Parlett, B., Dopico, F. M., Ferreira, C. (2016). The Inverse Eigenvector Problem for Real Tridiagonal Matrices. *SIAM Journal on Matrix Analysis and Applications*, 37(2), pp. 577–597.

DOI: https://doi.org/10.1137/15m1025293

THE INVERSE EIGENVECTOR PROBLEM FOR REAL TRIDIAGONAL MATRICES *

BERESFORD PARLETT†, FROILÁN M. DOPICO‡, AND CARLA FERREIRA§

Abstract. A little known property of a pair of eigenvectors (column and row) of a real tridiagonal matrix is presented. With its help we can define necessary and sufficient conditions for the unique real tridiagonal matrix for which an approximate pair of complex eigenvectors are exact. Similarly we can designate the unique real tridiagonal matrix for which two approximate real eigenvectors, with different real eigenvalues, are also exact. We close with an illustration that these unique "backward error" matrices are sensitive to small rounding errors in certain partial sums which play a key role in determining the matrices.

Key words. backward errors, tridiagonal matrices, eigenvector

AMS subject classifications. 65F15, 65F18, 15A29

1. Introduction. The real symmetric eigenvalue problem $A\boldsymbol{x} = \boldsymbol{x}\lambda$ is well posed, in an absolute sense, because an eigenvalue λ can change by no more than the spectral norm of the change in the matrix A. This is Weyl's Theorem. In the unsymmetric eigenvalue problem some eigenvalues may be robust while others may be extremely sensitive to uncertainty in the matrix entries. Consequently, the assessment of error becomes a major concern.

The best known, and useful, measure is the residual $\boldsymbol{r}$ defined by $\boldsymbol{r} := \hat{\boldsymbol{x}}\hat{\lambda} - A\hat{\boldsymbol{x}}$ for approximate eigenpair $(\hat{\lambda}, \hat{\boldsymbol{x}})$. An alternative approach which is very appealing, especially for ill-posed and ill-conditioned problems, is the so-called "backward error". Find a matrix $\widehat{A}$ such that $\widehat{A}\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}\hat{\lambda}$ and measure, or estimate, $\|\widehat{A} - A\|$. Ideal would be the $\widehat{A}$ that minimizes $\widehat{A} - A$ in some norm.

The finely crafted procedure Hessenberg QR (HQR) is called "backward stable" because the final triangular matrix is orthogonally similar to a matrix close, in norm, to the initial Hessenberg matrix. So a close problem has been solved and there is no incentive to seek a backward error in each particular case. Nevertheless, norm results do not usually respect any special structure in the initial matrix, e.g., [4, 7, 12]. This situation motivates a steady research effort on backward errors [1, 2, 5, 6, 8, 10, 11]. The problems tend to be so hard that the results are essentially theoretical because the equations determining all the $\widehat{A}$ are too forbidding to solve for $\min \|\widehat{A} - A\|$.

Tridiagonal matrices, see Section 2, are so special that even one complex triple $(\lambda, \boldsymbol{x}, \boldsymbol{y}^*)$ can, under suitable conditions, determine a unique real tridiagonal matrix C for which it is an eigentriple: $C\boldsymbol{x} = \boldsymbol{x}\lambda$, $\boldsymbol{y}^*C = \lambda\boldsymbol{y}^*$. More generally, we determine when a few vectors, complex and real, define a unique C for which they are eigenvectors. This suggests the possibility of exhibiting the backward error matrix along with a computed eigentriple.

In the course of this investigation we made two auxiliary observations that are of independent interest. The first is Theorem 3.2 (new to us) that describes two relations between the entries of the

*Carla Ferreira research was supported by the Research Centre of Mathematics of the University of Minho with the Portuguese Funds from the "Fundação para a Ciência e a Tecnologia", through the Project PEstOE/MAT/UI0013/2014. The research of F. Dopico was partially supported by the Ministerio de Economía y Competitividad of Spain through the research grant MTM2012-32542.

†Department of Mathematics and Computer Science Division of the EECS Department, University of California, Berkeley, California 94720, U.S.A. (parlett@math.berkeley.edu).

‡Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es).

§Mathematics Centre and Mathematics and Applications Department, University of Minho, 4710-057 Braga, Portugal (caferrei@math.uminho.pt).

column and row eigenvectors of any real tridiagonal matrix. The second is the significant advantage of replacing the given (C, I) form of the eigenproblem by the (T, S) form where T is real symmetric tridiagonal and S is a signature matrix, $S = \mathrm{diag}(s_i)$, $s_i = \pm 1$. When both T and S are indefinite there will usually be complex eigenvalues. The advantage comes in computing eigenvectors but this form permits simple statements and proofs of conditions for uniqueness and the unreduced property. In fact, ST will be a balanced matrix similar to C.

Most of the paper is concerned with exhibiting necessary and sufficient conditions for uniqueness and the unreduced property in various cases and we end with a warning. Our nice simple formulae involve certain partial sums which are sensitive to the rounding errors in just those partial sums which are very small. Special care is needed.

2. Notation. A signature matrix $S \in \mathbb{R}^{n\times n}$ has the form $S = \mathrm{diag}(s_1, \ldots, s_n)$, $s_i = \pm 1$. Throughout this paper real symmetric tridiagonal matrices are denoted by T and have the form

$$T := \mathrm{diag}(\boldsymbol{a}) + \mathrm{diag}(\boldsymbol{b}, -1) + \mathrm{diag}(\boldsymbol{b}, +1) = \begin{bmatrix} a_1 & b_1 & & & & & \\ b_1 & a_2 & b_2 & & & & \\ & b_2 & a_3 & b_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & & b_{n-1} & a_n \end{bmatrix}. \tag{2.1}$$

In order to avoid many repetitions of the phrase "real symmetric tridiagonal matrix" we use instead "T, as in (2.1)".

In the real unsymmetric tridiagonal case we use Greek letters instead of Roman letters,

$$C := \mathrm{diag}(\boldsymbol{\alpha}) + \mathrm{diag}(\boldsymbol{\beta}, -1) + \mathrm{diag}(\boldsymbol{\gamma}, +1) = \begin{bmatrix} \alpha_1 & \gamma_1 & & & & & \\ \beta_1 & \alpha_2 & \gamma_2 & & & & \\ & \beta_2 & \alpha_3 & \gamma_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \beta_{n-2} & \alpha_{n-1} & \gamma_{n-1} \\ & & & & \beta_{n-1} & \alpha_n \end{bmatrix}. \tag{2.2}$$

The proper definition of the term "reducible" requires the use of permutations of rows and columns. Such permutations are of no interest here and so, for C as in (2.1), we say that C is *reduced* (not simply reducible) if any $\beta_j\gamma_j$, $j = 1, \ldots, n-1$, vanishes. Otherwise it is *unreduced*.

For a column vector $\boldsymbol{x} \in \mathbb{C}^n$, we use $\boldsymbol{x}^T$ for its transpose and $\boldsymbol{x}^*$ for its conjugate transpose. We reserve capital letters for matrices. For any matrix $A \in \mathbb{C}^{n\times n}$ with eigenvalue λ, we write the eigenvector equations as $A\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*A = \lambda\boldsymbol{v}^*$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, and call $\boldsymbol{u}$ a *column eigenvector* and $\boldsymbol{v}^*$ a *row eigenvector*. Many authors use the terminology *right eigenvector* for $\boldsymbol{u}$ and *left eigenvector* for $\boldsymbol{v}$, instead of column and row eigenvectors.

3. Basic results.

3.1. Eigenvectors of real C with zero entries. If u_j is the only zero entry of the column eigenvector $\boldsymbol{u}$, then the eigenvector equation

$$C\boldsymbol{u} = \boldsymbol{u}\lambda, \qquad \boldsymbol{0} \neq \boldsymbol{u} \in \mathbb{C}^n, \tag{3.1}$$

splits naturally into two distinct smaller eigenvector equations. In a self explanatory notation, these equations are

$$\begin{aligned} C_{1:j-1}\boldsymbol{u}_{1:j-1} &= \boldsymbol{u}_{1:j-1}\lambda, \\ \beta_{j-1}u_{j-1} + 0 + \gamma_j u_{j+1} &= 0, \\ C_{j+1:n}\boldsymbol{u}_{j+1:n} &= \boldsymbol{u}_{j+1:n}\lambda. \end{aligned} \tag{3.2}$$

In the other direction, any two tridiagonal matrices C_1 and C_2 which have a shared eigenvalue λ with column eigenvectors $\boldsymbol{z} \in \mathbb{C}^{j-1}$ and $\boldsymbol{w} \in \mathbb{C}^{n-j}$, with no zero entries, may be put together into a larger tridiagonal matrix C with column eigenvector $\begin{bmatrix}\boldsymbol{z}^T & 0 & \boldsymbol{w}^T\end{bmatrix}^T \in \mathbb{C}^n$ by using a "link" equation

$$\beta_{j-1}z_{j-1} + \alpha_j z_j + \gamma_j w_1 = z_j\lambda$$

in which $z_j = 0$ and α_j is arbitrary. The other entries in column j of C, namely, γ_{j-1} and β_j, may also be chosen arbitrarily. From the "link" equation, β_{j-1} and γ_j satisfy

$$\beta_{j-1}z_{j-1} + \gamma_j w_1 = 0. \tag{3.3}$$

We can always choose $\beta_{j-1} = \gamma_j = 0$. To have $\beta_{j-1}\gamma_j \neq 0$, condition (3.3) constrains the ratio

$$\frac{z_{j-1}}{w_1} = -\frac{\gamma_j}{\beta_{j-1}} \tag{3.4}$$

to be real. If z_{j-1}/w_1 is not real, then the vector $\begin{bmatrix}\boldsymbol{z}^T & 0 & \boldsymbol{w}^T\end{bmatrix}^T$ must be changed by multiplying either $\boldsymbol{z}$ or $\boldsymbol{w}$ by a suitable scalar so that (3.4) is satisfied.

To sum up, each zero entry in an eigenvector of a tridiagonal C leads to the study of two smaller tridiagonal eigenvector problems, and so on, until we have a collection of small eigenvector problems each of whose eigenvectors has no zeros entries. In other words, there is no loss of generality in restricting attention to eigenvectors with no entries that vanish.

3.2. Complex eigenvectors of real C - the real property. We first recall standard properties of eigenvectors of a real tridiagonal matrix C. Consider the eigenvector equation (3.1) above.

LEMMA 3.1. *If C is unreduced (no $\beta_j\gamma_j$ vanishes) then*
(i) $u_1u_n \neq 0$;
(ii) two consecutive entries of $\boldsymbol{u}$ cannot both vanish.

Proof. Equation 1 in (3.1) shows that $u_1 = 0$ implies $u_2 = 0$. Equation n shows that $u_n = 0$ implies $u_{n-1} = 0$. If $u_{j-1} = 0$ and $u_j = 0$, then Equation j shows that $u_{j+1} = 0$ and, in sequence, all other entries of $\boldsymbol{u}$ must vanish. This contradicts the property $\boldsymbol{u} \neq \boldsymbol{0}$ for any eigenvector. □

In the course of our investigations we stumbled on a useful property (appears to be new) of column eigenvector $\boldsymbol{u}$ and row eigenvector $\boldsymbol{v}^*$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, when A is real tridiagonal. This result implies that there is no loss of generality in requiring that $u_jv_j \in \mathbb{R}$ for all j (see Section 4.2).

THEOREM 3.2. *Consider an unreduced real tridiagonal matrix C, as in (2.2). The column eigenvector $\boldsymbol{u}$ and the row eigenvector $\boldsymbol{v}^*$, with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, of an eigenvalue $\lambda \in \mathbb{C}$, with no zero entries, satisfying $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \lambda\boldsymbol{v}^*$, may always be chosen so that*

$P_1(j)$: $u_jv_j \in \mathbb{R}$, $j = 1, 2, \ldots, n$, *and*
$P_2(j)$: $\beta_ju_jv_{j+1}$ *and* $\gamma_jv_ju_{j+1}$ *are conjugates for* $j = 1, 2, \ldots, n-1$.

The analysis that follows is needed when λ is nonreal but it is also true, and sometimes trivial, when λ is real.

Proof. From the hypotheses we assume without explicit mention that $0 \neq \beta_j\gamma_j \in \mathbb{R}$, $u_1v_1 \neq 0$, $\overline{\lambda} - \alpha_j = \overline{\lambda - \alpha_j}$.

Since eigenvectors are unique only up to a scalar factor, $\boldsymbol{u}$ and $\boldsymbol{v}$ may be chosen so that

$$(0 \neq)u_1v_1 \in \mathbb{R}. \tag{3.5}$$

We prove $P_1(j)$ and $P_2(j)$ by induction. The base $P_1(1)$ is (3.5).

Multiply together the first equations in $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and in $C^T\boldsymbol{v} = \boldsymbol{v}\overline{\lambda}$, the conjugate transpose of $\boldsymbol{v}^*C = \lambda\boldsymbol{v}^*$,

$$\gamma_1u_2 = (\lambda - \alpha_1)u_1, \qquad \beta_1v_2 = (\overline{\lambda} - \alpha_1)v_1, \tag{3.6}$$

to find

$$\gamma_1\beta_1u_2v_2 = |\lambda - \alpha_1|^2u_1v_1 \in \mathbb{R}$$

and the unreduced property shows that $(0 \neq)u_2v_2 \in \mathbb{R}$, establishing $P_1(2)$. Next, multiply the first equation in (3.6) by v_1 and the second by u_1 to find

$$\gamma_1v_1u_2 = (\lambda - \alpha_1)u_1v_1, \qquad \beta_1u_1v_2 = (\overline{\lambda} - \alpha_1)u_1v_1 = (\overline{\lambda - \alpha_1})u_1v_1,$$

establishing $P_2(1)$.

Now assume that $P_1(j), P_1(j-1), P_2(j-1)$ hold for some $1 < j < n$. Rewrite equation j in $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and in $C^T\boldsymbol{v} = \boldsymbol{v}\overline{\lambda}$ as

$$\gamma_ju_{j+1} = (\lambda - \alpha_j)u_j - \beta_{j-1}u_{j-1}, \qquad \beta_jv_{j+1} = (\overline{\lambda} - \alpha_j)v_j - \gamma_{j-1}v_{j-1}. \tag{3.7}$$

Multiply together the two equations above and rearrange terms to find

$$\begin{aligned}\gamma_j\beta_ju_{j+1}v_{j+1} = |\lambda - \alpha_j|^2u_jv_j + \beta_{j-1}\gamma_{j-1}u_{j-1}v_{j-1} - \\ - \big[\beta_{j-1}u_{j-1}v_j(\overline{\lambda} - \alpha_j) + \gamma_{j-1}v_{j-1}u_j(\lambda - \alpha_j)\big].\end{aligned} \tag{3.8}$$

Note that $(\overline{\lambda} - \alpha_j)$ and $(\lambda - \alpha_j)$ are conjugates and invoke $P_2(j-1)$ to see that the two terms in $[\,\cdot\,]$ are conjugates and so their sum is real. $P_1(j)$ shows that the first term on the right hand side of (3.8) is real and $P_1(j-1)$ shows that the second term is real. Hence $P_1(j+1)$ holds.

To establish $P_2(j)$ multiply the first equation in (3.7) by v_j and the second by u_j to find

$$\gamma_jv_ju_{j+1} = (\lambda - \alpha_j)u_jv_j - \beta_{j-1}u_{j-1}v_j \tag{3.9}$$

$$\beta_ju_jv_{j+1} = (\overline{\lambda} - \alpha_j)u_jv_j - \gamma_{j-1}v_{j-1}u_j. \tag{3.10}$$

$P_1(j)$ and $\alpha_j \in \mathbb{R}$ show that the first terms on the right of (3.9) and (3.10) are conjugates while $P_2(j-1)$ shows that the the second terms are conjugates. This yields $P_2(j)$.

Now invoke the base and the principle of finite induction to conclude that $P_1(j)$ holds for $j = 1, 2, \ldots, n$ and $P_2(j)$ holds for $j = 1, 2, \ldots, n-1$. □

When the eigenvectors are real then P_1 holds automatically. Later we use this theorem in several places.

3.3. (T, S) form. Any unreduced real tridiagonal matrix C, i.e., $\beta_j\gamma_j \neq 0, j = 1, \ldots, n-1$, may be "balanced" by a diagonal similarity transformation $B = ECE^{-1}$ so that

$$|B_{i,i+1}| = |B_{i+1,i}|, \quad i = 1 : n-1.$$

See [9]. There is a unique positive definite diagonal matrix E with $E_{11} = 1$ that achieves this state,

$$E = \operatorname{diag}\left(1, \left|\frac{\gamma_1}{\beta_1}\right|^{1/2}, \left|\frac{\gamma_1\gamma_2}{\beta_1\beta_2}\right|^{1/2}, \ldots, \left|\frac{\gamma_1\gamma_2\cdots\gamma_{n-1}}{\beta_1\beta_2\ldots\beta_{n-1}}\right|^{1/2}\right).$$

Any real balanced tridiagonal B may be written as

$$B = ST, \tag{3.11}$$

where S is a signature matrix given by

$$s_1 = 1, \quad s_{j+1} = s_j \operatorname{sign}\left(\frac{\gamma_j}{\beta_j}\right), \quad j = 1, \ldots, n-1,$$

and T is real symmetric as in (2.1).

Note that $S^2 = I$ and so the spectrum of B is identical to the spectrum of the pair (T, S). The eigenvector equation is

$$T\boldsymbol{x} = S\boldsymbol{x}\lambda, \qquad \mathbf{0} \neq \boldsymbol{x} \in \mathbb{C}^n. \tag{3.12}$$

There are many advantages in computing with (T, S) instead of B and C.

3.4. Complex eigenvectors of a real (T, S) pair. The next elementary result is specifically relevant to our case of a real pair (T, S) with complex eigenvector $\boldsymbol{x}$. This can only occur if both T and S are indefinite.

LEMMA 3.3. *Let S be an indefinite signature matrix and let T be as in (2.1). If $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ with $\mathbf{0} \neq \boldsymbol{x} \in \mathbb{C}$ and λ nonreal, then*

$$\boldsymbol{x}^* S\boldsymbol{x} = 0. \tag{3.13}$$

Proof. Premultiply the eigenvector equation by $\boldsymbol{x}^*$ to find $\boldsymbol{x}^*T\boldsymbol{x} = \boldsymbol{x}^*S\boldsymbol{x}\lambda$. Note that $\boldsymbol{x}^*S\boldsymbol{x} = \sum_{k=1}^n s_k|x_k|^2$ is real as is $\boldsymbol{x}^*T\boldsymbol{x} = \boldsymbol{x}^TT\overline{\boldsymbol{x}} = \overline{\boldsymbol{x}^*T\boldsymbol{x}}$. Conjugate this equation and subtract to find $0 = \boldsymbol{x}^*S\boldsymbol{x}(\lambda - \overline{\lambda})$. By assumption $\lambda \neq \overline{\lambda}$. ∎

In what follows $\boldsymbol{x}^*S\boldsymbol{x} = 0$ will always be a necessary condition on complex eigenvectors of real pairs (T, S).

The strong attraction of using the (T, S) form to study eigenvectors is that, because of symmetry, it is only necessary to compute the column eigenvector $\boldsymbol{x}$ ($T\boldsymbol{x} = S\boldsymbol{x}\lambda$, $ST\boldsymbol{x} = \boldsymbol{x}\lambda$) since $\boldsymbol{x}^TS$ is a row eigenvector of ST: transposing we obtain $\boldsymbol{x}^TT = \lambda\boldsymbol{x}^TS$ and $(\boldsymbol{x}^TS)ST = \lambda\boldsymbol{x}^TS$. Observe that Lemma 3.3 does not contradict the fact that $\boldsymbol{x}^TS\boldsymbol{x} \neq 0$ when λ is simple.

4. Complex eigenvectors.

4.1. (T, S) pair with a given complex eigenvector. We begin our study of matrices T, given in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, for given nonreal λ and $\boldsymbol{x}$, with the generic case.

The next result is the key technical lemma in the paper.

LEMMA 4.1. *Let S be an indefinite signature matrix and let $\boldsymbol{x}$ be a complex vector with no zero entries that satisfies $\boldsymbol{x}^* S\boldsymbol{x} = 0$. A matrix T, as in (2.1), and nonreal $\lambda \in \mathbb{C}$ satisfy $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ if and only if*

(a) $b_k\, \mathcal{I}m(\overline{x_k}\, x_{k+1}) = \mathcal{I}m(\lambda) \sum_{j=1}^{k} s_j |x_j|^2, \quad k = 1, \ldots, n-1,$

(b) $b_{k-1}\mathcal{R}e(x_{k-1}\overline{x_k}) + a_k|x_k|^2 + b_k\mathcal{R}e(\overline{x_k}x_{k+1}) = s_k|x_k|^2\, \mathcal{R}e(\lambda)\,, \qquad k = 1, \ldots, n,$
(for brevity, let $b_0 = b_n = 0$).

The proof below is elementary for experts. Related proofs in the rest of the paper will be more succinct.

Proof. We prove first "necessity". Since $x_j \neq 0$, the jth equation in $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ may be multiplied by $\overline{x_j}$ to obtain

$$b_{j-1}x_{j-1}\overline{x_j} + a_j|x_j|^2 + b_j\overline{x_j}x_{j+1} = s_j|x_j|^2\lambda\,, \qquad j = 1, \ldots, n. \tag{4.1}$$

Since T is real, the imaginary part of (4.1) yields

$$b_{j-1}\mathcal{I}m(x_{j-1}\overline{x_j}) + 0 + b_j\mathcal{I}m(\overline{x_j}x_{j+1}) = s_j|x_j|^2\mathcal{I}m(\lambda)\,, \qquad j = 1, \ldots, n. \tag{4.2}$$

The key observation is that $\mathcal{I}m(x_{j-1}\overline{x_j}) + \mathcal{I}m(\overline{x_{j-1}}x_j) = 0$ and thus there is extensive cancellation in summing (4.2) for $j = 1, \ldots, k$, $k < n$, to find

$$b_k\, \mathcal{I}m(\overline{x_k}\, x_{k+1}) = \mathcal{I}m(\lambda) \sum_{j=1}^{k} s_j|x_j|^2, \quad \text{for } k < n.$$

This is conclusion *(a)* while conclusion *(b)* is just the real part of (4.1) for $j = k$ and thus is a real equation defining a_k uniquely.

Summing (4.2) for $j = 1, \ldots, n$ yields $0 = (\boldsymbol{x}^* S\boldsymbol{x})\mathcal{I}m(\lambda)$.

To prove "sufficiency", let us assume that *(a)* and *(b)* hold. Subtract *(a)* with index $k-1$ from *(a)* with index k to obtain (4.2) with index k. This is the imaginary part of (4.1) with index k. In addition, *(b)* is the real part of (4.1) with index k. Thus *(a)* and *(b)* imply (4.1) which is equivalent to $T\boldsymbol{x} = S\boldsymbol{x}\lambda$. ∎

THEOREM 4.2 (**generic case**). *Let S be an indefinite signature matrix and let $\boldsymbol{x} \in \mathbb{C}^n$ have no zero entries and satisfy $\boldsymbol{x}^* S\boldsymbol{x} = 0$. For each nonreal $\lambda \in \mathbb{C}$ there exists a unique unreduced T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ if and only if*

(A) $\mathcal{I}m(\overline{x_k}\, x_{k+1}) \neq 0, \quad k = 1, \ldots, n-1,$

(B) $\sum_{j=1}^{k} s_j|x_j|^2 \neq 0, \quad k = 1, \ldots, n-1.$

Proof. [Sufficiency]. Given *(A)* and *(B)* consider any matrix T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$. By Lemma 4.1*(a)*, b_k is determined uniquely by

$$b_k = \frac{\mathcal{I}m(\lambda)}{\mathcal{I}m(\overline{x_k}\, x_{k+1})} \sum_{j=1}^{k} s_j|x_j|^2, \quad k = 1, \ldots, n-1, \tag{4.3}$$

and does not vanish. Given unique b_k, $k = 1, \ldots, n-1$, Lemma 4.1*(b)* determines a_k uniquely since $x_k \neq 0$. So the matrix T given by *(a)* and *(b)* in Lemma 4.1 is unique and satisfies the eigenvector equation for the given λ.

[Necessity]. Given an unreduced T such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, then Lemma 4.1*(a)* shows that, for $k < n$, both $\mathcal{I}m(\overline{x_k}\, x_{k+1})$ and $\sum_{j=1}^{k} s_j|x_j|^2$, if they vanish, must vanish together, since $b_k \neq 0$ and $\mathcal{I}m(\lambda) \neq 0$.

The next step is to show that $\mathcal{I}m(\overline{x_k}\, x_{k+1}) = 0$, $k < n$, violates uniqueness. Observe that $\mathcal{I}m(\overline{x_{n-1}}\, x_n) \neq 0$; see (4.2) with $j = n$. So, let k be the smallest index for which $\mathcal{I}m(\overline{x_k}\, x_{k+1}) = -\mathcal{I}m(x_k\, \overline{x_{k+1}}) = 0$; necessarily $k < n-1$. Then $b_{k-1} \neq 0$ is uniquely determined by Lemma 4.1*(a)* with index $k-1$. Next consider Equation (4.2) in Lemma 4.1 with $j = k+1 < n$,

$$b_k \mathcal{I}m(x_k \overline{x_{k+1}}) + b_{k+1}\mathcal{I}m(\overline{x_{k+1}} x_{k+2}) = s_{k+1}|x_{k+1}|^2 \mathcal{I}m(\lambda).$$

With $\mathcal{I}m(x_k \overline{x_{k+1}}) = 0$,

$$b_{k+1}\mathcal{I}m(\overline{x_{k+1}} x_{k+2}) = s_{k+1}|x_{k+1}|^2 \mathcal{I}m(\lambda)$$

and the right hand side does not vanish and, since $b_{k+1} \neq 0$, $\mathcal{I}m(\overline{x_{k+1}} x_{k+2})$ can not vanish either and so b_{k+1} is uniquely determined by this equation. In contrast, *(a)* in Lemma 4.1 is vacuous,

$$b_k \cdot 0 = 0 \cdot \mathcal{I}m(\lambda),$$

and puts no constraint on b_k. So, let $\widetilde{b}_k$ represent any value other than the b_k given by T. Now, consider replacing b_k by $\widetilde{b}_k$ in Lemma 4.1. Conclusion *(b)* yields a unique new $\widetilde{a}_k$ in terms of $\widetilde{b}_k$, i.e.,

$$b_{k-1}\mathcal{R}e(x_{k-1}\overline{x_k}) + \widetilde{a}_k|x_k|^2 + \widetilde{b}_k \mathcal{R}e(\overline{x_k} x_{k+1}) = s_k|x_k|^2\, \mathcal{R}e(\lambda)$$

since $|x_k|^2 \neq 0$. Next, increasing the index by 1,

$$\widetilde{b}_k \mathcal{R}e(x_k \overline{x_{k+1}}) + \widetilde{a}_{k+1}|x_{k+1}|^2 + b_{k+1}\mathcal{R}e(\overline{x_{k+1}} x_{k+2}) = s_{k+1}|x_{k+1}|^2\, \mathcal{R}e(\lambda)$$

determines a unique new $\widetilde{a}_{k+1}$ in terms of $\widetilde{b}_k$, since $|x_{k+1}|^2 \neq 0$. No other equations involve $\widetilde{b}_k$.

Thus there is another T, as in (2.1), differing from the given T in $\widetilde{b}_k$, $\widetilde{a}_k$, $\widetilde{a}_{k+1}$, only, that has $(\lambda, \boldsymbol{x})$ as an eigenpair. This violates uniqueness.

If $\displaystyle\sum_{j=1}^{k} s_j|x_j|^2 = 0$, then $\mathcal{I}m(\overline{x_k}\, x_{k+1}) \neq 0$ implies that $b_k = 0$ and T is reduced. Thus *(A)* and *(B)* must hold for all $k < n$ when T is unique and unreduced. □

THEOREM 4.3 (**nongeneric case**). *Let S be an indefinite signature matrix and let $\boldsymbol{x} \in \mathbb{C}^n$ have no zero entries and satisfy $\boldsymbol{x}^* S\boldsymbol{x} = 0$. For each nonreal $\lambda \in \mathbb{C}$, there exists*

(I) a unique T, as in (2.1), satisfying $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ if and only if

$$\mathcal{I}m(\overline{x_k}\, x_{k+1}) \neq 0, \quad k = 1, \ldots, n-1. \tag{4.4}$$

(II) an unreduced T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ if and only if for each $k = 1, \ldots, n-1$,

$$\text{either both or neither of } \mathcal{I}m(\overline{x_k}\, x_{k+1}) \text{ and } \sum_{j=1}^{k} s_j|x_j|^2 \text{ vanish.} \tag{4.5}$$

Proof. [Sufficiency for *(I)*]. Given (4.4), according to *(a)* in Lemma 4.1, b_k is uniquely defined by (4.3) and vanishes if and only if $\sum_{j=1}^{k} s_j|x_j|^2 = 0$.

Since $x_k \neq 0$ and the $\{b_k\}$ are unique, condition *(b)* in Lemma 4.1 defines a_k uniquely.

[Necessity for *(I)*]. Note, from Equation (4.2) in Lemma 4.1, that

$$b_{n-1}\mathcal{I}m(x_{n-1}\overline{x_n}) = s_n|x_n|^2\mathcal{I}m(\lambda) \neq 0.$$

Let $k < n-1$ be the smallest index such that $\mathcal{I}m(\overline{x_k}x_{k+1}) = 0$. Then, by *(a)* in Lemma 4.1,

$$\sum_{j=1}^{k} s_j|x_j|^2 = b_k\,\mathcal{I}m(\overline{x_k}\,x_{k+1})/\mathcal{I}m(\lambda) = 0.$$

Thus, there are no constraints on b_k and, exactly as in the proof of Theorem 4.2, there is a family of T matrices with $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ contradicting uniqueness. Hence, uniqueness implies (4.4).

[Sufficiency for *(II)*]. When $\mathcal{I}m(\overline{x_k}\,x_{k+1}) \neq 0$ and $\sum_{j=1}^{k} s_j|x_j|^2 \neq 0$, then *(a)* in Lemma 4.1 determines a unique nonzero b_k. On the other hand, when both vanish, then, as in the proof of Theorem 4.2, there is no constraint on b_k and it may be taken as a free parameter in forming a family of T with $T\boldsymbol{x} = S\boldsymbol{x}\lambda$. Any nonzero value for b_k gives an unreduced T and it will not be unique whenever $\mathcal{I}m(\overline{x_k}\,x_{k+1}) = 0$.

[Necessity for *(II)*]. Consider any unreduced T, as in (2.1), that satisfies $T\boldsymbol{x} = S\boldsymbol{x}\lambda$. Since $b_k \neq 0$, *(a)* in Lemma 4.1 shows that $\mathcal{I}m(\overline{x_k}\,x_{k+1})$ and $\sum_{j=1}^{k} s_j|x_j|^2$ are both zero or neither is zero. This is (4.5). □

REMARK 4.1. *It is the condition that no x_k can vanish that permits these simple proofs. The reader is referred to Section 3.1 to see that there is no loss of generality in this assumption.*

REMARK 4.2. *Condition (4.4) shows that, in the general case, an eigenvector $\boldsymbol{x}$ of a pair (T, S) with no zero entries must have the property that the ratio $x_{k+1}/x_k (= \overline{x_k}\,x_{k+1}/|x_k|^2), k = 1, \ldots, n-1$, is not real whenever its eigenvalue is not real.*

Finally, a reader may ask for mere existence of T. The proof of *(II)* in Theorem 4.3 shows

THEOREM 4.4. *Let S be an indefinite signature matrix and let $\boldsymbol{x} \in \mathbb{C}^n$ have no zero entries and satisfy $\boldsymbol{x}^*S\boldsymbol{x} = 0$. For each nonreal $\lambda \in \mathbb{C}$,*

(III) there exists a T, as in (2.1), with $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ if and only if, for $k = 1, \ldots, n-1$,

$$\mathcal{I}m(\overline{x_k}\,x_{k+1}) = 0 \text{ implies } \sum_{j=1}^{k} s_j|x_j|^2 = 0.$$

Proof. Condition *(a)* in Lemma 4.1 determines a b_k but with no constraints when both sides vanish. Condition *(b)* determines a_k in terms of b_{k-1} and b_k since $x_k \neq 0$. □

4.2. C with a given pair of complex eigenvectors. The next step is to extend the results of Section 4.1 from the (T, S) form to the (C, I) form, as given in (2.2), with

$$C = \mathrm{diag}(\boldsymbol{\alpha}) + \mathrm{diag}(\boldsymbol{\beta}, -1) + \mathrm{diag}(\boldsymbol{\gamma}, +1) \tag{4.6}$$

for a real n-array $\boldsymbol{\alpha}$ and two real $n-1$ arrays $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. Our interest is mainly in unreduced matrices where $\beta_j\gamma_j \neq 0, j = 1, \ldots, n-1$.

In this section we seek necessary and sufficient conditions on a pair of vectors, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$, so that, for any nonreal λ, there exists a real C, as in (4.6), with $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \lambda\boldsymbol{v}^*$.

By Theorem 3.2 in Section 3.2, $\boldsymbol{u}$ and $\boldsymbol{v}$ may always be chosen to satisfy certain conditions and henceforth we only consider complex vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ which satisfy these simplifying conditions.

Necessary and sufficient conditions may be found by following similar techniques to those used in Section 4.1 but we claim that, despite first appearances, the problems in the two sections are equivalent and so the results in Section 4.1 may be translated into the correct results for this section.

The first appearances mentioned above are that there are $3n-2$ real unkonwns $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}$ and two given complex vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ amount to approximately $4n$ real conditions. The apparent mismatch of conditions and unknowns is illusory because $\boldsymbol{u}$ and $\boldsymbol{v}$ are far from independent. See Theorem 3.2.

To see the equivalence of the two problems (Section 4.1 and Section 4.2), let us start with any real unreduced C, as in (4.6). From Section 3.3 we know that C may be balanced by a unique diagonal similarity transformation

$$ST = ECE^{-1}$$

with $E = \operatorname{diag}(e_1, \ldots, e_n)$, positive definite and $e_1 = 1$, determined by C. These properties of E are used below. Recall that S is a signature matrix, $S = \operatorname{diag}(s_1, s_2, \ldots, s_n)$ with $s_i = \pm 1$.

Let $\boldsymbol{x}$, with no zero entries, be the complex eigenvector of ST for λ,

$$ST\boldsymbol{x} = \boldsymbol{x}\lambda, \qquad (\boldsymbol{x}^T S)ST = \lambda \boldsymbol{x}^T S.$$

Now substitute ECE^{-1} for ST to find

$$CE^{-1}\boldsymbol{x} = E^{-1}\boldsymbol{x}\lambda \qquad \text{and} \qquad (\boldsymbol{x}^T SE)C = \boldsymbol{x}^T SE\lambda. \tag{4.7}$$

Thus, given $\boldsymbol{x}$, with E and S known, we can take $\boldsymbol{u} = E^{-1}\boldsymbol{x}$ and $\boldsymbol{v}^* = \boldsymbol{x}^T SE$, ignoring scalar factors, with $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^* C = \lambda \boldsymbol{v}^*$, since the real property $P_1(j)$ in Theorem 3.2 is satisfied,

$$u_j v_j = (x_j/e_j)(\overline{x_j} s_j e_j) = s_j |x_j|^2 \in \mathbb{R}.$$

Now, given two complex eigenvectors $\boldsymbol{u}$ and $\boldsymbol{v}^*$ of C for a nonreal λ, with no zero entries and satisfying the real condition $P_1(j)$ in Theorem 3.2, to have consistency in defining $\boldsymbol{x}$ from (4.7), either as a multiple of $E\boldsymbol{u}$ or as a multiple of $SE^{-1}\overline{\boldsymbol{v}}$, with the signature matrix S defined from

$$s_j := \operatorname{sign}(u_j v_j), \tag{4.8}$$

we need a scaling factor μ so that

$$x_j = e_j u_j, \qquad x_j = (s_j \overline{v_j}\mu)/e_j.$$

Hence, since $u_j v_j$ is real and $e_1 = 1$,

$$\begin{aligned} \mu &= e_j^2 s_j (u_j/\overline{v_j}) = e_j^2 (s_j u_j v_j)/|v_j|^2 = e_j^2 |u_j v_j|/|v_j|^2 = e_j^2 |u_j|/|v_j|, \quad j = 1, 2, \ldots, n \\ &= |u_1|/|v_1|. \end{aligned} \tag{4.9}$$

Also, from the equalities above,

$$e_j^2 = |v_j/v_1|/|u_j/u_1| > 0, \quad j = 1, 2, \ldots, n. \tag{4.10}$$

Observe that (4.9) fixes μ and then (4.10) fixes E, positive definite, as

$$e_j = \big(|v_j/v_1|/|u_j/u_1|\big)^{1/2}.$$

Thus, under suitable conditions, for any nonreal λ, S and $\boldsymbol{x}$ define real T and $C = E^{-1}STE$. Indeed, the two problems (T, S) and (C, I) are equivallent.

Theorem 4.2 gives the conditions on $\boldsymbol{x}$ to have a unique real unreduced symmetric T satisfying $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ for any nonreal λ. These conditons must be translated into the conditions on $\boldsymbol{u}$ and $\boldsymbol{v}$. In practice we normalize $\boldsymbol{u}$ and $\boldsymbol{v}$ so that $|u_1| = |v_1|$ and then $\mu = 1$. Formulae become simpler but, for the theorem to follow, we do not need $\mu = 1$.

Observe that

$$s_j|x_j|^2 = s_j(e_ju_j)(s_jv_j\overline{\mu}/e_j) = \mu u_jv_j, \qquad \mu \neq 0, \tag{4.11}$$

and Condition *(B)* in Theorem 4.2 gives us

$$\sum_{j=1}^{k} u_jv_j \neq 0, \quad k = 1, \ldots, n-1.$$

Set $k = n$ to recover the necessary condition $0 = \boldsymbol{v}^T\boldsymbol{u} = \mu(\boldsymbol{x}^*SE)(E^{-1}\boldsymbol{x}) = \mu\boldsymbol{x}^*S\boldsymbol{x}$ when λ is nonreal. This does not contradict the fact that $\boldsymbol{v}^*\boldsymbol{u} \neq 0$ when λ is simple.

For Condition *(A)* in Theorem 4.2 we must connect the two sets of off-diagonal entries $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ in (4.6) to one set of off-diagonal entries $\boldsymbol{b}$ of T. Again, $C = E^{-1}STE$ yields

$$\gamma_j = e_j^{-1}s_jb_je_{j+1}, \quad \beta_j = e_{j+1}^{-1}s_{j+1}b_je_j \tag{4.12}$$

so that both γ_j and β_j are simply related to b_j. Also

$$\alpha_j = s_ja_j$$

for the diagonal entries.

The important quantity $b_k\overline{x_k}x_{k+1}$ from Lemma 4.1 satisfies

$$b_k\overline{x_k}x_{k+1} = (\gamma_ks_ke_k/e_{k+1})[v_k/(s_ke_k)\mu](u_{k+1}e_{k+1}) = \mu\gamma_kv_ku_{k+1}. \tag{4.13}$$

Note that its conjugate $b_k\overline{x_{k+1}}x_k$ translates into

$$b_k\overline{x_{k+1}}x_k = (\beta_ks_{k+1}e_{k+1}/e_k)[v_{k+1}/(s_{k+1}e_{k+1})\mu](u_ke_k) = \mu\beta_ku_kv_{k+1}. \tag{4.14}$$

Again, the $\{e_j\}$ need not be known since they cancel out.

We have recovered property $P_2(j)$ in Theorem 3.2 in Section 3.2.

There are nice formulae for the entries of C. Substitute (4.11), (4.13) and (4.14) into Lemma 4.1 *(a)*, for $k = 1, \ldots, n-1$, to obtain

$$\gamma_k\,\mathcal{I}m(v_k\,u_{k+1}) = \mathcal{I}m(\lambda)\sum_{j=1}^{k} u_jv_j, \tag{4.15}$$

$$\beta_k\,\mathcal{I}m(u_k\,v_{k+1}) = -\mathcal{I}m(\lambda)\sum_{j=1}^{k} u_jv_j. \tag{4.16}$$

Equate real parts to find $\mu\beta_k\mathcal{R}e(u_kv_{k+1}) = \mu\gamma_k\mathcal{R}e(v_ku_{k+1}) = b_k\mathcal{R}e(\overline{x_k}x_{k+1}) = b_k\mathcal{R}e(\overline{x_{k+1}}x_k)$. Lemma 4.1 *(b)*, for $k = 1, \ldots, n$, gives

$$\beta_{k-1}\mathcal{R}e(u_{k-1}v_k) + \alpha_ku_kv_k + \gamma_k\mathcal{R}e(v_ku_{k+1}) = u_kv_k\,\mathcal{R}e(\lambda). \tag{4.17}$$

Other versions of (4.17) are

$$\gamma_{k-1}\mathcal{R}e(u_k v_{k-1}) + \alpha_k u_k v_k + \gamma_k \mathcal{R}e(v_k u_{k+1}) = u_k v_k \, \mathcal{R}e(\lambda)$$

and

$$\beta_{k-1}\mathcal{R}e(u_{k-1} v_k) + \alpha_k u_k v_k + \beta_k \mathcal{R}e(u_k v_{k+1}) = u_k v_k \, \mathcal{R}e(\lambda).$$

Since $\gamma_k v_k u_{k+1}$ and $\beta_k u_k v_{k+1}$ vanish together, if either vanishes, either $\mathcal{I}m(v_k\, u_{k+1}) \neq 0$ or $\mathcal{I}m(u_k v_{k+1}) \neq 0$ may be substituted for the condition $\mathcal{I}m(\overline{x_k} x_{k+1}) \neq 0$ from Theorem 4.2. We are now in a position to restate Theorem 4.2 as it applies to real C.

THEOREM 4.5 (**generic case**). *Let $\boldsymbol{u} \in \mathbb{C}^n, \boldsymbol{v} \in \mathbb{C}^n$ be complex vectors with no zero entries satisfying $\boldsymbol{v}^T\boldsymbol{u} = 0$ and $u_j v_j \in \mathbb{R}, j = 1, \ldots, n$. For each nonreal $\lambda \in \mathbb{C}$ there exists a unique unreduced $C \in \mathbb{R}^{n\times n}$, as in (2.2), such that $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \lambda\boldsymbol{v}^*$ if and only if*

(A) $\mathcal{I}m(v_k\, u_{k+1}) \neq 0, \quad k = 1, \ldots, n-1,$

(B) $\displaystyle\sum_{j=1}^{k} u_j v_j \neq 0, \quad k = 1, \ldots, n-1.$

For the nongeneric case Theorems 4.3 and 4.4 applied to real C gives

THEOREM 4.6 (**nongeneric case**). *Let $\boldsymbol{u} \in \mathbb{C}^n, \boldsymbol{v} \in \mathbb{C}^n$ be complex vectors with no zero entries satisfying $\boldsymbol{v}^T\boldsymbol{u} = 0$ and $u_j v_j \in \mathbb{R}, j = 1, \ldots, n$. For each nonreal $\lambda \in \mathbb{C}$, there exists*

(I) a unique $C \in \mathbb{R}^{n\times n}$, as in (2.2), satisfying $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \boldsymbol{v}^*\lambda$ if and only if

$$\mathcal{I}m(v_k\, u_{k+1}) \neq 0, \quad k = 1, \ldots, n-1. \tag{4.18}$$

However C might be reduced.

(II) an unreduced $C \in \mathbb{R}^{n\times n}$, as in (2.2), satisfying $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \boldsymbol{v}^*\lambda$ if and only if for each $k = 1, \ldots, n-1$,

$$\text{either both or neither of } \mathcal{I}m(v_k\, u_{k+1}) \text{ and } \sum_{j=1}^{k} u_j v_j \text{ vanish.} \tag{4.19}$$

(III) (mere existence) $C \in \mathbb{R}^{n\times n}$, as in (2.2), satisfying $C\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^*C = \boldsymbol{v}^*\lambda$ if and only if for $k = 1, \ldots, n-1$,

$$\mathcal{I}m(v_k\, u_{k+1}) = 0 \text{ implies } \sum_{j=1}^{k} u_j v_j = 0.$$

When both terms vanish then both γ_k and β_k are unconstrained by (4.15) and (4.16).

4.3. *J* matrices. In many applications the matrix C in (2.2) has the special property that $\gamma_j = 1, j = 1, 2, \ldots, n-1$. This is often called a J matrix and has the virtue of the form being invariant under the LR and dqds transforms,

$$J = \begin{bmatrix} \alpha_1 & 1 & & & & \\ \beta_1 & \alpha_2 & 1 & & & \\ & \beta_2 & \alpha_3 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_{n-2} & \alpha_{n-1} & 1 \\ & & & & \beta_{n-1} & \alpha_n \end{bmatrix}. \tag{4.20}$$

In this case relation (4.15), with $\gamma_k = 1$, imposes a tight connection between $\mathcal{I}m(v_k u_{k+1})$ and $\sum_{j=1}^{k} u_j v_j$ for each $k = 1, 2, \ldots, n-1$. Their ratios must all be equal and thus fix $\mathcal{I}m(\lambda)$.

From (4.13) and (4.14) we see that

$$\beta_k u_k v_{k+1} = \overline{v_k u_{k+1}}$$

and, thus,

$$\beta_k = \frac{(\overline{v_k} v_k)(\overline{u_{k+1}} u_{k+1})}{(u_k v_k)(u_{k+1} v_{k+1})} = \frac{|v_k|^2 |u_{k+1}|^2}{(u_k v_k)(u_{k+1} v_{k+1})} \neq 0, \quad k = 1, 2, \ldots, n-1. \tag{4.21}$$

So, β_k is well defined, by the real property, and is unique. Moreover, on adding (4.15) and (4.16), we have

$$\beta_k = -\mathcal{I}m(v_k u_{k+1})/\mathcal{I}m(u_k v_{k+1}), \quad k = 1, 2, \ldots, n-1.$$

Then the relation (4.17) fixes α_k for any given $\mathcal{R}e(\lambda)$ and $\mathcal{I}m(\lambda) = \mathcal{I}m(v_1 u_2)/u_1 v_1 \neq 0$.

Turning to the nongeneric case we note that relation (4.15), with $\gamma_k = 1$, implies Condition III in Theorem 4.6 and so uniqueness and the unreduced property must accompany (mere) existence of a real J matrix with the given complex eigenvector. We state this result formally.

THEOREM 4.7. *Let $\boldsymbol{u} \in \mathbb{C}^n$, $\boldsymbol{v} \in \mathbb{C}^n$ be complex vectors with no zero entries satisfying $\boldsymbol{v}^T \boldsymbol{u} = 0$ and $u_i v_i \in \mathbb{R}$, $i = 1, \ldots, n$. For a suitable real value $\mathcal{I}m(\lambda) \neq 0$ there exists a real J matrix, as in (4.20), such that $J\boldsymbol{u} = \boldsymbol{u}\lambda$, $\boldsymbol{v}^* J = \lambda \boldsymbol{v}^*$ if and only if*

(A) $\mathcal{I}m(v_k\, u_{k+1}) = \mathcal{I}m(\lambda) \sum_{j=1}^{k} u_j v_j, \quad k = 1, \ldots, n-1.$

In this case, for each value of $\mathcal{R}e(\lambda)$, (4.21) and (4.17), with $\gamma_k = 1$, determine a unique unreduced J matrix with $J\boldsymbol{u} = \boldsymbol{u}\lambda$ and $\boldsymbol{v}^* J = \lambda \boldsymbol{v}^*$.

5. Real eigenvectors.

5.1. (T, S) with a given pair of real eigenvectors. There are $2n-1$ free real degrees of freedom in T, as in (2.1), while 2 real n-vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ impose $2n$ real conditions. However, in the problem considered here $\boldsymbol{y}^T S \boldsymbol{x} = 0$ and thus the number of constraints matches the degrees of freedom. This condition is true in general, not just in the tridiagonal case.

The analysis follows the pattern of the complex case and begins with a key technical lemma.

REMARK 5.1. *To see that determinants have been presented in Section 4 but hidden, observe that*

$$\mathcal{I}m(\overline{x_k}\, x_{k+1}) = \det \begin{bmatrix} \mathcal{R}e(x_k) & \mathcal{I}m(x_k) \\ \mathcal{R}e(x_{k+1}) & \mathcal{I}m(x_{k+1}) \end{bmatrix}.$$

LEMMA 5.1. *Consider a signature matrix S, which could be definite or indefinite, and two linearly independent vectors $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^n$ satisfying $\boldsymbol{y}^T S \boldsymbol{x} = 0$ and $(x_j, y_j) \neq (0, 0)$, $j = 1, \ldots, n$. If T, as in (2.1), satisfies $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, $T\boldsymbol{y} = S\boldsymbol{y}\mu$, $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\lambda \neq \mu$, then*

(a) $b_k \det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} = (\mu - \lambda) \sum_{j=1}^{k} s_j x_j y_j, \quad k = 1, \ldots, n-1.$

(b) $b_{k-1}(x_{k-1} y_k + x_k y_{k-1}) + 2a_k x_k y_k + b_k(x_k y_{k+1} + x_{k+1} y_k) = s_k x_k y_k (\mu + \lambda), \quad k = 1, \ldots, n,$
(for brevity, let $b_0 = b_n = 0$.)

Proof. When $x_j y_j \neq 0$ it is legitimate to multiply the jth equation in $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ by y_j and the jth equation in $T\boldsymbol{y} = S\boldsymbol{y}\mu$ by x_j to obtain

$$b_{j-1}x_{j-1}y_j + a_j x_j y_j + b_j x_{j+1} y_j = s_j x_j y_j \lambda, \tag{5.1}$$
$$b_{j-1}x_j y_{j-1} + a_j x_j y_j + b_j x_j y_{j+1} = s_j x_j y_j \mu. \tag{5.2}$$

Subtract the second from the first to find

$$b_{j-1}(x_{j-1}y_j - x_j y_{j-1}) + 0 + b_j(x_{j+1}y_j - x_j y_{j+1}) = s_j x_j y_j(\lambda - \mu), \tag{5.3}$$

that is,

$$b_{j-1} \det \begin{bmatrix} x_{j-1} & y_{j-1} \\ x_j & y_j \end{bmatrix} - b_j \det \begin{bmatrix} x_j & y_j \\ x_{j+1} & y_{j+1} \end{bmatrix} = s_j x_j y_j(\lambda - \mu), \quad j = 1, \ldots, n-1. \tag{5.4}$$

Observe that the second term on the left hand side of (5.4) for index $j-1$ is the negative of the first term for index j. Now, sum equation (5.4) for $j = 1, \ldots, k$, $k < n$, to obtain

$$-b_k \det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} = (\lambda - \mu) \sum_{j=1}^{k} s_j x_j y_j.$$

This is conclusion *(a)* when $x_j y_j \neq 0$.

Summing (5.3) for $j = 1, \ldots, n-1$, yields $0 = (\lambda - \mu) \sum_{j=1}^{n} s_j x_j y_j$ which shows that $\boldsymbol{y}^T S \boldsymbol{x} = 0$ is a necessary condition.

Summing equations (5.1) and (5.2) yields

$$b_{j-1}(x_{j-1}y_j + x_j y_{j-1}) + 2a_j x_j y_j + b_j(x_j y_{j+1} + x_{j+1} y_j) = s_j x_j y_j(\lambda + \mu), \quad j = 1, \ldots, n, \tag{5.5}$$

which is conclusion *(b)* when $x_j y_j \neq 0$.

Next we seek to relax the condition $x_j y_j \neq 0$ to the weaker one $(x_j, y_j) \neq (0, 0)$. Suppose that $x_j = 0, y_j \neq 0$. Then (5.1) becomes

$$b_{j-1}x_{j-1}y_j + 0 + b_j x_{j+1} y_j = 0 \tag{5.6}$$

and the jth equation in $T\boldsymbol{y} = S\boldsymbol{y}\mu$ is ignored because it is now multiplied by $x_j = 0$. However, (5.6) is precisely both (5.4) and (5.5) when $x_j = 0$.

The situation is the same when $x_j \neq 0$, $y_j = 0$. We recover (5.4) and (5.5). The rest of the argument is unchanged and the proof is complete for conclusions *(a)* and *(b)*. □

THEOREM 5.2 (**generic case**). *Consider a signature matrix S, which could be definite or indefinite, and two linearly independent vectors $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^n$ with $(x_j, y_j) \neq (0, 0)$, $j = 1, \ldots, n$, and $\boldsymbol{y}^T S \boldsymbol{x} = 0$. For any $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\lambda \neq \mu$, there exists a unique unreduced T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda, T\boldsymbol{y} = S\boldsymbol{y}\mu$ if and only if*

(A) $\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} \neq 0, \quad k = 1, \ldots, n-1,$

(B) $\sum_{j=1}^{k} s_j x_j y_j \neq 0, \quad k = 1, \ldots, n-1.$

Proof. [Sufficiency]. When *(A)* and *(B)* hold, then *(a)* in Lemma 5.1 determines b_k as unique and nonzero. If $x_k y_k \neq 0$, then *(b)* in Lemma 5.1 defines a_k uniquely. For the cases $x_k = 0, y_k \neq 0$ and $x_k \neq 0, y_k = 0$, the uniqueness of a_k is confirmed by the k-th equation in $T\boldsymbol{y} = S\boldsymbol{y}\lambda$ and by the k-th equation in $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, respectively.

[Necessity].The argument follows the one in Theorem 4.2 with the determinant in *(A)* taking place of $\mathcal{I}m(\overline{x_k}\, x_{k+1})$. When this determinant vanishes then b_k may be taken as a free parameter and both a_k and a_{k+1} are determined uniquely given b_k in all three cases: $x_k y_k \neq 0$; $x_k = 0, y_k \neq 0$; $x_k \neq 0, y_k = 0$. Uniqueness is lost. Thus, when T is unique, the determinant cannot vanish and then, using the unreduced property ($b_k \neq 0$), *(a)* in Lemma 5.1 shows that the determinant in *(A)* and the corresponding partial sum $\sum_{j=1}^{k} s_j x_j y_j$ must vanish together if either of them does vanish. Thus, *(A)* and *(B)* hold. □

The nongeneric case also follows the pattern of the complex case.

THEOREM 5.3 (**nongeneric case**). *Let S be a signature matrix, definite or indefinite, and let two linearly independent vectors $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^n$ satisfy $(x_j, y_j) \neq (0,0)$, $j = 1, \ldots, n$, and $\boldsymbol{y}^T S\boldsymbol{x} = 0$. For any $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\lambda \neq \mu$, there exists*

(I) a unique T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, $T\boldsymbol{y} = S\boldsymbol{y}\mu$ if and only if

$$\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} \neq 0, \quad k = 1, \ldots, n-1. \tag{5.7}$$

(II) an unreduced T, as in (2.1), such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, $T\boldsymbol{y} = S\boldsymbol{y}\mu$ if and only if for each $k = 1, \ldots, n-1$,

$$\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} \text{ and } \sum_{j=1}^{k} s_j x_j y_j \text{ vanish together, if either vanishes.} \tag{5.8}$$

Proof. [Sufficiency for *(I)*]. When (5.7) holds, then *(a)* in Lemma 5.1 gives b_k, $k = 1, \ldots, n-1$, uniquely and b_k vanishes if and only if $\displaystyle\sum_{j=1}^{k} s_j x_j y_j = 0$. When one of x_k, y_k vanishes, then a_k is determined uniquely from either $T\boldsymbol{x} = S\boldsymbol{x}\lambda$ or $T\boldsymbol{y} = S\boldsymbol{y}\mu$ accordingly. When neither vanishes, then a_k is determined by Lemma 5.1 *(b)*. Consistency follows from the exclusive use of elementary operations in the proof of Lemma 5.1.

[Necessity for *(I)*]. The argument in the proof of "necessity" in Theorem 5.2 does not require the unreduced property and so may be invoked here to show that if $\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} = 0$ for any $1 \leq k \leq n-1$, then uniqueness of T is lost. Lemma 5.1 *(a)* shows that the partial sum $\sum_{j=1}^{k} s_j x_j y_j$ vanishes in this case.

[Sufficiency for *(II)*]. Assume that (5.8) holds. If $\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} \neq 0$, then b_k is determined and nonzero by Lemma 5.1 *(a)*. If the determinant vanishes, then b_k is unconstrained and we can choose $b_k \neq 0$ which then determines a_k and a_{k+1} as in the proof of Theorem 5.2. In such cases T is not unique.

[Necessity for *(II)*]. Given an unreduced T satisfying $T\boldsymbol{x} = S\boldsymbol{x}\lambda, T\boldsymbol{y} = S\boldsymbol{y}\mu$, then Lemma 5.1 *(a)* shows that $\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix}$ and $\displaystyle\sum_{j=1}^{k} s_j x_j y_j$ vanish together if either does vanishes. Thus, (5.8) holds. □

Finally, we state without proof that given $S, \boldsymbol{x}, \lambda, \boldsymbol{y}, \mu$ as in Theorem 5.3, then T, as in (2.1), exists such that $T\boldsymbol{x} = S\boldsymbol{x}\lambda$, $T\boldsymbol{y} = S\boldsymbol{y}\mu$ if and only if for $k = 1, \ldots, n-1$,

$$\det \begin{bmatrix} x_k & y_k \\ x_{k+1} & y_{k+1} \end{bmatrix} = 0 \quad \text{implies} \quad \sum_{j=1}^{k} s_j x_j y_j = 0.$$

5.2. C with two given pairs of real eigenvectors. We first study the necessary conditions that constrain two pairs of real eigenvectors associated with different real eigenvalues of a real unreduced tridiagonal matrix C as in (4.6),

$$C\boldsymbol{u} = \boldsymbol{u}\lambda, \qquad \boldsymbol{v}^T C = \lambda \boldsymbol{v}^T, \qquad C\boldsymbol{w} = \boldsymbol{w}\mu, \qquad \boldsymbol{z}^T C = \mu \boldsymbol{z}^T, \qquad \lambda \neq \mu.$$

Note that, as for any real square matrix,

$$\boldsymbol{v}^T \boldsymbol{w} = \boldsymbol{z}^T \boldsymbol{u} = 0. \tag{5.9}$$

For simplicity we assume no zero entries in $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ and $\boldsymbol{z}$. This assumption is used extensively and will not be mentioned each time.

We also assume the necessary conditions found in property $P_2(j)$ in Theorem 3.2, with no mention of β_j and γ_j, which take the simple form

$$\frac{u_{j+1}}{u_j} \Big/ \frac{v_{j+1}}{v_j} = \frac{w_{j+1}}{w_j} \Big/ \frac{z_{j+1}}{z_j}, \quad j = 1, \ldots, n-1. \tag{5.10}$$

The eigenvector equations yield $4n$ equations. For $j = 1, \ldots, n$,

$$\begin{aligned} \beta_{j-1}u_{j-1} + \alpha_j u_j + \gamma_j u_{j+1} &= u_j\lambda, & \beta_{j-1}w_{j-1} + \alpha_j w_j + \gamma_j w_{j+1} &= w_j\mu, \\ \gamma_{j-1}v_{j-1} + \alpha_j v_j + \beta_j v_{j+1} &= v_j\lambda, & \gamma_{j-1}z_{j-1} + \alpha_j z_j + \beta_j z_{j+1} &= z_j\mu. \end{aligned}$$

All entries with index 0 and $n+1$ vanish. We now perform elementary operations on this system of equations to obtain an equivalent but more informative system. With condition (5.9) in mind we combine equations for $\boldsymbol{u}$, $\boldsymbol{z}$ and for $\boldsymbol{w}, \boldsymbol{v}$. Multiply the j-th equation for $\boldsymbol{u}$ by z_j, the j-th equation for $\boldsymbol{z}$ by u_j and subtract. Do the same for $\boldsymbol{v}, \boldsymbol{w}$ to find, for $j = 1, \ldots, n$,

$$(\beta_{j-1}u_{j-1}z_j - \gamma_{j-1}z_{j-1}u_j) + (\gamma_j z_j u_{j+1} - \beta_j u_j z_{j+1}) = u_j z_j(\lambda - \mu), \tag{5.11}$$

$$(\gamma_{j-1}v_{j-1}w_j - \beta_{j-1}w_{j-1}v_j) + (\beta_j w_j v_{j+1} - \gamma_j v_j w_{j+1}) = w_j v_j(\lambda - \mu). \tag{5.12}$$

Note that, in both equations above, the right $(\cdot)$ for j is the negative of the left $(\cdot)$ for $j+1$. Sum for $j = 1, 2, \ldots, k < n$ and use the cancellation to find

$$\begin{bmatrix} z_k u_{k+1} & -u_k z_{k+1} \\ -v_k w_{k+1} & w_k v_{k+1} \end{bmatrix} \begin{bmatrix} \gamma_k \\ \beta_k \end{bmatrix} = (\lambda - \mu) \begin{bmatrix} \sum_{j=1}^{k} u_j z_j \\ \sum_{j=1}^{k} w_j v_j \end{bmatrix}. \tag{5.13}$$

Now return to the original eigenvector equations and proceed as above but add rather than subtract. The result is, for $j = 1, \ldots, n$,

$$(\beta_{j-1}u_{j-1}z_j + \gamma_{j-1}z_{j-1}u_j) + 2\alpha_j u_j z_j + (\gamma_j z_j u_{j+1} + \beta_j u_j z_{j+1}) = u_j z_j(\lambda + \mu), \tag{5.14}$$

$$(\gamma_{j-1}v_{j-1}w_j + \beta_{j-1}w_{j-1}v_j) + 2\alpha_j v_j w_j + (\beta_j w_j v_{j+1} + \gamma_j v_j w_{j+1}) = w_j v_j(\lambda + \mu). \tag{5.15}$$

Observe that α_j is determined uniquely by (5.14) or (5.15). Consistency is shown below.

The necessary condition (5.13) defines a unique nonzero vector $\begin{bmatrix} \gamma_k & \beta_k \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \end{bmatrix}$, for $k = 1, \ldots, n-1$, provided that

(A) $\det \begin{bmatrix} z_k u_{k+1} & -u_k z_{k+1} \\ -v_k w_{k+1} & w_k v_{k+1} \end{bmatrix} \neq 0, \quad k = 1, \ldots, n-1,$

(B) $\sum_{j=1}^{k} u_j z_j$ and $\sum_{j=1}^{k} w_j v_j$ do not vanish simultaneously for $k = 1, \ldots, n-1$.

By condition (5.9), $\sum_{j=1}^{n} w_j v_j = 0$ and $\sum_{j=1}^{n} u_j z_j = 0$.

The $4n$ linear equations (5.11), (5.12), (5.14) and (5.15) are equivalent to the original eigenvector equations since only elementary operations, which are reversible, were used to obtain them. Consequently, the conditions *(A)* and *(B)* are necessary conditions on $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ and $\boldsymbol{z}$ to have a unique unreduced C to satisfy the eigenvector equations. What about sufficiency?

When C is not given, equations (5.13), under conditions *(A)* and *(B)*, determine unique $[\gamma_k\ ,\ \beta_k] \neq [0, 0]$, $k = 1, \ldots, n-1$. We have yet to show that $\beta_k \gamma_k \neq 0$, i.e., C is unreduced. We must also show that (5.14) and (5.15), the other half of our equations, are consistent, i.e., determine the same value for α_j.

At this point we need the unused necessary conditions (5.10). By multiplication we obtain

$$\frac{u_{j+1} z_{j+1}}{u_j z_j} = \frac{v_{j+1} w_{j+1}}{v_j w_j} =: \frac{1}{\sigma_j}, \quad j = 1, \ldots, n-1, \tag{5.16}$$

and thus define σ_j. By the nonzero entry condition, rewrite (5.14) and (5.15) as

$$(\beta_{j-1} u_{j-1}/u_j + \gamma_{j-1} z_{j-1}/z_j) + 2\alpha_j + (\gamma_j u_{j+1}/u_j + \beta_j z_{j+1}/z_j) = \lambda + \mu, \tag{5.17}$$

$$(\gamma_{j-1} v_{j-1}/v_j + \beta_{j-1} w_{j-1}/w_j) + 2\alpha_j + (\beta_j v_{j+1}/v_j + \gamma_j w_{j+1}/w_j) = \lambda + \mu. \tag{5.18}$$

Next consider the partial sums in (5.13) and extract the last term in each, using (5.16), for $j = k-1$ down to $j = 1$. This gives

$$\sum_{j=1}^{k} u_j z_j = u_k z_k (1 + \sigma_{k-1} + \sigma_{k-1}\sigma_{k-2} + \ldots + \sigma_{k-1}\sigma_{k-2} \cdots \sigma_1),$$

$$\sum_{j=1}^{k} v_j w_j = v_k w_k (1 + \sigma_{k-1} + \sigma_{k-1}\sigma_{k-2} + \ldots + \sigma_{k-1}\sigma_{k-2} \cdots \sigma_1),$$

which suggests rewriting (5.13) in a more illuminating way as

$$\gamma_k \frac{u_{k+1}}{u_k} - \beta_k \frac{z_{k+1}}{z_k} = \frac{\lambda - \mu}{u_k z_k} \sum_{j=1}^{k} u_j z_j,$$

$$-\gamma_k \frac{w_{k+1}}{w_k} + \beta_k \frac{v_{k+1}}{v_k} = \frac{\lambda - \mu}{w_k v_k} \sum_{j=1}^{k} w_j v_j.$$

Rearranging the two left sides, which were shown to be equal above, yields

$$\gamma_k \left(\frac{u_{k+1}}{u_k} + \frac{w_{k+1}}{w_k} \right) = \beta_k \left(\frac{v_{k+1}}{v_k} + \frac{z_{k+1}}{z_k} \right). \tag{5.19}$$

Now use (5.10) to find that

$$\frac{u_{k+1}}{u_k} / \frac{v_{k+1}}{v_k} = \frac{w_{k+1}}{w_k} / \frac{z_{k+1}}{z_k} = \left(\frac{u_{k+1}}{u_k} + \frac{w_{k+1}}{w_k} \right) / \left(\frac{v_{k+1}}{v_k} + \frac{z_{k+1}}{z_k} \right)$$

and the crucial result

LEMMA 5.4.

$$\gamma_k u_{k+1} v_k = \beta_k v_{k+1} u_k, \qquad \gamma_k w_{k+1} z_k = \beta_k z_{k+1} w_k, \quad k = 1, \ldots, n-1.$$

Thus, with the necessary condition (5.10), Lemma 5.4 shows that β_k and γ_k can only vanish together, while conditions *(A)* and *(B)* forbid that both quantities vanish.

Finally, inspect the two equations (5.17) and (5.18) for α_j and use Lemma 5.4 to see that

$$\gamma_j u_{j+1}/u_j + \beta_j z_{j+1}/z_j = \beta_j v_{j+1}/v_j + \gamma_j w_{j+1}/w_j$$

and

$$\beta_{j-1} u_{j-1}/u_j + \gamma_{j-1} z_{j-1}/z_j = \gamma_{j-1} v_{j-1}/v_j + \beta_{j-1} w_{j-1}/w_j.$$

Thus equations (5.17) and (5.18) are the same and determine the same unique value of α_j, $j = 1, \ldots, n$.

So (A) and (B) are sufficient conditions for the given vectors to be suitable eigenvectors for a unique unreduced C. We have proved

THEOREM 5.5 (**generic case**). *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{z}$ be four vectors in $\mathbb{R}^n$ with no zero entries satisfying $\boldsymbol{v}^T\boldsymbol{w} = \boldsymbol{z}^T\boldsymbol{u} = 0$ and*

$$\frac{u_{j+1}}{u_j}/\frac{v_{j+1}}{v_j} = \frac{w_{j+1}}{w_j}/\frac{z_{j+1}}{z_j}, \quad j = 1, \ldots, n-1.$$

For any $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\lambda \neq \mu$, there exists a unique unreduced $C \in \mathbb{R}^{n\times n}$, as in (2.2), such that $C\boldsymbol{u} = \boldsymbol{u}\lambda$, $\boldsymbol{v}^T C = \lambda \boldsymbol{v}^T$, $C\boldsymbol{w} = \boldsymbol{w}\mu$, $\boldsymbol{z}^T C = \mu \boldsymbol{z}^T$, if and only if

(A) $\det \begin{bmatrix} z_k u_{k+1} & -u_k z_{k+1} \\ -v_k w_{k+1} & w_k v_{k+1} \end{bmatrix} \neq 0, \quad k = 1, \ldots, n-1,$

(B) $\sum_{j=1}^{k} u_j z_j$ *and* $\sum_{j=1}^{k} w_j v_j$ *do not vanish simultaneously for* $k = 1, \ldots, n-1$.

5.3. J matrix with two given real eigenvectors. The result for two real vectors is a little different from the symmetric (T, S) form case and we include the proof.

THEOREM 5.6. *Consider two vectors $\boldsymbol{u} \in \mathbb{R}^n$, $\boldsymbol{v} \in \mathbb{R}^n$ with no zero entries and satisfying $\boldsymbol{v}^T\boldsymbol{u} = 0$. For any distinct $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\lambda \neq \mu$, there is always a unique J matrix, as in (4.20), satisfying $J\boldsymbol{u} = \boldsymbol{u}\lambda$, $\boldsymbol{v}^T J = \mu \boldsymbol{v}^T$, and J is unreduced if and only if*

$$v_k u_{k+1} \neq (\lambda - \mu) \sum_{j=1}^{k} v_j u_j, \quad k = 1, \ldots, n-1.$$

Proof. For brevity define $\beta_0 = \beta_n = 0$, $u_{n+1} = 0$ and $v_0 = 0$.

Since $u_j v_j \neq 0$, multiply the jth equation of $J\boldsymbol{u} = \boldsymbol{u}\lambda$ by v_j and the jth equation of $\boldsymbol{v}^T J = \mu \boldsymbol{v}^T$ by u_j and first subtract and then add to find an equivalent pair of equations,

$$(v_{j-1}u_j - \beta_{j-1} v_j u_{j-1}) + (\beta_j v_{j+1} u_j - v_j u_{j+1}) = v_j u_j (\mu - \lambda) \tag{5.20}$$

$$(v_{j-1}u_j + \beta_{j-1} v_j u_{j-1}) + 2\alpha_j v_j u_j + (\beta_j v_{j+1} u_j + v_j u_{j+1}) = v_j u_j (\mu + \lambda). \tag{5.21}$$

Observe that the second $(\cdot)$ on the left hand side of (5.20) for index $j-1$ is the negative of the first $(\cdot)$ for index j . Now sum (5.20) for $j = 1, \ldots, k$, and use the cancelation to find

$$\beta_k v_{k+1} u_k - v_k u_{k+1} = (\mu - \lambda) \sum_{j=1}^{k} v_j u_j. \tag{5.22}$$

Since no entry of $\boldsymbol{u}$ nor $\boldsymbol{v}$ vanishes, (5.22) determines uniquely each β_k, $k = 1, \ldots, n-1$. Since $\boldsymbol{v}^T \boldsymbol{u} = 0$, an alternative form of (5.22) is

$$\beta_k v_{k+1} u_k - v_k u_{k+1} = (\mu - \lambda) \left(- \sum_{j=k+1}^{n} v_j u_j \right), \tag{5.23}$$

and this arises from summing (5.20) for $j = k+1, \ldots, n$.

Finally, each α_j is determined uniquely by (5.21).

Note that β_k vanishes if and only if

$$v_k u_{k+1} = (\lambda - \mu) \sum_{j=1}^{k} v_j u_j,$$

in which case there is a strong constraint on $\lambda - \mu$. □

6. Up and Down. If digital computers are to be used for computing the partial sums $\sum_{j=1}^{k} s_j |x_j|^2$, $j = 1, \ldots, n$, needed in Lemma 4.1 to compute T from the given eigenvector $\boldsymbol{x}$, then roundoff error is a serious concern because some cancellation in the sums is to be expected. Ideal would be to compute the $|x_i|^2$ exactly and sum them in twice working precision. If extra precision is not available, we have to compute the partial sums with care. One approach is to compute them in two ways, as follows:

$$\begin{aligned} \mathrm{sum}_{\mathrm{down}}(1) &:= s_1 |x_1|^2, \quad \mathrm{sum}_{\mathrm{down}}(k) := \mathrm{sum}_{\mathrm{down}}(k-1) + s_k |x_k|^2, \quad k = 2, \ldots, n, \\ \mathrm{sum}_{\mathrm{up}}(n) &:= s_n |x_n|^2, \quad \mathrm{sum}_{\mathrm{up}}(k) := \mathrm{sum}_{\mathrm{up}}(k+1) + s_k |x_k|^2, \quad k = n-1, \ldots, 1. \end{aligned}$$

In exact arithmetic

$$\mathrm{sum}_{\mathrm{down}}(n) = 0 \qquad \text{and} \qquad \mathrm{sum}_{\mathrm{up}}(1) = 0.$$

With this notation, for each k, $k = 1, \ldots, n-1$, $\mathrm{sum}_{\mathrm{down}}(k) + \mathrm{sum}_{\mathrm{up}}(k+1) = \boldsymbol{x}^* S \boldsymbol{x} = 0$.

We need to combine the two sequences with care. A sensible choice is to switch from $\mathrm{sum}_{\mathrm{down}}$ to $\mathrm{sum}_{\mathrm{up}}$ at any index where $\mathrm{sum}_{\mathrm{down}}$ assumes its maximum absolute value. After this index $\mathrm{sum}_{\mathrm{down}}$ looses information with each cancellation. Similarly, $\mathrm{sum}_{\mathrm{up}}$ looses information with each cancellation going up. In practice, we compute l such that

$$l := \arg\max_k \left| \mathrm{sum}_{\mathrm{down}}(k) \right|$$

and create a spliced array by

$$\begin{aligned} \mathrm{sum}_{\mathrm{spl}}(m) &= \begin{cases} \mathrm{sum}_{\mathrm{down}}(m), & m \le l, \\ -\mathrm{sum}_{\mathrm{up}}(m+1), & l < m < n, \end{cases} \\ \mathrm{sum}_{\mathrm{spl}}(n) &= 0. \end{aligned}$$

More elaborated schemes are possible. One of these uses "compensated summation". This valuable technique is not as universally used as it should be. A good reference is [4, p.93]. In brief, some extra variables and extra operations are used to estimate the quantity that is discarded in each addition. Compensated summation is not exact; nevertheless, in our case, it has proved beneficial for large values of n.

Below we give an example to show how necessary it is to form the sum with care.

EXAMPLE 6.1. *This matrix of order $n = 14$ was created in ST form with*

$$\begin{aligned} T &= \operatorname{diag}(\boldsymbol{a}) + \operatorname{diag}(\boldsymbol{b}, -1) + \operatorname{diag}(\boldsymbol{b}, +1), \\ a_j &= (-1)^{j-1}(2j+3), \quad j = 1, \ldots, 14, \\ \boldsymbol{b} &= [\ 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1], \end{aligned}$$

and $S = \operatorname{diag}\left([1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1]\right)$. *Given the approximate eigenvalue* $\lambda = 16.0 + 7.65 \times 10^{-2}\,\mathrm{i}$ *and the approximate corresponding eigenvector*

$$\boldsymbol{x} = \begin{bmatrix} -3.56 \times 10^{-7} + 3.59 \times 10^{-8}\,\mathrm{i} \\ 3.93 \times 10^{-6} - 3.69 \times 10^{-7}\,\mathrm{i} \\ 3.59 \times 10^{-5} - 3.07 \times 10^{-6}\,\mathrm{i} \\ 8.95 \times 10^{-4} - 7.37 \times 10^{-5}\,\mathrm{i} \\ 2.42 \times 10^{-2} - 1.93 \times 10^{-3}\,\mathrm{i} \\ -7.05 \times 10^{-1} + 5.41 \times 10^{-2}\,\mathrm{i} \\ 7.07 \times 10^{-1} \\ 2.02 \times 10^{-2} - 4.40 \times 10^{-5}\,\mathrm{i} \\ -5.42 \times 10^{-4} + 2.32 \times 10^{-6}\,\mathrm{i} \\ -7.66 \times 10^{-5} - 4.90 \times 10^{-7}\,\mathrm{i} \\ 8.52 \times 10^{-6} + 1.27 \times 10^{-7}\,\mathrm{i} \\ 1.98 \times 10^{-7} + 2.60 \times 10^{-9}\,\mathrm{i} \\ -4.39 \times 10^{-9} - 5.02 \times 10^{-11}\,\mathrm{i} \\ 9.34 \times 10^{-11} + 9.16 \times 10^{-13}\,\mathrm{i} \end{bmatrix},$$

we compute the new $\boldsymbol{b}$ *according to the formula in (4.3) using the values of* $\mathrm{sum}_{\mathrm{down}}$, $\mathrm{sum}_{\mathrm{up}}$ *and* $\mathrm{sum}_{\mathrm{spl}}$. *We show the results in Table 6.1. The new* $\boldsymbol{b}$ *spliced at index 6 is shown in boldface (as well as the relative errors). Table 6.2 shows the different partial sums.*

initial $\boldsymbol{b}$	down $\boldsymbol{b}$	down error	up $\boldsymbol{b}$	up error
-1	$\mathbf{-1.00}$	$\mathbf{3.02 \times 10^{11}}$	-1.06	6.33×10^{-2}
-1	$\mathbf{-1.00}$	$\mathbf{6.61 \times 10^{12}}$	-9.99×10^{-1}	5.24×10^{-4}
-1	$\mathbf{-1.00}$	$\mathbf{4.91 \times 10^{11}}$	-1.00	6.18×10^{-6}
1	$\mathbf{1.00}$	$\mathbf{3.33 \times 10^{16}}$	1.00	1.00×10^{-8}
1	$\mathbf{1.00}$	$\mathbf{2.73 \times 10^{13}}$	1.00	1.35×10^{-11}
1	$\mathbf{1.00}$	$\mathbf{6.44 \times 10^{15}}$	1.00	9.88×10^{-15}
1	1.00	1.99×10^{-11}	$\mathbf{1.00}$	$\mathbf{2.34 \times 10^{14}}$
1	1.00	2.70×10^{-8}	$\mathbf{1.00}$	$\mathbf{8.88 \times 10^{15}}$
1	1.00	1.40×10^{-6}	$\mathbf{1.00}$	$\mathbf{3.22 \times 10^{15}}$
1	1.00	1.12×10^{-4}	$\mathbf{1.00}$	$\mathbf{1.31 \times 10^{14}}$
1	1.21	2.07×10^{-1}	$\mathbf{1.00}$	$\mathbf{9.99 \times 10^{15}}$
1	-4.19×10^{2}	4.20×10^{2}	$\mathbf{1.00}$	$\mathbf{1.44 \times 10^{14}}$
-1	-9.29×10^{5}	9.29×10^{5}	$\mathbf{-1.00}$	$\mathbf{1.11 \times 10^{15}}$

TABLE 6.1
Relative errors for reconstructed $\boldsymbol{b}$ *with* $\mathrm{sum}_{\mathrm{down}}$ *and* $\mathrm{sum}_{\mathrm{up}}$.

k	$\mathrm{sum}_{\mathrm{down}}(k)$	$\mathrm{sum}_{\mathrm{up}}(k+1)$	$\mathrm{sum}_{\mathrm{spl}}(k)$
1	1.28×10^{-13}	-1.36×10^{-13}	1.28×10^{-13}
2	-1.55×10^{-11}	1.55×10^{-11}	-1.55×10^{-11}
3	-1.31×10^{-9}	1.31×10^{-9}	-1.31×10^{-9}
4	8.05×10^{-7}	-8.05×10^{-7}	8.05×10^{-7}
5	-5.90×10^{-4}	5.90×10^{-4}	-5.90×10^{-4}
6	-5.00×10^{-1}	5.00×10^{-1}	-5.00×10^{-1}
7	-4.06×10^{-4}	4.06×10^{-4}	-4.06×10^{-4}
8	3.00×10^{-7}	-3.00×10^{-7}	3.00×10^{-7}
9	5.80×10^{-9}	-5.80×10^{-9}	5.80×10^{-9}
10	-7.26×10^{-11}	7.26×10^{-11}	-7.26×10^{-11}
11	-4.72×10^{-14}	3.91×10^{-14}	-3.91×10^{-14}
12	-8.08×10^{-15}	-1.93×10^{-17}	1.93×10^{-17}
13	-8.10×10^{-15}	8.72×10^{-21}	-8.72×10^{-21}

TABLE 6.2
Partial sums - splice index equals 6.

7. Conclusion. We have given a rather thorough account of all the cases when a unique real tridiagonal matrix can make an approximate eigentriple $(\widehat{\lambda}, \widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}^*)$ exact. Our results suggest two related avenues for future work extending the results in [3]. The first is to exploit the apparent locality of the formula in (4.3) for the off-diagonal entries in the (T, S) formulation to make well chosen perturbations to just a few entries of the computed eigenvector in order to reduce the backward error even further. This work is in its early stages. The second direction is to compare the backward error from several, or all, the computed eigenvectors of a given tridiagonal matrix.

So far the technique shown in Section 6 have given adequate accuracy to the partial sums but our results are limited.

REFERENCES

[1] S. Chandrasekaran and I. C. F. Ipsen, *Backward errors for eigenvalue and singular value decompositions*, Numer. Math., 68 (1994), pp. 215-223.

[2] A. S. Deif, *Realistic a priori and a posteriori error bounds for computed eigenvalues*, IMA J. Numer. Anal., 9 (1990), pp. 323-329.

[3] C. Ferreira, B. Parlett, and F. M. Dopico, *Sensitivity of eigenvalues of an unsymmetric tridiagonal matrix*, Numer. Math., 122 (2012), pp. 527-555.

[4] N. J. Higham, *Accuracy and Stability of Numerical Algorithms, Second Edition*, Society for Industrial and Applied Mathematics, 2002.

[5] D. J. Higham and N. J. Higham, *Structured backward error and condition of generalized eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 493-512.

[6] M. E. Hochstenbach and B. Plestenjak, *Backward error, condition numbers, and pseudospectra for the multiparameter eigenvalue problem*, Linear Algebra Appl., 375 (2003), pp. 63-81.

[7] W. Kahan, B N. Parlett, and E. Jiang, *Residual bounds on approximate eigensystems of nonnormal matrices*, SIAM J. Numer. Anal., 9(3) (1982), pp. 470-484.

[8] Xin-guo Liu and Ze-xi Wang, *A note on the backward errors for Hermite eigenvalue problems*, Appl. Math. Comput., 165(2) (2005), pp. 405-417.

[9] B. N. Parlett and C. Reinsch, *Balancing a matrix for calculation of eigenvalues and eigenvectors*, Numer. Math., 13 (1969), pp. 292-304.

[10] F. Tisseur, *A chart of backward errors for singly and doubly structured eigenvalue problems*, SIAM J. Matrix Anal. Appl., 24(3):877-897, 2003.

[11] F. Tisseur. *Backward error and condition of polynomial eigenvalue problems.* Linear Algebra Appl., 309 (2000), pp. 339-361.

[12] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.