H\mathbf{B}\mathbf{z}$ is real since $\mathbf{B}$ is symmetric and positive definite. This results in solving a system of n complex and one real nonlinear equation for the $(n+1)$ complex unknowns $\mathbf{v} = [\mathbf{z}, \lambda]^T$. Note that, if $\mathbf{z}$ from $(\mathbf{z}, \lambda)$ solves (3), then so does $e^{i\theta}\mathbf{z}$ for any $\theta \in [0, 2\pi)$. Hence, (3) does not have a unique solution. Another drawback of the normalisation (2) is that $\bar{\mathbf{z}}$ in $\mathbf{z}^H\mathbf{B}\mathbf{z} = \bar{\mathbf{z}}^T\mathbf{B}\mathbf{z}$ is not differentiable. Therefore, we cannot just differentiate (3) and apply the standard Newton's method. In this article, we shall show how these drawbacks can be overcome, at least for the $\mathbf{B} = \mathbf{I}$ case.

Parlett and Saad in [2], studied inverse iteration with a complex shift $\sigma = \alpha + i\beta$ where α and β are real. They showed that by replacing the shifted complex system $(\mathbf{A} - \sigma\mathbf{B})\boldsymbol{\phi} = \mathbf{B}\boldsymbol{\varphi}$, with a real one, the size of the problem is doubled, where $\boldsymbol{\varphi} = \boldsymbol{\varphi}_1 + i\boldsymbol{\varphi}_2$, $\boldsymbol{\phi} = \boldsymbol{\phi}_1 + i\boldsymbol{\phi}_2$ for $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \boldsymbol{\phi}_1, \boldsymbol{\phi}_2 \in \mathbb{R}^n$ and $i = \sqrt{-1}$ is the imaginary unit of a complex number. This is because solving a complex linear system of equations takes twice the storage and is roughly three times the cost of solving a real system [3]. When real arithmetic rather than complex arithmetic is used, we lose any band structure in $\mathbf{A}$ and $\mathbf{B}$ [2]. The numerical examples in [2], show linear convergence to the eigenvalue closest to the fixed shift.

Next, Tisseur in [4] considered the symmetric definite generalised eigenvalue problem $\mathbf{A}\boldsymbol{\phi} = \lambda\mathbf{B}\boldsymbol{\phi}$, $\lambda \in \mathbb{R}$ as a special case of (1) where $\mathbf{A}$ is symmetric and $\mathbf{B}$ is symmetric positive definite but with the real normalisation

$$\tau\mathbf{e}_s^T\boldsymbol{\phi} = \tau; \quad \text{for some fixed s},$$

where $\tau = \max(\|\mathbf{A}\|, \|\mathbf{B}\|)$, (see, for example, [4, p. 1049]) and $\mathbf{e}_j$ is the jth column of the identity matrix. The real scalar τ is introduced to scale $\mathbf{F}(\mathbf{w})$ and $\mathbf{F}_{\mathbf{w}}(\mathbf{w})$ when $\mathbf{A}$ and $\mathbf{B}$ are multiplied by a scalar. In this case,

$$\mathbf{F}(\mathbf{w}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B})\boldsymbol{\phi} \\ \tau\mathbf{e}_s^T\boldsymbol{\phi} - \tau \end{bmatrix}, \quad \text{and} \quad \mathbf{F}_{\mathbf{w}}(\mathbf{w}) = \begin{bmatrix} (\mathbf{A} - \lambda\mathbf{B}) & -\mathbf{B}\boldsymbol{\phi} \\ \tau\mathbf{e}_s^T & 0 \end{bmatrix}.$$

Tisseur [4], showed that the Jacobian $\mathbf{F}_{\mathbf{w}}(\mathbf{w})$ above is singular at the root if and only if λ^* is a finite multiple eigenvalue of the pencil $(\mathbf{A}, \mathbf{B})$. The main result in [4] is Theorem 2.4 [4, pp. 1044-1046]. It shows that if the linear system to be solved is not too ill conditioned, the solver is not completely unstable, the Jacobian is approximated accurately enough and we have a good initial guess very close to the solution, then the norm of the residual reduces after one step of Newton's method in floating point arithmetic. The main point is that both [5] and [4] used two different differentiable normalisations, while in this paper we analyse the natural extension of the distance norm, which is a non differentiable normalisation and so leads to interesting theoretical questions.

Our approach for analysing the solution of (3) for $\mathbf{v}$ begins by splitting the eigenpair $(\mathbf{z}, \lambda)$ into their real and imaginary parts: $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$, $\lambda = \alpha + i\beta$ where $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^n$, and $\alpha, \beta \in \mathbb{R}$. After expanding (3), we obtain a real system of $(2n+1)$ under-determined nonlinear equations in $(2n+2)$ real unknowns $\mathbf{v} = [\mathbf{z}_1, \mathbf{z}_2, \alpha, \beta]^T$, and it is natural to use the Gauss-Newton method (see, for example, Deuflhard [6, pp. 222-223]) to obtain a solution. By linearising the system of under-determined nonlinear equations, we obtain a system of under-determined linear equations involving the corresponding Jacobian. This idea has been properly developed in an earlier work by the same authors [8]. Here, we show that by adding an extra equation, the augmented system becomes square which can be solved by LU factorization at a cheaper rate and quadratic convergence is guaranteed. We show both theoretically and numerically that the algorithm presented in [8] and the present work are equivalent in the absence of roundoff errors. The key result in this paper is Theorem 3.1 and Algorithm 1 is given. Throughout this paper, $\|.\| = \|.\|_2$.

2 Under-determined system of linear Equations for the computation of the complex eigenpair of $(\mathbf{A}, \mathbf{B})$

In this section, we will expand the system of n complex and one real nonlinear equations in $(n+1)$ complex unknowns (3) by writing $\mathbf{z}$ and λ as $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$ 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 $\mathbf{z}^H\mathbf{B}\mathbf{z} = 1$ gives only one real equation, since $\mathbf{B}$ is symmetric positive definite, while $(\mathbf{A} - \lambda\mathbf{B})\mathbf{z} = \mathbf{0}$ results in $2n$ real equations. This results in a real $(2n+1)$ under-determined system of nonlinear equations in $(2n+2)$ real unknowns. This will then be followed by presenting the real under-determined system of nonlinear equations and an explicit expression for its Jacobian. If you have read the previous two papers in this series, skip this section and the next and move to the section 3, otherwise keep reading.

If we let $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$ and $\lambda = \alpha + i\beta$, then the nonlinear system of equations (3) can be written as

$$\begin{aligned}(\mathbf{A} - \lambda\mathbf{B})\mathbf{z} &= \left[\mathbf{A} - (\alpha + i\beta)\mathbf{B}\right](\mathbf{z}_1 + i\mathbf{z}_2) \\ &= (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 + i[(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1], \end{aligned} \tag{4}$$

and

$$\mathbf{z}^H\mathbf{B}\mathbf{z} = \mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2.$$

Hence, (2) implies

$$-\frac{1}{2}\mathbf{z}^H\mathbf{B}\mathbf{z} + \frac{1}{2} = -\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} = \mathbf{0}.$$

Since $(\mathbf{A} - \lambda\mathbf{B})\mathbf{z} = \mathbf{0}$, we equate the real and imaginary parts of (4) to zero and obtain the $2n$ real equations

$$(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 = \mathbf{0}, \quad \text{and} \quad (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 = \mathbf{0}.$$

This means, $\mathbf{F}(\mathbf{v})$ consists of the $2n$ real equations arising from (4) and one real equation $-\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} = 0$;

$$\mathbf{F}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B}\mathbf{z}_1 + (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 \\ -\frac{1}{2}(\mathbf{z}_1^T\mathbf{B}\mathbf{z}_1 + \mathbf{z}_2^T\mathbf{B}\mathbf{z}_2) + \frac{1}{2} \end{bmatrix} = \mathbf{0}, \tag{5}$$

where $\mathbf{F} : \mathbb{R}^{(2n+2)} \rightarrow \mathbb{R}^{(2n+1)}$. The Jacobian, $\mathbf{F_v}(\mathbf{v})$ of $\mathbf{F}(\mathbf{v})$ with $\mathbf{v} = [\mathbf{z}_1, \mathbf{z}_2, \alpha.\beta]^T$ has the following explicit expression

$$\mathbf{F_v}(\mathbf{v}) = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} & -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) & -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \\ -(\mathbf{B}\mathbf{z}_1)^T & -(\mathbf{B}\mathbf{z}_2)^T & 0 & 0 \end{bmatrix}, \tag{6}$$

and is a $(2n+1)$ by $(2n+2)$ real matrix. We define the real $2n$ by $2n$ matrix $\mathbf{M}$ as

$$\mathbf{M} = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) \end{bmatrix}. \tag{7}$$

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

$$\mathbf{N} = \begin{bmatrix} -\mathbf{B}\mathbf{z}_1 & \mathbf{B}\mathbf{z}_2 \\ -\mathbf{B}\mathbf{z}_2 & -\mathbf{B}\mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{w}_1 \end{bmatrix}, \tag{8}$$

consisting of the product of $\mathbf{B}_2 = \begin{bmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{bmatrix}$ and the matrix of right nullvectors (given in the next equation) of $\mathbf{M}$ at the root, where

$$\mathbf{w} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix}, \qquad \mathbf{w}_1 = \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix}, \tag{9}$$

and $\mathbf{O}$ is the n by n zero matrix. The Jacobian (6) can be rewritten in the following partitioned form

$$\mathbf{F_v}(\mathbf{v}) = \begin{bmatrix} \mathbf{M} & -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{w}_1 \\ -(\mathbf{B}_2\mathbf{w})^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{M} & \mathbf{N} \\ -(\mathbf{B}_2\mathbf{w})^T & \mathbf{0}^T \end{bmatrix}, \tag{10}$$

with $\mathbf{M}$, $\mathbf{N}$ defined in (7) and (8) respectively. Note that because at the root,

$$\begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2 \\ (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 \end{bmatrix} = \mathbf{0},$$

this implies that $\mathbf{w}$ or its nonzero scalar multiple is a right nullvector of $\mathbf{M}$. In the same vein, we find

$$\begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B}) & \beta\mathbf{B} \\ -\beta\mathbf{B} & (\mathbf{A} - \alpha\mathbf{B}) \end{bmatrix} \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix} = \begin{bmatrix} (\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_2 - \beta\mathbf{B}\mathbf{z}_1 \\ -\{(\mathbf{A} - \alpha\mathbf{B})\mathbf{z}_1 + \beta\mathbf{B}\mathbf{z}_2\} \end{bmatrix} = \mathbf{0},$$

and $\mathbf{w}_1$ or its nonzero scalar multiple is also a right nullvector of $\mathbf{M}$ at the root.

Theorem 2.1. *Assume that the eigenpair $(\mathbf{z}, \lambda)$ of the pencil $(\mathbf{A}, \mathbf{B})$ is algebraically simple. If $\mathbf{z}_1$ and $\mathbf{z}_2$ are nonzero vectors, then $\boldsymbol{\phi} = \{\tau[\mathbf{w}_1^T, 0, 0], \tau \in \mathbb{R}\}$ is the eigenspace corresponding to the zero eigenvalue of $\mathbf{F_v}(\mathbf{v})$ at the root.*

Proof: See [8] ■

Corollary 2.1. : *If the eigenpair $(\mathbf{z}, \lambda)$ of $(\mathbf{A}, \mathbf{B})$ is algebraically simple, then the Jacobian $\mathbf{F_v}(\mathbf{v})$ in* (10) *is of full rank at the root.*

Proof. See [8]. □

Next, in order to solve the under-determined system of nonlinear equations (5), we need to linearize $\mathbf{F}(\mathbf{v}) = \mathbf{0}$. After linearizing $\mathbf{F}(\mathbf{v}) = \mathbf{0}$, we have to solve the following under-determined linear system of equations

$$\mathbf{F_v}(\mathbf{v}^{(k)})\Delta\mathbf{v}^{(k)} = -\mathbf{F}(\mathbf{v}^{(k)}). \tag{11}$$

The following result which we state without a proof will be used in the later part of this paper.

Lemma 2.1. : *[7, p. 6] Let* $\mathbf{F_w}(\mathbf{w})$ *be of full rank. If*

$$\mathbf{F_w}(\mathbf{w})\Delta\mathbf{w} = \mathbf{F}(\mathbf{w}),$$

is an under-determined linear system of equations, then its least squares solution

$$\Delta\mathbf{w} = -\mathbf{F_w}(\mathbf{w})^T[\mathbf{F_w}(\mathbf{w})\mathbf{F_w}(\mathbf{w})^T]^{-1}\mathbf{F}(\mathbf{w}),$$

is orthogonal to the nullspace of $\mathbf{F_w}(\mathbf{w})$.

Proof. See [8] □

Next, we state the following result which was proved in [9] and shows that the solution $\Delta\mathbf{v}^{(k)}$ obtained by solving the underdetermined system of nonlinear equations (11) is equivalent to those obtained by solving a square, augmented linear system.

Lemma 2.2. : *Let* $\boldsymbol{n}^{(k)}$ *be the exact nullvector of* $\mathbf{F_v}(\mathbf{v}^{(k)})$. *The solution* $\Delta\mathbf{v}^{(k)}$ *can be obtained via:*

(a). solving the under-determined linear system of $(2n+1)$ *real equations for the* $(2n+2)$ *real unknowns* $\Delta\mathbf{v}^{(k)}$ (11) *and updating* $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$, *or*

(b). solving the square linear system of $(2n+2)$ *real equations*

$$\begin{bmatrix} \mathbf{F_v}(\mathbf{v}^{(k)}) \\ \boldsymbol{n}^{(k)T} \end{bmatrix} \Delta\mathbf{v}^{(k)} = -\begin{bmatrix} \mathbf{F}(\mathbf{v}^{(k)}) \\ 0 \end{bmatrix}. \tag{12}$$

and updating $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$. *(Here, we neglect round off errors).*

Proof. See [9]. □

The next section contains useful theoretical expressions that will help us in section 3 of this paper.

2.1 Theoretical form for the Nullvector of the Jacobian (6)

In the proof of Lemma 2.2 at the tail end of last section, we made use of the exact nullvector (which we do not compute in practice) of the Jacobian (6). In this section, we give a theoretical expression for the exact nullvector of the Jacobian (6) when not at the root. To do this, we rewrite the under-determined linear system of equations (11) in a compressed form, present two important theoretical relationships: (18) and (19) for the exact nullvector of the Jacobian.

Note that the matrix $\mathbf{M}$ defined by (7) is singular at the root. However, this section is anchored on the assumption that when $\mathbf{v}$ is not at the root, $\mathbf{M}$ is nonsingular. First, we define the $2n$ by $2n$ matrix $\mathbf{J}$ as (see, for example [10])

$$\mathbf{J} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix}, \tag{13}$$

and note that

$$\mathbf{Jw} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{z}_2 \\ -\mathbf{z}_1 \end{bmatrix} = \mathbf{w}_1, \tag{14}$$

defined by (9). The matrix $\mathbf{J}$ satisfies the following properties:

1. $\mathbf{J}^T = -\mathbf{J}$.
2. $\mathbf{J}^T\mathbf{J} = \mathbf{I}_{2n}$, where $\mathbf{I}_{2n}$ is the $2n$ by $2n$ identity matrix.
3. $\mathbf{J}^2 = -\mathbf{I}_{2n}$.
4. $\mathbf{J}$ commutes with $\mathbf{M}$ and $\mathbf{B}_2$, *i.e.*, $\mathbf{JM} = \mathbf{MJ}$ and $\mathbf{JB}_2 = \mathbf{B}_2\mathbf{J}$.
5. For $\mathbf{w} \in \mathbb{R}^{2n}$, $\mathbf{w}^T\mathbf{B}_2\mathbf{Jw} = \mathbf{w}^T\mathbf{JB}_2\mathbf{w} = 0$.
6. Let $\mathbf{u}$ be an unknown vector that solves $\mathbf{Mu} = \mathbf{B}_2\mathbf{w}$. By premultiplying both sides by $\mathbf{J}$ we obtain $\mathbf{JMu} = \mathbf{JB}_2\mathbf{w}$ and hence $\mathbf{MJu} = \mathbf{JB}_2\mathbf{w}$ by the commutativity of $\mathbf{M}$ and $\mathbf{J}$. Therefore,
$$\mathbf{Mu} = \mathbf{B}_2\mathbf{w}, \quad \text{implies} \quad \mathbf{M}(\mathbf{Ju}) = \mathbf{JB}_2\mathbf{w}. \tag{15}$$

The equation $\mathbf{Mu} = \mathbf{B}_2\mathbf{w}$ stems from expanding the shifted system $(\mathbf{A} - \sigma\mathbf{B})\mathbf{y} = \mathbf{Bz}$, into its real and imaginary parts as in [2] for $\sigma = \alpha + i\beta$ and $\mathbf{z} = \mathbf{z}_1 + i\mathbf{z}_2$. For ease of notation and for the rest of this section, we shall drop the superscripts $^{(k)}$ and write $\mathbf{w}^+ = \mathbf{w} + \Delta\mathbf{w}$ where $\mathbf{w}^+ = \mathbf{w}^{(k+1)}$, replace $\mathbf{w}^{(k)}$ and $[\Delta\mathbf{z}_1^{(k)^T}, \Delta\mathbf{z}_2^{(k)^T}]$ with $\mathbf{w}$ and $\Delta\mathbf{w}$ respectively *e.t.c.* As earlier stated, we assume that the $2n$ by $2n$ matrix $\mathbf{M}$ is nonsingular except at the root. For the rest of this section, our aim is to give an explicit theoretical expression for the nullvector of (6).

Let the exact nullvector $\boldsymbol{n}$ of

$$\mathbf{F_v}(\mathbf{v}) = \begin{bmatrix} \mathbf{M} & -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{Jw} \\ -(\mathbf{B}_2\mathbf{w})^T & 0 & 0 \end{bmatrix},$$

be defined as $\boldsymbol{n} = [\boldsymbol{n}_{\mathbf{w}}^T, \boldsymbol{n}_\alpha, \boldsymbol{n}_\beta]$, where $\boldsymbol{n}_{\mathbf{w}} \in \mathbb{R}^{2n}$, $\boldsymbol{n}_\alpha$ and $\boldsymbol{n}_\beta$ are real scalars, $\mathbf{Jw}$ and $\mathbf{M}$ are defined respectively by (14) and (7). Hence,

$$\begin{bmatrix} \mathbf{M} & -\mathbf{B}_2\mathbf{w} & \mathbf{B}_2\mathbf{Jw} \\ -(\mathbf{B}_2\mathbf{w})^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{n}_{\mathbf{w}} \\ \boldsymbol{n}_\alpha \\ \boldsymbol{n}_\beta \end{bmatrix} = \mathbf{0},$$

then after expanding the matrix-vector multiplication, we obtain

$$\mathbf{M}\boldsymbol{n}_{\mathbf{w}} - \boldsymbol{n}_\alpha\mathbf{B}_2\mathbf{w} + \boldsymbol{n}_\beta(\mathbf{B}_2\mathbf{Jw}) = \mathbf{0} \tag{16}$$
$$\mathbf{w}^T\mathbf{B}_2\boldsymbol{n}_{\mathbf{w}} = 0. \tag{17}$$

From (16), $\mathbf{M}\boldsymbol{n}_{\mathbf{w}} = \boldsymbol{n}_\alpha \mathbf{B}_2\mathbf{w} - \boldsymbol{n}_\beta(\mathbf{B}_2\mathbf{J}\mathbf{w})$, using the fact that $\mathbf{J}$ commutes with $\mathbf{B}_2$ and $\mathbf{M}$, and using (15) with $\mathbf{B}_2 = \mathbf{I}_{2n}$ we obtain

$$\boldsymbol{n}_{\mathbf{w}} = \boldsymbol{n}_\alpha \mathbf{u} - \boldsymbol{n}_\beta \mathbf{J}\mathbf{u}. \tag{18}$$

Since $\mathbf{w}$ is $\mathbf{B}_2$-orthogonal to $\boldsymbol{n}_{\mathbf{w}}$ by virtue of (17), taking the $\mathbf{B}_2$-inner product of both sides of the above with $\mathbf{w}$ yields

$$\mathbf{w}^T\mathbf{B}_2\boldsymbol{n}_{\mathbf{w}} = \boldsymbol{n}_\alpha(\mathbf{w}^T\mathbf{B}_2\mathbf{u}) - \boldsymbol{n}_\beta(\mathbf{w}^T\mathbf{B}_2\mathbf{J}\mathbf{u}) = \mathbf{0}.$$

We may choose

$$\boldsymbol{n}_\alpha = \mathbf{w}^T\mathbf{B}_2\mathbf{J}\mathbf{u}, \quad \text{and} \quad \boldsymbol{n}_\beta = \mathbf{w}^T\mathbf{B}_2\mathbf{u}, \tag{19}$$

since we never normalise $\boldsymbol{n}$. Hence, $\boldsymbol{n}_{\mathbf{w}}$ is given by (18) with $\boldsymbol{n}_\alpha$ and $\boldsymbol{n}_\beta$ by (19). So we have a formula for $\boldsymbol{n}_{\mathbf{w}}$ in terms of $\mathbf{w}$ and $\mathbf{u}$ obtained from (15). Therefore,

$$\boldsymbol{n} = [\boldsymbol{n}_{\mathbf{w}}^T, \boldsymbol{n}_\alpha, \boldsymbol{n}_\beta] = [(\boldsymbol{n}_\alpha\mathbf{u} - \boldsymbol{n}_\beta\mathbf{J}\mathbf{u})^T, (\mathbf{w}^T\mathbf{B}_2\mathbf{J}\mathbf{u}), (\mathbf{w}^T\mathbf{B}_2\mathbf{u})].$$

We emphasise that in practice, we would never compute the solution of (15). It will be used for purely theoretical purposes since we know that the Gauss-Newton solution, $\Delta\mathbf{v}$, is orthogonal to $\boldsymbol{n}$.

3 Square System of Equations for The Numerical Computation of the Complex Eigenvalues of a Matrix for $\mathbf{B} = \mathbf{I}$

In the preceding section, we presented two main important theoretical relationships, (18) and (19). In this section, we will make use of these relationships in our discussion but only in the special case in which $\mathbf{B} = \mathbf{I}$. Moreover, in Section (2), we saw that the solution to the under-determined system of nonlinear equations (5) for the numerical computation of the complex eigenpair $(\mathbf{z}, \lambda)$ of the pencil $(\mathbf{A}, \mathbf{B})$ can be solved by the Gauss-Newton method via QR factorization. It was also stated in Lemma 2.1 that the minimum norm solution to the resulting linear system of equations is orthogonal to the nullspace. However, in Section 2, we used the result of Lemma 2.1 to add an extra equation to the under-determined linear system of equations, so as to obtain a square one. This is because, at each iteration of the computation, ${\boldsymbol{n}^{(k)}}^T\Delta\mathbf{v}^{(k)} = 0$ and so it does not change the solution, even though the square linear system of equations gives a unique solution because the augmented Jacobian is nonsingular.

Nevertheless, as mentioned in the last section, we would never compute $\boldsymbol{n}$ in practice, but Theorem 2.1 guarantees the existence of a unique nullvector $\boldsymbol{\phi}$ at the root. We will use $\boldsymbol{\phi}^{(k)}$ defined by $\boldsymbol{\phi}^{(k)} = [\mathbf{z}_2^{(k)}, -\mathbf{z}_1^{(k)}, 0, 0]$ as an approximation to the exact nullvector $\boldsymbol{n}$ and show that the solution obtained by solving (11) is equivalent to the solution obtained by solving

$$\begin{bmatrix} \mathbf{F}_{\mathbf{v}}(\mathbf{v}^{(k)}) \\ {\boldsymbol{\phi}^{(k)}}^T \end{bmatrix} \Delta\mathbf{v}^{(k)} = -\begin{bmatrix} \mathbf{F}(\mathbf{v}^{(k)}) \\ 0 \end{bmatrix}, \tag{20}$$

in the absence of round off errors. To do this, we will show that ${\boldsymbol{\phi}^{(k)}}^T\Delta\mathbf{v}^{(k)} = 0$ for each k, where $\Delta\mathbf{v}^{(k)}$ is given by (11) and this is the key result in this section.

This section is structured as follows, we begin by adding the extra equation ${\boldsymbol{n}^{(k)}}^T\Delta\mathbf{v}^{(k)} = 0$ to (11) in order to obtain the square linear system of equations (12). The main result in this section is Theorem 3.1, and Algorithm 1 is presented for computing the algebraically simple eigenpair of $\mathbf{A}$. Note that since $\mathbf{M}$ has been shown to be singular at the root in section 2, this section is anchored on the assumption that when $\mathbf{v}$ is not at the root, $\mathbf{M}$ is nonsingular, but this is acceptable since we use the construction here to prove a theoretical result about the correction $\Delta\mathbf{v}^{(k)}$ while not at the root.

Consider the problem of solving the under-determined linear system of equations (11) for the $(2n+2)$ real unknowns $\Delta\mathbf{v} = [\Delta\mathbf{w}^T, \Delta\alpha, \Delta\beta]$. It was stated in Lemma 2.1 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,

$$0 = \boldsymbol{n}^T\Delta\mathbf{v} = \boldsymbol{n}_{\mathbf{w}}^T\Delta\mathbf{w} + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta, \tag{21}$$

where we have dropped the superscript $^{(k)}$ in $\alpha, \beta, \boldsymbol{n}, \mathbf{w}$ and $\mathbf{v}$. We begin by writing the linear system of equations (11) in expanded form as

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} -\mathbf{Mw} \\ \frac{1}{2}(\mathbf{w}^T\mathbf{w} - 1) \end{bmatrix}, \tag{22}$$

or,

$$\begin{aligned} \mathbf{M}\Delta\mathbf{w} - \Delta\alpha\mathbf{w} + \Delta\beta\mathbf{Jw} &= -\mathbf{Mw} \\ -\mathbf{w}^T\Delta\mathbf{w} &= \frac{1}{2}\mathbf{w}^T\mathbf{w} - \frac{1}{2}. \end{aligned}$$

After rearrangement, the first equation reduces to

$$\mathbf{Mw}^+ - \Delta\alpha\mathbf{w} + \Delta\beta\mathbf{Jw} = \mathbf{0}. \tag{23}$$

By multiplying both sides of the second equation by 2, we obtain:

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

This in turn reduces to

$$\mathbf{w}^T(\mathbf{w} + 2\Delta\mathbf{w}) = 1. \tag{24}$$

Since $\mathbf{w}^+ = \mathbf{w} + \Delta\mathbf{w}$, $2\Delta\mathbf{w} = 2\mathbf{w}^+ - 2\mathbf{w}$ and $\mathbf{w} + 2\Delta\mathbf{w} = 2\mathbf{w}^+ - \mathbf{w}$, then $\mathbf{w}^T(\mathbf{w} + 2\Delta\mathbf{w}) = \mathbf{w}^T(2\mathbf{w}^+ - \mathbf{w}) = 2\mathbf{w}^T\mathbf{w}^+ - \mathbf{w}^T\mathbf{w}$. Consequently,

$$\mathbf{w}^T\mathbf{w}^+ = \frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1). \tag{25}$$

The combined set of equations (23) and (25), which is the simplified form of (22), can be expressed as:

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{w}^+ \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1) \end{bmatrix}. \tag{26}$$

Now, if we expand along the first row of (26), then

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

This means that we could solve (26) by solving

$$\mathbf{Mu} = \mathbf{w}, \quad \text{and} \quad \mathbf{MJu} = \mathbf{Jw},$$

(by Property 6 of $\mathbf{J}$ after (14)), for $\mathbf{u}$, after which the solution of (27) is given by

$$\mathbf{w}^{+} = \Delta\alpha\mathbf{u} - \Delta\beta\mathbf{Ju}. \tag{28}$$

If we add the nullvector $\boldsymbol{n}$ to the last row of (22) with $\mathbf{B} = \mathbf{I}$ and using (21), then

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \\ \boldsymbol{n}_{\mathbf{w}}^T & \boldsymbol{n}_\alpha & \boldsymbol{n}_\beta \end{bmatrix} \begin{bmatrix} \Delta\mathbf{w} \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} -\mathbf{Mw} \\ \frac{1}{2}(\mathbf{w}^T\mathbf{w} - 1) \\ 0 \end{bmatrix}.$$

One can also add $\boldsymbol{n}$ to the last row of (26) to yield

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \\ \boldsymbol{n}_{\mathbf{w}}^T & \boldsymbol{n}_\alpha & \boldsymbol{n}_\beta \end{bmatrix} \begin{bmatrix} \mathbf{w}^{+} \\ \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ -\frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1) \\ \boldsymbol{n}_{\mathbf{w}}^T\mathbf{w} \end{bmatrix}. \tag{29}$$

By expanding the middle row of (29), $\mathbf{w}^T\mathbf{w}^{+} = \frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1)$. But from (28), $\mathbf{w}^{+} = \Delta\alpha\mathbf{u} - \Delta\beta\mathbf{Ju}$. This implies that, by taking the inner product of both sides with $\mathbf{w}$, yields

$$\mathbf{w}^T\mathbf{w}^{+} = \Delta\alpha(\mathbf{w}^T\mathbf{u}) - \Delta\beta(\mathbf{w}^T\mathbf{Ju}) = \frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1).$$

Using the definition (19) for $\boldsymbol{n}_\alpha$ and $\boldsymbol{n}_\beta$ with $\mathbf{B} = \mathbf{I}$, we obtain

$$\boldsymbol{n}_\beta\Delta\alpha - \boldsymbol{n}_\alpha\Delta\beta = \frac{1}{2}(\mathbf{w}^T\mathbf{w} + 1), \tag{30}$$

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. Note that by using $\boldsymbol{n}_{\mathbf{w}} = \boldsymbol{n}_\alpha\mathbf{u} - \boldsymbol{n}_\beta\mathbf{Ju}$, and (19) we can simplify

$$\begin{aligned} \boldsymbol{n}_{\mathbf{w}}^T\mathbf{w} &= \boldsymbol{n}_\alpha\mathbf{u}^T\mathbf{w} - \boldsymbol{n}_\beta\mathbf{u}^T\mathbf{J}^T\mathbf{w} \\ &= \boldsymbol{n}_\alpha\mathbf{u}^T\mathbf{w} + \boldsymbol{n}_\beta\mathbf{u}^T\mathbf{Jw} \\ &= (\mathbf{w}^T\mathbf{Ju})(\mathbf{u}^T\mathbf{w}) + (\mathbf{w}^T\mathbf{u})(\mathbf{u}^T\mathbf{Jw}) \\ &= -(\mathbf{w}^T\mathbf{J}^T\mathbf{u})(\mathbf{u}^T\mathbf{w}) + (\mathbf{w}^T\mathbf{u})(\mathbf{u}^T\mathbf{Jw}) \\ &= -[(\mathbf{Jw})^T\mathbf{u}](\mathbf{w}^T\mathbf{u}) + (\mathbf{w}^T\mathbf{u})[\mathbf{u}^T(\mathbf{Jw})] \\ &= -(\mathbf{w}_1^T\mathbf{u})(\mathbf{w}^T\mathbf{u}) + (\mathbf{w}^T\mathbf{u})(\mathbf{u}^T\mathbf{w}_1) \\ &= 0. \end{aligned}$$

Now, after expanding along the third row of (29), we have

$$\begin{aligned} \boldsymbol{n}_{\mathbf{w}}^T\mathbf{w}^{+} + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta &= \boldsymbol{n}_{\mathbf{w}}^T(\mathbf{w} + \Delta\mathbf{w}) + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta \\ &= \boldsymbol{n}_{\mathbf{w}}^T\mathbf{w} + \underbrace{(\boldsymbol{n}_{\mathbf{w}}^T\Delta\mathbf{w} + \boldsymbol{n}_\alpha\Delta\alpha + \boldsymbol{n}_\beta\Delta\beta)}_{=0} \\ &= \boldsymbol{n}_{\mathbf{w}}^T\mathbf{w} \\ &= 0. \end{aligned}$$

If we substitute the expression (18) for $n_{\mathbf{w}}$ and (28) for $\mathbf{w}^+$ into the left hand side, then one obtains

$$\begin{aligned} 0 &= n_{\mathbf{w}}^T \mathbf{w}^+ + n_\alpha \Delta\alpha + n_\beta \Delta\beta \\ &= \left[n_\alpha \mathbf{u}^T - n_\beta (\mathbf{J}\mathbf{u})^T\right]\left[\Delta\alpha \mathbf{u} - \Delta\beta \mathbf{J}\mathbf{u}\right] + n_\alpha \Delta\alpha + n_\beta \Delta\beta. \end{aligned} \tag{31}$$

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

$$\begin{aligned} \left[n_\alpha \mathbf{u}^T - n_\beta (\mathbf{J}\mathbf{u})^T\right](\Delta\alpha \mathbf{u} - \Delta\beta \mathbf{J}\mathbf{u}) &= n_\alpha \Delta\alpha \mathbf{u}^T \mathbf{u} + n_\beta \Delta\beta \mathbf{u}^T \mathbf{J}^T \mathbf{J}\mathbf{u} \\ &= n_\alpha \Delta\alpha \|\mathbf{u}\|^2 + n_\beta \Delta\beta \|\mathbf{u}\|^2 \\ &= \left(n_\alpha \Delta\alpha + n_\beta \Delta\beta\right)\|\mathbf{u}\|^2. \end{aligned}$$

Consequently, (31) becomes

$$\left(n_\alpha \Delta\alpha + n_\beta \Delta\beta\right)\|\mathbf{u}\|^2 + n_\alpha \Delta\alpha + n_\beta \Delta\beta = (1 + \|\mathbf{u}\|^2)(n_\alpha \Delta\alpha + n_\beta \Delta\beta) = 0.$$

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

$$n_\alpha \Delta\alpha + n_\beta \Delta\beta = 0. \tag{32}$$

We combine the two equations (30) and (32) below

$$\begin{bmatrix} n_\beta & -n_\alpha \\ n_\alpha & n_\beta \end{bmatrix} \begin{bmatrix} \Delta\alpha \\ \Delta\beta \end{bmatrix} = \begin{bmatrix} \frac{1}{2}(\mathbf{w}^T \mathbf{w} + 1) \\ 0 \end{bmatrix},$$

and compute $\Delta\alpha$, $\Delta\beta$ simultaneously. The matrix on the left hand side is always nonsingular except at the root (in which case all entries are zero). Observe that

$$\begin{aligned} \mathbf{w}^T \mathbf{J}^T \Delta\mathbf{w} &= -\mathbf{w}^T \mathbf{J} \Delta\mathbf{w} \\ &= -\mathbf{w}^T \mathbf{J}(\mathbf{w}^+ - \mathbf{w}) \\ &= -\mathbf{w}^T \mathbf{J}\mathbf{w}^+ + \mathbf{w}^T \mathbf{J}\mathbf{w} \\ &= -\mathbf{w}^T \mathbf{J}\mathbf{w}^+, \end{aligned}$$

where we have used the fact that $\mathbf{w}^T \mathbf{J}\mathbf{w} = 0$ for all $\mathbf{w}$, so that (32) can now be applied to simplify $\mathbf{w}^T \mathbf{J}^T \Delta\mathbf{w}$ as

$$\begin{aligned} \mathbf{w}^T \mathbf{J}^T \Delta\mathbf{w} &= -\mathbf{w}^T \mathbf{J}\mathbf{w}^+ \\ &= -\mathbf{w}^T \mathbf{J}(\Delta\alpha \mathbf{u} - \Delta\beta \mathbf{J}\mathbf{u}) \\ &= -\mathbf{w}^T (\Delta\alpha \mathbf{J}\mathbf{u} + \Delta\beta \mathbf{u}) \\ &= -\left[\Delta\alpha(\mathbf{w}^T \mathbf{J}\mathbf{u}) + \Delta\beta(\mathbf{w}^T \mathbf{u})\right] \\ &= -\left[n_\alpha \Delta\alpha + n_\beta \Delta\beta\right] \\ &= 0. \end{aligned} \tag{33}$$

Notice that we have used the property $\mathbf{J}^2 = -\mathbf{I}_{2n}$ to arrive at the third to the last step above and the definition (28) for $\mathbf{w}^+$. Therefore, we have proved the key result

$$\mathbf{w}^T \mathbf{J}^T \Delta\mathbf{w} = 0.$$

The above analysis leads to the following fundamental result.

Theorem 3.1. *Let* $\boldsymbol{\phi}^{(k)} = [(\mathbf{Jw})^T, 0, 0]$ *be an approximation to the exact nullvector* $\boldsymbol{n}^{(k)}$ *of the Jacobian*

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

for $k = 0, 1, 2, 3, \cdots$ *.*

(a). The augmented Jacobian matrix

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \\ (\mathbf{Jw})^T & 0 & 0 \end{bmatrix}, \tag{34}$$

is nonsingular at an algebraically simple eigenvalue of $\mathbf{Az} = \lambda \mathbf{z}$*.*

(b). The (unique) solution of

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \\ (\mathbf{Jw})^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \alpha \\ \Delta \beta \end{bmatrix} = \begin{bmatrix} -\mathbf{Mw} \\ \frac{1}{2}(\mathbf{w}^T \mathbf{w} - 1) \\ 0 \end{bmatrix}, \tag{35}$$

is identical to the least squares solution of the under-determined system

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{Jw} \\ -\mathbf{w}^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{w} \\ \Delta \alpha \\ \Delta \beta \end{bmatrix} = \begin{bmatrix} -\mathbf{Mw} \\ \frac{1}{2}(\mathbf{w}^T \mathbf{w} - 1) \end{bmatrix}. \tag{36}$$

Proof:

(a). At the root $\boldsymbol{\phi} = \boldsymbol{n}$ and since the real $(2n+1)$ by $(2n+2)$ Jacobian (6) has been shown to be of full rank in Corrolary 2.1, so adding the $(2n+2)$th row, $\boldsymbol{n}^T$ to the Jacobian (6) increases the row rank by one (since the nullvector, $\boldsymbol{n}$ is orthogonal to every row of $\mathbf{F_v}(\mathbf{v})$). Hence,

$$\text{rank}\left(\begin{bmatrix} \mathbf{F_v}(\mathbf{v}) \\ \boldsymbol{n}^T \end{bmatrix}\right) = 2n + 2.$$

Therefore, the matrix in (34) is nonsingular at the root.

(b). Recall that $\Delta \mathbf{v}^{(k)} = [\Delta \mathbf{w}^T, \Delta \alpha, \Delta \beta]$. By using (33), this implies

$$\boldsymbol{\phi}^{(k)^T} \Delta \mathbf{v}^{(k)} = (\mathbf{Jw})^T \Delta \mathbf{w} = \mathbf{w}^T \mathbf{J}^T \Delta \mathbf{w} = 0.$$

Hence, showing that both (35) and (36) are equivalent for $k = 0, 1, 2, 3, \cdots$ ..

■

The above result means that instead of solving (11) or (36) via QR factorisation at a cost of approximately $\frac{32}{3}n^3$ floating point operations, we could use LU factorisation to solve (35) more efficiently at a cost of approximately $\frac{16}{3}n^3$. We now present Algorithm 1 for computing the algebraically simple complex eigenpair of $\mathbf{A}$.

Stop Algorithm 1 as soon as

$$\|\Delta \mathbf{v}^{(k)}\| \leq \text{tol}.$$

We consider the same example as in [8] with the same starting guesses but with a different algorithm: Algorithm 1. For comparison sake, we present the result table in that paper for ease of reference as follows.

Algorithm 1 Eigenpair Computation using Newton's method

Require: $\mathbf{A}, \mathbf{w}^{(0)} = [\mathbf{z}_1^{(0)}, \mathbf{z}_2^{(0)}], \mathbf{v}^{(0)} = [\mathbf{w}^{(0)}, \alpha^{(0)}, \beta^{(0)}]^T, k_{\max}$ and *tol*.

1: **for** $k = 0, 1, 2, \ldots$ until convergence **do**

2: Compute the LU factorisation of

$$\begin{bmatrix} \mathbf{M} & -\mathbf{w} & \mathbf{J}\mathbf{w} \\ -\mathbf{w}^T & 0 & 0 \\ (\mathbf{J}\mathbf{w})^T & 0 & 0 \end{bmatrix}.$$

3: Form

$$\mathbf{d}^{(k)} = \begin{bmatrix} -\mathbf{M}\mathbf{w} \\ \frac{1}{2}(\mathbf{w}^T\mathbf{w} - 1) \\ 0 \end{bmatrix}.$$

4: Solve the lower triangular system $L\mathbf{c}^{(k)} = \mathbf{d}^{(k)}$ for $\mathbf{c}^{(k)}$.

5: Solve the upper triangular system $U\Delta\mathbf{v}^{(k)} = \mathbf{c}^{(k)}$ for $\Delta\mathbf{v}^{(k)}$.

6: Update $\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \Delta\mathbf{v}^{(k)}$.

7: **end for**

Ensure: $\mathbf{v}^{(k_{\max})} = [\mathbf{w}^{(k_{\max})}, \alpha^{(k_{\max})}, \beta^{(k_{\max})}]^T$.

k	$\alpha^{(k)}$	$\beta^{(k)}$	$\Vert\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\Vert$	$\Vert\lambda^{(k+1)} - \lambda^{(k)}\Vert$	$\Vert\Delta\mathbf{v}^{(k)}\Vert$	$\Vert\mathbf{F}(\mathbf{v}^{(k)})\Vert$
0	0.00000e+00	2.50000	3.8e+00	7.8e-01	3.9e+00	3.6e+01
1	2.34253e-01	1.75371	1.8e+00	2.2e-01	1.8e+00	7.8e+00
2	1.18745e-01	1.94460	8.1e-01	1.4e-01	8.2e-01	1.7e+00
3	4.47044e-02	2.06484	2.5e-01	7.0e-02	2.6e-01	3.4e-01
4	8.82702e-03	2.12479	3.1e-02	1.7e-02	3.5e-02	3.7e-02
5	2.48114e-04	2.13905	4.8e-04	5.2e-04	7.1e-04	7.1e-04
6	1.80714e-05	2.13950	1.2e-07	2.5e-07	2.8e-07	2.8e-07
7	1.81999e-05	2.13950	2.1e-14	2.9e-14	3.6e-14	6.0e-14

Table 1: Values of $\alpha^{(k)}$ and $\beta^{(k)}$ using the algorithm in [8]. Columns 6 and 7 show that the results converged quadratically for $k = 3, 4, 5, 6$ and 7.

Example 3.1. *Consider the 200 by 200 matrix* $\mathbf{A}$ `bwm200.mtx` *from the matrix market library [11]. It is the discretised Jacobian of the Brusselator wave model for a chemical reaction. The resulting eigenvalue problem with* $\mathbf{B} = \mathbf{I}$ *was also studied in [2] and we are interested in finding the rightmost eigenvalue of* $\mathbf{A}$ *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 [2] and took* $\mathbf{z}_1^{(0)} = \mathbf{1}/2\Vert\mathbf{1}\Vert$ *and* $\mathbf{z}_2^{(0)} = \frac{\sqrt{3}}{2}\mathbf{1}/\Vert\mathbf{1}\Vert$*, where* $\mathbf{1}$ *is the vector of all ones. We stopped Algorithm 1, when*

$$\Vert\Delta\mathbf{v}^{(k)}\Vert \le 5.6 \times 10^{-14}.$$

The results of Table 2 agree with those of Table 1 but with little disparities in the last two columns. This indeed show that the solution obtained by solving the under-determined system (11) *is equivalent to those obtained by solving the square system* (35)*, the disparities in the eighth and nineth rows are caused by round off errors. It also shows that the algorithm presented in [8] which involves solving an under- determined system of linear equatioins and Algorithm 1 are equivalent which is our aim.*

k	$\alpha^{(k)}$	$\beta^{(k)}$	$\|\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\|$	$\|\lambda^{(k+1)} - \lambda^{(k)}\|$	$\|\Delta\mathbf{v}^{(k)}\|$	$\|\mathbf{F}(\mathbf{v}^{(k)})\|$
0	0.00000e+00	2.50000	3.8e+00	7.8e-01	3.9e+00	3.6e+01
1	2.34253e-01	1.75371	1.8e+00	2.2e-01	1.8e+00	7.8e+00
2	1.18745e-01	1.94460	8.1e-01	1.4e-01	8.2e-01	1.7e+00
3	4.47044e-02	2.06484	2.5e-01	7.0e-02	2.6e-01	3.4e-01
4	8.82702e-03	2.12479	3.1e-02	1.7e-02	3.5e-02	3.7e-02
5	2.48114e-04	2.13905	4.8e-04	5.2e-04	7.1e-04	7.1e-04
6	1.80714e-05	2.13950	1.2e-07	2.5e-07	2.8e-07	2.8e-07
7	1.81999e-05	2.13950	1.3e-14	8.4e-14	8.5e-14	6.3e-14
8	1.81999e-05	2.13950	1.0e-14	4.8e-14	4.9e-14	5.3e-14

Table 2: Values of $\alpha^{(k)}$ and $\beta^{(k)}$ of Example 3.1. Columns 5 and 6 show that the results converged quadratically for $k = 3, 4, 5, 6$ and 7.

Conclusion

In this work, we have shown both theoretically and computationally with a numerical example that the solution obtained by solving an under-determined linear system of equations and a square system are equivalent in the absence of round off errors. This means that instead of solving (11) or (36) via QR factorisation at a cost of approximately $\frac{32}{3}n^3$ floating point operations, we could use LU factorisation to solve (35) 'more efficiently' at a cost of approximately $\frac{16}{3}n^3$. It will be interesting to see what can be proved to link the real square system of equations with the original complex one, that is (3); at least in the case $\mathbf{B} = \mathbf{I}$.

Acknowledgements

The first author acknowledges financial support from the University of Bath, United Kingdom, in the form of a University Research Studentship. This work was finalized when the first author was a post doctoral fellow at the Computational Biology Group, Faculty of Health Sciences, University of Cape Town and duely acknowledges funds provided by the National Research Foundation of South Africa.

References

[1] G. W. Stewart. *Matrix Algorithms*, volume II: Eigensystems. SIAM, 2001.

[2] B. N. Parlett and Y. Saad. Complex Shift and Invert Strategies for Real Matrices. *Lin. Alg. Appl.*, pages 575–595, 1987.

[3] K. Meerbergen, and D. Roose. Matrix Transformations for Computing Rightmost Eigenvalues of Large Sparse Non-Symmetric Eigenvalue Problems. *IMA Journal of Numerical Analysis*, 16:297–346, 1996.

[4] F. Tisseur. Newton's Method in Floating Point Arithmetic and Iterative Refinement of Generalized Eigenvalue Problems. *SIAM J. Matrix Anal. Appl.*, 22:1038–1057, 2001.

[5] A. Ruhe. Algorithms for the Nonlinear Eigenvalue Problem. *SIAM J. Matrix Anal. Appl.*, 10(4):674–689, 1973.

[6] P. Deuflhard. *Newton Methods for Nonlinear Problems*, chapter 4, pages 174–175. Springer, 2004.

[7] S. Boyd. *Lecture 8: Least Norm Solutions of Under-determined Equations*, EE263 Autumn 2008-09.

[8] R. O. Akinola and A. Spence. Two-norm normalization for the matrix pencil: Inverse iteration with a complex shift. *International Journal of Innovation in Science and Mathematics*, 2(5):435–439, 2014.

[9] R. O. Akinola. Theoretical expression for the nullvector of the jacobian: Inverse iteration with a complex shift. *International Journal of Innovation in Science and Mathematics*, 2(4):367–371, 2014.

[10] M. A. Freitag, and A. Spence. The Calculation of the Distance to Instability by the Computation of a Jordan Block. *Submitted: Linear Algebra and its Application*, August 2009.

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