SENSITIVITY OF EIGENVALUES OF AN UNSYMMETRIC TRIDIAGONAL MATRIX *

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

Abstract. Several *relative* eigenvalue condition numbers that exploit tridiagonal form are derived. Some of them use triangular factorizations instead of the matrix entries and so they shed light on when eigenvalues are less sensitive to perturbations of factored forms than to perturbations of the matrix entries. A novel empirical condition number is used to show when perturbations are so large that the eigenvalue response is not linear. Some interesting examples are examined in detail.

Key words. condition numbers, unsymmetric tridiagonal matrices, factored forms

AMS subject classifications. 65F15

1. Introduction. A thoughtful look at Figure 2.1 should persuade the reader that it is not sufficient to produce software to compute eigenvalues; instead each eigenvalue should be acompanied by an estimate of its sensitivity to uncertainty in the matrix. This measure should be realistic, not pessimistic. To this end our measure should respect the zero structure of the matrix and, in line with today's high aspirations, should measure the relative, not absolute, sensitivity in an eigenvalue, that is, it should measure the ratio between the relative variation of an eigenvalue and the largest of the relative variations of each of the parameters defining the matrix.

This paper is mainly focused on infinitesimal perturbations of matrices and their effect on simple eigenvalues. Therefore, we will use derivatives of eigenvalues. For a simple eigenvalue λ of a given matrix $A = (a_{ij})$, the absolute sensitivity with respect to an entry is $\left|\dfrac{\partial\lambda}{\partial a_{ij}}\right|$, whereas the relative sensitivity is $\left|\dfrac{a_{ij}}{\lambda}\dfrac{\partial\lambda}{\partial a_{ij}}\right|$, for λ and a_{ij} nonzero. The relative sensitivity of λ to the whole matrix is the sum of the above sensitivities over the nonzero entries. See section 6.1 for details.

We mention previous work in a later section but here we list what our paper offers: several simple relative eigenvalue condition numbers, a novel empirical condition number which tells us the largest perturbation level at which the response of the eigenvalue is linear in the perturbation, a detailed study of a few challenging matrices. In addition we present some theoretical results that relate our condition numbers to each other. Our main results are given in: (a) Theorem 6.1, where a relative eigenvalue condition number with respect to infinitesimal perturbations of the matrix entries is presented; (b) Theorems 6.3, 6.4 and 6.5, where relative eigenvalue condition numbers with respect to infinitesimal perturbations of the entries of several triangular factorizations of the matrix are presented; (c) Lemmas 6.2 and 6.6, where it is proved that several of the condition numbers that we present are in fact equivalent and that, as a consequence, some representations of tridiagonals do not represent any improvement from the point of view of relative sensitivity of

*The first author is supported by FEDER Funds through "Programa Operacional Factores de Competitividade - COMPETE" and by Portuguese Funds through FCT - "Fundação para a Ciência e a Tecnologia", within the Project PEst-C/MAT/UI0013/2011. The research of F. Dopico was partially supported by the Ministerio de Ciencia e Innovación of Spain through the research grant MTM-2009-09281.

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

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

[§]Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM and Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es).

eigenvalues; and (d) equation (7.1), where the novel empirical condition number is defined.

2. Notation and basic results. We try to follow a widespread practice in Numerical Linear Algebra of reserving capital Roman letters, $A, B, \ldots$, for matrices, (boldfaced) lower case Roman letters, $\boldsymbol{x}, \boldsymbol{y}, \ldots$, for column vectors and lower case Greek letters, $\alpha, \beta, \ldots$, for scalars. We consider real matrices only but some eigenvalues and eigenvectors may be complex. Given a complex column vector $\boldsymbol{y}$ we will be using both $\boldsymbol{y}^T$ and $\boldsymbol{y}^* := \bar{\boldsymbol{y}}^T$, where $\bar{\alpha}$ is the conjugate of α. In order to have vector and matrix norms consistent,

$$\|\boldsymbol{y}\|_\infty = \max_i |y_i| \qquad \text{but} \qquad \|\boldsymbol{y}^*\|_\infty = \|\boldsymbol{y}^T\|_\infty = \|\boldsymbol{y}\|_1 = \sum_i |\boldsymbol{y}_i|.$$

We assume that the reader has some knowledge of matrix eigenvalue problems. We write the eigenvalue/eigenvector equation for an $n \times n$ matrix M as

$$M\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* M = \lambda \boldsymbol{y}^*.$$

Some authors prefer to write the second equation as

$$M^*\boldsymbol{y} = \boldsymbol{y}\bar{\lambda} \qquad (\text{or } M^T\bar{\boldsymbol{y}} = \bar{\boldsymbol{y}}\lambda)$$

but that seems unnatural to us because λ has been replaced by $\bar{\lambda}$. When λ is simple then $\boldsymbol{y}^*\boldsymbol{x} \neq 0$ and this case will be our main focus. The (oblique) spectral projector onto λ's eigenspace is

$$P_\lambda = \boldsymbol{x}(\boldsymbol{y}^*\boldsymbol{x})^{-1}\boldsymbol{y}^* \tag{2.1}$$

and its spectral norm is the *Wilkinson condition number* for λ,

$$\kappa_\lambda := \|P_\lambda\|_2 = \frac{\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2}{|\boldsymbol{y}^*\boldsymbol{x}|} = \frac{1}{|\cos\angle(x,y)|}. \tag{2.2}$$

It can be proved [6] that

$$\kappa_\lambda = \lim_{\eta\to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta} \,:\, (\lambda+\delta\lambda) \text{ is an eigenvalue of } (M+\delta M),\ \|\delta M\|_2 \le \eta \right\}. \tag{2.3}$$

This equality explicitly shows that the Wilkinson condition number measures the absolute sensitivity of a simple eigenvalue with respect to absolute normwise perturbations of the matrix. Since our studies concern relative changes, the standard Wilkinson condition number needs to be replaced by a relative measure with respect to relative normwise perturbations of the matrix, the one used by Bini, Gemignani and Tisseur in [1],

$$BGT(\lambda; M) := \kappa_\lambda \frac{\|M\|_2}{|\lambda|}, \quad \lambda \neq 0. \tag{2.4}$$

Note that κ_λ is invariant under translation $M \to M - \sigma I$, while BGT is not.

Multiple eigenvalues can occur and we will discuss one case later. The *spectrum* of M is its set of eigenvalues and complex eigenvalues occur in conjugate pairs, $M\bar{\boldsymbol{x}} = \bar{\boldsymbol{x}}\bar{\lambda}$. In exact arithmetic the spectrum of real M and M^T are identical; but see Figure 2.1 for a shock. For our special topic of tridiagonal matrices we can use methods that guarantee this spectral identity for M and M^T but we do not know how to achieve it in the general case.

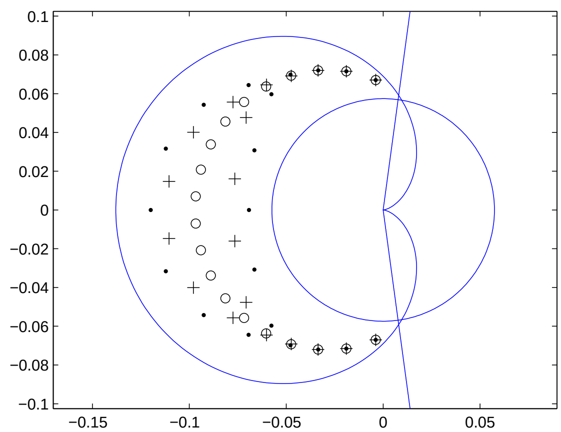

FIGURE 2.1. *Eigenvalues of a generalized Bessel matrix of order $n = 20$, $B_{20}^{(-4.5,2)}$: exact ($\circ$) by Mathematica, approximations by Matlab using the matrix ($\cdot$) and its transpose ($+$).*

3. Relevant studies. There has been a lot of work on relative eigenvalue perturbation theory in the last decades. Some works [7, 10] have developed global bounds for the variation of eigenvalues under finite perturbations of nonsymmetric matrices. In these papers the condition number, for inversion, of the eigenvector matrix plays a crucial role in the bounds. However, our desire is to produce a posteriori sensitivity measures under infinitesimal perturbations for individual eigentriples, i. e., after computing them. These measures are linked to the concept of eigenvalue condition numbers. Formulae for structured condition numbers, both absolute and relative, have been produced by several authors [6, 8, 12] but our goal is to exploit specifically the tridiagonal property, both in terms of the entries of the matrix and in terms of its triangular factors, and the resulting simplifications.

The interesting paper [12] also considers structured relative condition numbers with respect to perturbations of the entries that preserve the zero pattern of the matrix but covers very general patterns. The consequence is that their formulae are less transparent and, in fact, the authors produce code that computes their condition numbers given *any* pattern of zero entries and so cannot exploit the tridiagonal form explicitly. A novel feature of their paper is the use of λ's spectral projector (see (2.1)) to point to entries to which λ is sensitive. Their framework does not permit them to study the factored representation of the matrix.

In the last section of [13] the entries of the relative gradient vector *relgrad*, defined in (6.3), are exhibited but the 1-norm was not taken and so no condition number was defined.

4. Tridiagonal matrices. For such matrices the (i, j) entry vanishes when $|i - j| > 1$. We shall insist that no entry with $|i-j| = 1$ vanishes. Such matrices are said to be *unreduced*. It follows that there is only one linearly independent column (or row) eigenvector for each eigenvalue λ even when λ has eigenvectors $\boldsymbol{v}_j$ of higher grade, i. e., $(M - \lambda I)^{j-1}\boldsymbol{v}_j \neq \mathbf{0}$, $(M - \lambda I)^j \boldsymbol{v}_j = \mathbf{0}$, $j > 1$. In other words, only one Jordan block per eigenvalue. A useful property of tridiagonals is that an eigenvector (grade 1) has nonzero first and last entries. This follows from the fact that the entries in an eigenvector are linked by a three term recurrence. If $C\boldsymbol{x} = \boldsymbol{x}\lambda$, C tridiagonal, then

$$C_{i,i-1}x_{i-1} + (C_{i,i} - \lambda)x_i + C_{i,i+1}x_{i+1} = 0, \quad i = 2, \ldots, n-1.$$

4.1. Balancing. When no off-diagonal entry of a tridiagonal matrix C vanishes then it can be "balanced" easily to give the matrix B shown below, which has the property that, for all i, the ith row and the ith column have the same norm ($\|\cdot\|_1, \|\cdot\|_2$, or $\|\cdot\|_\infty$). In practice balancing is done using only changes in the exponent, so there are no rounding errors. See [14]. We show three unreduced tridiagonal matrices, all related by diagonal similarity transformations,

$$C = \begin{pmatrix} a_1 & c_1 & & & \\ b_1 & a_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & c_{n-1} \\ & & & b_{n-1} & a_n \end{pmatrix}, \quad J = \begin{pmatrix} a_1 & 1 & & & \\ b_1c_1 & a_2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2}c_{n-2} & a_{n-1} & 1 \\ & & & b_{n-1}c_{n-1} & a_n \end{pmatrix} \tag{4.1}$$

and

$$B = \begin{pmatrix} a_1 & \sqrt{|b_1c_1|} & & & \\ \gamma_1\sqrt{|b_1c_1|} & a_2 & \sqrt{|b_2c_2|} & & \\ & \ddots & \ddots & \ddots & \\ & & \gamma_{n-2}\sqrt{|b_{n-2}c_{n-2}|} & a_{n-1} & \sqrt{|b_{n-1}c_{n-1}|} \\ & & & \gamma_{n-1}\sqrt{|b_{n-1}c_{n-1}|} & a_n \end{pmatrix} \tag{4.2}$$

where $\gamma_i = \mathrm{sign}(b_ic_i)$, $i = 1, \ldots, n-1$. It is a good exercise to find the matrices F that specify the transformation. For general applications, C is the input matrix and J uses the fewest parameters.

4.2. Multiple eigenvalues. For most of this paper we will assume that each eigenvalue differs from its neighbors in most of the digits used in the computer. For example, if the precision is 16 decimals then we assume that the nearest neighbors agree to not more than 8. Computing accurate eigenvectors when eigenvalues are much closer than that is a difficult research area and beyond the scope of this paper. These eigenvectors are needed to compute condition numbers. However we must mention that innocent looking unreduced tridiagonals can have multiple eigenvalues and, at the extreme, the spectrum may contain just one eigenvalue. Because the matrix is unreduced that eigenvalue belongs to one $n \times n$ Jordan block.

This extreme case is so sensitive as to defy satisfactory computation using, say, Matlab. Figure 4.1 shows Matlab's output for the following one-point spectrum *Liu* matrix ($n = 14$) with 0 as its only (multiple) eigenvalue:

$$\begin{aligned} Liu_{14} = {} & \mathrm{diag}(0,0,0,0,0,0,-1,1,0,0,0,0,0,0) + \mathrm{diag}_{-1}(1,1,\ldots 1,1) + \\ & + \mathrm{diag}_{+1}(-1,1,1,-1,1,-1,-1,-1,1,-1,1,1,-1). \end{aligned}$$

See Liu [11]. The output shows exactly what is predicted by perturbation theory. See Golub and Van Loan [4], Wilkinson [19] or Stewart and Sun [18]. The computed eigenvalues lie evenly distributed on quite a large circle centered on the exact single eigenvalue, a circle of radius close to $\varepsilon^{1/n}$ with ε the *roundoff unit*. The mean of the computed eigenvalues would be a good approximation - if only we knew that the input matrix had this special property.

It is interesting, and little known, that, in the tridiagonal case, such eigenvalues can be computed in an efficient and accurate way. See Section 3.2 of [3]. We compute the 0 eigenvalue of Liu_{14} exactly and determine its multiplicity. We cannot do this for general matrices.

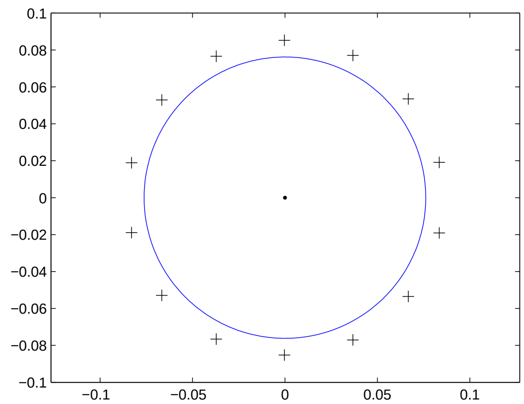

FIGURE 4.1. *Matlab's eigenvalues of the Liu matrix ($n = 14$) and the circle of radius $\sqrt[n]{\varepsilon}$.*

5. Factored forms. A salient property of real tridiagonal matrices is that they can be balanced so easily. See (4.2). However once the matrix is balanced it is not hard to see that it can be made real symmetric simply by changing the signs of certain rows in a clever way. This is not a similarity transformation and would not preserve the eigenvalues. However changing the signs of certain rows is accomplished by premultiplying by a so-called *signature matrix* Δ,

$$\Delta = \operatorname{diag}(\delta_1, \delta_2, \ldots, \delta_n), \qquad \delta_i = \pm 1.$$

Note that Δ is its own inverse, $\Delta^2 = I_n$.

Let B in (4.2) denote the balanced real tridiagonal matrix. Then we have

$$\Delta B = T,$$

T is real symmetric tridiagonal. Consider

$$(B - \lambda I)\boldsymbol{x} = \mathbf{0}$$

and premultiply by Δ to find

$$(T - \lambda \Delta)\boldsymbol{x} = \mathbf{0}. \tag{5.1}$$

If B has complex eigenvalues then both T and Δ are indefinite, i.e., $\boldsymbol{v}^T T \boldsymbol{v}$ and $\boldsymbol{u}^T \Delta \boldsymbol{u}$ are positive for some vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and negative for others. Often, but not always, T admits triangular factorization as

$$T = LDL^T \tag{5.2}$$

where L is unit lower bidiagonal,

$$L = \begin{pmatrix} 1 & & & & \\ l_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & l_{n-2} & 1 & \\ & & & l_{n-1} & 1 \end{pmatrix},$$

and $D = \mathrm{diag}(d_1, d_2, \dots, d_n)$ is often called the matrix of pivots. See [5]. This gives a new representation,

$$
\begin{aligned}
B &= \Delta T = \Delta L D L^T, && (5.3)\\
B^T &= T\Delta = L D L^T \Delta. && (5.4)
\end{aligned}
$$

Even when the factorization (5.2) does not breakdown there can be large element growth, i.e.,

$$\|D\|_2 \gg \|T\|_2, \qquad \|L\|_2 \gg \|T\|_2.$$

Such growth can be avoided by allowing D to have 2×2 blocks on the diagonal but we don't pursue this technique here.

Recall that B has the same eigenvalues (in exact arithmetic) as the input matrix C and their eigenvectors are simply related.

$$B = FCF^{-1} \quad \text{and} \quad B\boldsymbol{x} = \boldsymbol{x}\lambda$$

imply

$$C\left(F^{-1}\boldsymbol{x}\right) = F^{-1}B\boldsymbol{x} = \left(F^{-1}\boldsymbol{x}\right)\lambda, \qquad F \text{ is diagonal.}$$

It is not obligatory to balance a matrix and so we consider a related representation. We call J in (4.1) the J-form of C. Suppose that J admits triangular factorization

$$J = \mathcal{L}\mathcal{U}$$

where $\mathcal{L}$ ($\neq L$ above) and $\mathcal{U}$ are lower and upper bidiagonals, respectively, of the form

$$\mathcal{L} = \begin{pmatrix} 1 & & & & \\ \mathfrak{l}_1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \mathfrak{l}_{n-2} & 1 & \\ & & & \mathfrak{l}_{n-1} & 1 \end{pmatrix}, \qquad \mathcal{U} = \begin{pmatrix} \mathfrak{u}_1 & 1 & & & \\ & \mathfrak{u}_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \mathfrak{u}_{n-1} & 1 \\ & & & & \mathfrak{u}_n \end{pmatrix}. \qquad (5.5)$$

Warning. The difference between the symbols for the parameters in L and in $\mathcal{L}$ is subtle.

Among other questions, we ask when do the parameters in $\mathcal{L}$ and $\mathcal{U}$ determine J's eigenvalues better than the entries in J.

We could derive relative condition numbers for eigenvalues of C, J and B with respect to tiny relative perturbations of each entry but they turn out to be equal (see Lemma 6.2). It also turns out that the relative condition numbers for the various factored forms ($J = \mathcal{L}\mathcal{U}, B = \Delta T = \Delta LDL^T$) are equivalent. See Section 6.6. So we only derive relative condition numbers for C and for the factors $\mathcal{L}, \mathcal{U}$ and present the formulae for the other factored forms.

6. Derivatives and condition numbers. A simple eigenvalue λ is a smooth function of the matrix entries [18, Ch IV, Theorem 2.3] and the natural measure of λ's sensitivity is some suitable norm of the gradient vector. See (6.1) for the tridiagonal case. In [16] J. Rice published a very abstract theory of condition, both absolute and relative, but, as far as we can tell, it fails to explain why, for the relative condition of λ, we need the 1-norm of the relative gradient vector *relgrad* (see (6.3)). For this reason we have included the natural derivation of *relcond*(λ), the relative condition number for λ, which depends strongly on (6.4).

6.1. Relative versus absolute. Entries versus norms. The classical viewpoint suggests that the best that can be expected in numerical practice is "absolute" sensitivity, i.e., a measure of the absolute variation of an eigenvalue with respect to the norm of the matrix. This is what the Wilkinson condition number reflects according to (2.3). We develop here a measure of "relative" sensitivity, i.e., a measure of the relative variation of an eigenvalue with respect to the largest relative perturbation of each of the nonzero entries of the matrix. If the three diagonals defining tridiagonal C are in arrays $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ (see (4.1)), then for a simple eigenvalue λ one considers all the partial derivatives $\left\{\frac{\partial\lambda}{\partial a_i}, \frac{\partial\lambda}{\partial b_i}, \frac{\partial\lambda}{\partial c_i}\right\}$ and forms the absolute gradient vector

$$\mathrm{grad}_C(\lambda) = \left(\frac{\partial\lambda}{\partial a_1}, \dots, \frac{\partial\lambda}{\partial a_n}, \frac{\partial\lambda}{\partial b_1}, \dots, \frac{\partial\lambda}{\partial b_{n-1}}, \frac{\partial\lambda}{\partial c_1}, \dots, \frac{\partial\lambda}{\partial c_{n-1}}\right)^T. \tag{6.1}$$

For infinitesimal absolute changes $(\delta a_1, \dots, \delta a_n, \delta b_1, \dots, \delta b_{n-1}, \delta c_1, \dots, \delta c_{n-1})^T =: \delta C$, we have

$$\delta\lambda = \mathrm{grad}_C(\lambda)^T \cdot \delta C + \textit{higher order terms (h.o.t.)}. \tag{6.2}$$

In order to turn (6.2) into relative terms we want to relate $|\delta\lambda/\lambda|$ to $|\delta p_j/p_j|, j = 1, \dots, 3n-2$, $\boldsymbol{p} = \big(a_1, \dots, a_n, b_1, \dots, b_{n-1}, c_1, \dots, c_{n-1}\big)$. Assuming no zeros, rewrite (6.2) as

$$\begin{aligned}\frac{\delta\lambda}{\lambda} &= \left(\frac{a_1}{\lambda}\frac{\partial\lambda}{\partial a_1}, \dots, \frac{c_{n-1}}{\lambda}\frac{\partial\lambda}{\partial c_{n-1}}\right) \cdot \left(\frac{\delta a_1}{a_1}, \dots, \frac{\delta c_{n-1}}{c_{n-1}}\right)^T + h.o.t. \\ &=: \mathit{relgrad}_C(\lambda)^T \cdot \mathit{rel}\,\delta C + h.o.t.,\end{aligned} \tag{6.3}$$

defining the *relative gradient* and the *relative perturbation*. When a parameter vanishes we should omit the corresponding term in the inner product.

The perturbations we consider are of the form $|\delta p_i| \le \eta |p_i|$. As a matrix, entry by entry,

$$|\delta C| \le \eta |C|, \qquad 0 < \eta \ll 1, \tag{6.4}$$

and

$$|\,\mathit{rel}\,\delta C| \le \eta\,(1, 1, \dots 1)^T.$$

Recall that $|C|_{ij} = |C_{ij}|$. We say that the *level of relative perturbation* is η.

To avoid unnecessary factors of n we use $\|\cdot\|_\infty$ for $\mathit{rel}\,\delta C$. Thus

$$\|\,\mathit{rel}\,\delta C\|_\infty \le \eta.$$

and, since $|\boldsymbol{u}^T\boldsymbol{v}| \le \|\boldsymbol{u}\|_1\|\boldsymbol{v}\|_\infty$ (Hölder inequality),

$$\left|\frac{\delta\lambda}{\lambda}\right| \le \|\,\mathit{relgrad}_C(\lambda)\|_1 \|\,\mathit{rel}\,\delta C\|_\infty + h.o.t. \le \eta \|\,\mathit{relgrad}_C(\lambda)\|_1 + h.o.t. \tag{6.5}$$

and we define the *structured relative condition number* for λ as a function of C by

$$\boxed{\mathit{relcond}(\lambda; C) := \|\,\mathit{relgrad}_C(\lambda)\|_1, \quad \lambda \ne 0.}$$

There is no reason to expect $\mathit{relcond}(\lambda; C) > 1$; any value in $[0, +\infty[$ could occur. For representations other than the matrix entries we also use (6.3) with the appropriate *relgrad* including derivatives with respect to the parameters defining the representation.

Observe that standard properties of norms [5] guarantee that there exist particular vectors $rel\,\delta C$ with $\|\, rel\,\delta C\|_\infty = \eta$ such that $|\, relgrad_C(\lambda)^T \cdot rel\,\delta C| = \|\, relgrad_C(\lambda)\|_1 \|\, rel\,\delta C\|_\infty$. For these vectors $rel\,\delta C$, the bound (6.5) is attained to first order in η, and this allows us to prove immediately that

$$\begin{aligned} relcond(\lambda; C) &:= \|\, relgrad_C(\lambda)\|_1 \\ &= \lim_{\eta\to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} \,:\, (\lambda+\delta\lambda) \text{ is an eigenvalue of } (C+\delta C),\ |\delta C| \le \eta |C| \right\}. \end{aligned} \quad (6.6)$$

The expression in (6.6) is often given as definition of componentwise condition number for simple eigenvalues [6]. Our contribution here has been to prove that the abstract expression (6.6) is equal to $\|\, relgrad_C(\lambda)\|_1$, which is easily computable. This property extends directly to representations different than the matrix entries.

Should $C \neq O$ be singular then appropriate independent relative changes to the entries will destroy singularity. So we set $relcond(0; C) = \infty$. Our other representations will have finite values for $\|\, relgrad(\lambda)\|_1$ and thus will define tiny eigenvalues to high relative accuracy, a very desirable property.

Warning. We do not know in advance when η is small enough to warrant the neglect of *h.o.t.*. Our numerical examples shed light on this topic. We know of no other study that addresses it.

6.2. Representation 1 - entries of C. We now derive an explicit expression for $relcond(\lambda; C)$. We treat each component of C as an independent variable. Thus, with $I = \big(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\big)$,

$$\frac{\partial C}{\partial a_j} = \boldsymbol{e}_j \boldsymbol{e}_j^T, \quad \frac{\partial C}{\partial b_j} = \boldsymbol{e}_{j+1}\boldsymbol{e}_j^T \quad \text{and} \quad \frac{\partial C}{\partial c_j} = \boldsymbol{e}_j \boldsymbol{e}_{j+1}^T.$$

Let λ be a simple nonzero eigenvalue of C and

$$C\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* C = \lambda \boldsymbol{y}^*.$$

Then, for $p_j = a_j, b_j, c_j$, we differentiate $C\boldsymbol{x} = \boldsymbol{x}\lambda$ to get

$$\frac{\partial C}{\partial p_j}\boldsymbol{x} + C\frac{\partial \boldsymbol{x}}{\partial p_j} = \frac{\partial \boldsymbol{x}}{\partial p_j}\lambda + \boldsymbol{x}\frac{\partial \lambda}{\partial p_j}.$$

Multiply by $\boldsymbol{y}^*$ and cancel equal terms to find

$$\frac{\partial \lambda}{\partial p_j}\boldsymbol{y}^*\boldsymbol{x} = \boldsymbol{y}^*\frac{\partial C}{\partial p_j}\boldsymbol{x}, \quad p_j = a_j, b_j, c_j.$$

Thus,

$$\frac{\partial \lambda}{\partial a_j} = \frac{\overline{y_j} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial b_j} = \frac{\overline{y_{j+1}} x_j}{\boldsymbol{y}^*\boldsymbol{x}}, \quad \frac{\partial \lambda}{\partial c_j} = \frac{\overline{y_j} x_{j+1}}{\boldsymbol{y}^*\boldsymbol{x}}$$

and

$$relgrad_C(\lambda) = \frac{1}{\lambda \boldsymbol{y}^*\boldsymbol{x}}\big(a_1\overline{y_1}x_1, \ldots, a_n\overline{y_n}x_n, b_1\overline{y_2}x_1, \ldots, b_{n-1}\overline{y_n}x_{n-1}, c_1\overline{y_1}x_2, \ldots, c_{n-1}\overline{y_{n-1}}x_n\big)^T.$$

Finally, observe that

$$\|\, relgrad_C(\lambda)\|_1 = \frac{1}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}\left(\sum_{j=1}^{n}|a_j||y_j||x_j| + \sum_{j=1}^{n-1}\big(|b_j||y_{j+1}||x_j| + |c_j||y_j||x_{j+1}|\big)\right) = \frac{|\boldsymbol{y}|^T|C||\boldsymbol{x}|}{|\lambda||\boldsymbol{y}^*\boldsymbol{x}|}.$$

According to equation (6.6), the arguments above prove Theorem 6.1 that provides an *eigenvalue relative condition number* for C.

THEOREM 6.1. *Let $\lambda \neq 0$ be a simple eigenvalue of an unreduced real tridiagonal matrix C with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. Then $relcond(\lambda; C) := \| relgrad_C(\lambda)\|_1$ is equal to*

$$relcond(\lambda; C) = \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} : (\lambda + \delta\lambda) \textit{ is an eigenvalue of } (C + \delta C),\ |\delta C| \leq \eta |C| \right\}$$
$$= \frac{|\boldsymbol{y}|^T |C| |\boldsymbol{x}|}{|\lambda| |\boldsymbol{y}^* \boldsymbol{x}|}.$$

Note that the numerator of $relcond(\lambda; C)$ is of the form $(row) \cdot (matrix) \cdot (column)$. If $\lambda = 0$, $relcond(\lambda; C) = \infty$. This expression for $relcond(\lambda; C)$ turns out to be an instance of Theorem 3.2 in Higham and Higham [6]. A rather similar expression has also appeared in [12].

The form of $relcond(\lambda; C)$ yields the following result.

LEMMA 6.2. *For any scaling matrix S invertible and diagonal,*

$$relcond(\lambda; SCS^{-1}) = relcond(\lambda; C).$$

Proof. Let $G = SCS^{-1}$ so that, in the notation above,

$$\boldsymbol{y}^* = \boldsymbol{y}_G^* S \quad \text{and} \quad \boldsymbol{x} = S^{-1} \boldsymbol{x}_G.$$

Consequently, $|\boldsymbol{y}^* \boldsymbol{x}| = |\boldsymbol{y}_G^* \boldsymbol{x}_G|$ and

$$|\boldsymbol{y}^*||C||\boldsymbol{x}| = |\boldsymbol{y}_G^* S||C||S^{-1}\boldsymbol{x}_G| = |\boldsymbol{y}_G^*||S||C||S^{-1}||\boldsymbol{x}_G| = |\boldsymbol{y}_G^*||G||\boldsymbol{x}_G| = |\boldsymbol{y}_G|^T |G||\boldsymbol{x}_G|,$$

because S is diagonal and no additions occur in SCS^{-1}. $\square$

This result shows that the relative condition number we derive is invariant under diagonal similarity transformations. In contrast, neither κ_λ nor BGT are invariant under these transformations. Thus, for matrices J and B in Section 4.1,

$$relcond(\lambda; C) = relcond(\lambda; J) = relcond(\lambda; B)$$

and there is no improvement in the relative condition number by balancing a tridiagonal matrix C to get B. It does not follow that it is a waste of time to balance a tridiagonal matrix before computing its eigentriples [14]. Some tridiagonal test matrices (for example, Lesp and Clement matrices [9, 2]) have big Wilkinson condition numbers but are symmetrizable. Eigenvalue techniques for symmetric tridiagonals are more efficient and accurate than those for general tridiagonals.

6.3. Representation 2 - $\mathcal{L}$, $\mathcal{U}$ representation of J. Assume that J permits triangular factorization $J = \mathcal{L}\mathcal{U}$ (see (5.5)) and λ is a simple eigenvalue,

$$\mathcal{L}\mathcal{U}\boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^*\mathcal{L}\mathcal{U} = \lambda \boldsymbol{y}^*, \qquad \lambda \neq 0.$$

Recall that $\mathcal{L} = I + \mathring{\mathcal{L}}$ and $\mathcal{U} = \mathrm{diag}(\mathbf{u}) + N$ with

$$\mathring{\mathcal{L}} = \begin{pmatrix} 0 & & & & \\ \mathfrak{l}_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \mathfrak{l}_{n-2} & 0 & \\ & & & \mathfrak{l}_{n-1} & 0 \end{pmatrix} \quad \text{and} \quad N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

The perturbations at level η are given by

$$\begin{aligned} |\delta \mathfrak{l}_i| &\le \eta |\mathfrak{l}_i|, \qquad 0 < \eta \ll 1,\\ |\delta \mathfrak{u}_i| &\le \eta |\mathfrak{u}_i|, \qquad 0 < \eta \ll 1. \end{aligned}$$

The 1's are not changed.

Next we derive an explicit expression for the relative condition number for $J = \mathcal{L}\mathcal{U}$,

$$\mathit{relcond}(\lambda; \mathcal{L}, \mathcal{U}) := \| \mathit{relgrad}_{\mathcal{L},\mathcal{U}}(\lambda)\|_1.$$

For $\mathfrak{u}_j$ we find

$$\frac{\partial \lambda}{\partial \mathfrak{u}_j} \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{y}^* \mathcal{L} \frac{\partial \mathcal{U}}{\partial \mathfrak{u}_j} \boldsymbol{x} = \boldsymbol{y}^* \mathcal{L} \boldsymbol{e}_j \boldsymbol{e}_j^T \boldsymbol{x} = (\boldsymbol{y}^* \mathcal{L})_j x_j, \quad j = 1, \dots, n,$$

and for $\mathfrak{l}_j$,

$$\frac{\partial \lambda}{\partial \mathfrak{l}_j} \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{y}^* \frac{\partial \mathcal{L}}{\partial \mathfrak{l}_j} \mathcal{U} \boldsymbol{x} = \boldsymbol{y}^* \boldsymbol{e}_{j+1} \boldsymbol{e}_j^T \mathcal{U} \boldsymbol{x} = \overline{y_{j+1}} (\mathcal{U}\boldsymbol{x})_j, \quad j = 1, \dots, n-1.$$

Then

$$\operatorname{grad}_{\mathcal{L},\mathcal{U}}(\lambda) = \left(\frac{\partial \lambda}{\partial \mathfrak{u}_1}, \dots, \frac{\partial \lambda}{\partial \mathfrak{u}_n}, \frac{\partial \lambda}{\partial \mathfrak{l}_1}, \dots, \frac{\partial \lambda}{\partial \mathfrak{l}_{n-1}} \right)^T \tag{6.7}$$

and, inserting the parameters $\mathfrak{l}_j$ and $\mathfrak{u}_j$ appropriately,

$$\begin{aligned} \lambda(\boldsymbol{y}^*\boldsymbol{x})\, \mathit{relgrad}_{\mathcal{L}\mathcal{U}}(\lambda) =& \big((\boldsymbol{y}^*\mathcal{L})_1 \mathfrak{u}_1 x_1, \dots, (\boldsymbol{y}^*\mathcal{L})_n \mathfrak{u}_n x_n, \overline{y_2}\mathfrak{l}_1 (\mathcal{U}\boldsymbol{x})_1, \dots, \overline{y_n}\mathfrak{l}_{n-1} (\mathcal{U}\boldsymbol{x})_{n-1} \big)^T \\ =& \big((\boldsymbol{y}^*\mathcal{L})_1 \mathfrak{u}_1 x_1, \dots, (\boldsymbol{y}^*\mathcal{L})_n \mathfrak{u}_n x_n, (\boldsymbol{y}^*\mathring{\mathcal{L}})_1 (\mathcal{U}\boldsymbol{x})_1, \dots, (\boldsymbol{y}^*\mathring{\mathcal{L}})_{n-1} (\mathcal{U}\boldsymbol{x})_{n-1} \big)^T. \end{aligned}$$

So, observing that $|\boldsymbol{y}^*\mathring{\mathcal{L}}||\mathcal{U}\boldsymbol{x}| = |\boldsymbol{y}|^T |\mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x}|$ in the last line below,

$$\begin{aligned} |\lambda||\boldsymbol{y}^*\boldsymbol{x}| \| \mathit{relgrad}_{\mathcal{L},\mathcal{U}}(\lambda)\|_1 &= \sum_{j=1}^{n} |(\boldsymbol{y}^*\mathcal{L})_j \mathfrak{u}_j x_j| + \sum_{j=1}^{n-1} |(\boldsymbol{y}^*\mathring{\mathcal{L}})_j (\mathcal{U}\boldsymbol{x})_j| \\ &= |\boldsymbol{y}^*\mathcal{L}||\operatorname{diag}(\mathbf{u})\boldsymbol{x}| + |\boldsymbol{y}^*\mathring{\mathcal{L}}||\mathcal{U}\boldsymbol{x}| \\ &= |\boldsymbol{y}^*\mathcal{L}\operatorname{diag}(\mathbf{u})||\boldsymbol{x}| + |\boldsymbol{y}|^T|\mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x}|. \end{aligned} \tag{6.8}$$

As we did in Theorem 6.1, we can summarize the arguments above in Theorem 6.3.

THEOREM 6.3. *Let J be an unreduced real tridiagonal matrix that permits a triangular factorization $J = \mathcal{L}\mathcal{U}$ with factors as in* (5.5)*, let $\mathring{\mathcal{L}} = \mathcal{L} - \mathcal{I}$, and let* $\operatorname{diag}(\mathbf{u}) = \operatorname{diag}(\mathfrak{u}_1, \dots, \mathfrak{u}_n)$*. Let $\lambda \neq 0$ be a simple eigenvalue of J with left eigenvector $\boldsymbol{y}$ and right eigenvector $\boldsymbol{x}$. Then $\mathit{relcond}(\lambda; \mathcal{L}, \mathcal{U}) := \| \mathit{relgrad}_{\mathcal{L},\mathcal{U}}(\lambda)\|_1$ is equal to*

$$\begin{aligned} \mathit{relcond}(\lambda; \mathcal{L}, \mathcal{U}) = \lim_{\eta \to 0} \sup \Big\{ \frac{|\delta\lambda|}{\eta|\lambda|} \,:\, &(\lambda + \delta\lambda) \text{ is an eigenvalue of } (\mathcal{L} + \delta\mathcal{L})(\mathcal{U} + \delta\mathcal{U}), \\ &|\delta\mathcal{L}| \le \eta|\mathring{\mathcal{L}}|, \quad |\delta\mathcal{U}| \le \eta|\operatorname{diag}(\mathbf{u})| \Big\} \\ = \frac{|\boldsymbol{y}^*\mathcal{L}\operatorname{diag}(\mathbf{u})||\boldsymbol{x}| + |\boldsymbol{y}|^T|\mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x}|}{|\lambda|\,|\boldsymbol{y}^*\boldsymbol{x}|}&. \end{aligned}$$

Observe that neither the 1's nor the 0's of $\mathcal{L}$ and $\mathcal{U}$ are changed in the perturbations.

Next, we express $relcond(\lambda; \mathcal{L}, \mathcal{U})$ in a form that is more convenient for computing it. To this end, write

$$\mathcal{U} = \operatorname{diag}(\mathfrak{u}) \left(I + \mathring{\mathcal{U}}\right)$$

where

$$\mathring{\mathcal{U}} = \begin{pmatrix} 0 & \mathfrak{u}_1^{-1} & & & \\ & 0 & \mathfrak{u}_2^{-1} & & \\ & & \ddots & \ddots & \\ & & & 0 & \mathfrak{u}_{n-1}^{-1} \\ & & & & 0 \end{pmatrix}.$$

In order to extract a factor of $|\lambda|$ on the right in (6.8), use

$$\boldsymbol{y}^* \mathcal{L}\mathcal{U} = \lambda \boldsymbol{y}^*, \quad \mathcal{U}\boldsymbol{x} = \mathcal{L}^{-1}\boldsymbol{x}\lambda, \quad \lambda \neq 0,$$

to find

$$\boldsymbol{y}^* \mathcal{L}\operatorname{diag}(\mathfrak{u}) = \lambda \boldsymbol{y}^* \left(I + \mathring{\mathcal{U}}\right)^{-1}, \quad \mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x} = \mathring{\mathcal{L}}\mathcal{L}^{-1}\boldsymbol{x}\lambda, \quad \lambda \neq 0. \tag{6.9}$$

Also use $\mathring{\mathcal{L}} = \mathcal{L} - I$ to see that $\mathring{\mathcal{L}}\mathcal{L}^{-1} = I - \mathcal{L}^{-1} = \mathcal{L}^{-1}\mathring{\mathcal{L}}$. Substitute this relation in (6.9) to obtain

$$\boldsymbol{y}^* \mathcal{L}\operatorname{diag}(\mathfrak{u}) = \lambda \boldsymbol{y}^* \left(I + \mathring{\mathcal{U}}\right)^{-1}, \quad \mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x} = \mathcal{L}^{-1}\mathring{\mathcal{L}}\boldsymbol{x}\lambda, \quad \lambda \neq 0. \tag{6.10}$$

If $\mathcal{L}\mathcal{U}$ exists, $\mathfrak{u}_j \neq 0$, $j = 1, \ldots, n-1$. Fortunately $\mathfrak{u}_n$ does not appear in $\mathring{\mathcal{U}}$. Substitute the expressions in (6.10) into (6.8) and cancel $|\lambda|$ ($\neq 0$) to find

$$|\boldsymbol{y}^*\boldsymbol{x}| \| \, relgrad_{\mathcal{L}\mathcal{U}}(\lambda)\|_1 = \left|\boldsymbol{y}^* \left(I + \mathring{\mathcal{U}}\right)^{-1}\right| |\boldsymbol{x}| + |\boldsymbol{y}|^T |\mathcal{L}^{-1}\mathring{\mathcal{L}}\boldsymbol{x}|.$$

For the cost of solving two bidiagonal linear systems

$$\boldsymbol{v}^* \left(I + \mathring{\mathcal{U}}\right) = \boldsymbol{y}^* \text{ for } \boldsymbol{v}^* \quad \text{and} \quad \mathcal{L}\boldsymbol{w} = \mathring{\mathcal{L}}\boldsymbol{x} \text{ for } \boldsymbol{w}$$

we obtain the following expression of the relative condition number for $J = \mathcal{L}\mathcal{U}$

$$\boxed{relcond(\lambda; \mathcal{L}, \mathcal{U}) := \frac{|\boldsymbol{v}|^T |\boldsymbol{x}| + |\boldsymbol{y}|^T |\boldsymbol{w}|}{|\boldsymbol{y}^*\boldsymbol{x}|}.} \tag{6.11}$$

Although the right hand side of (6.11) is a nonzero finite number for a simple eigenvalue $\lambda = 0$, observe that the perturbations we consider for $\mathcal{U}$, that is, $|\delta \mathfrak{u}_i| \leq \eta |\mathfrak{u}_i|$, produce $\mathfrak{u}_n + \delta\mathfrak{u}_n = 0$ whenever $\mathfrak{u}_n = 0$. This means that singularity is preserved or, equivalently, that the zero eigenvalue is preserved. Therefore, it seems appropriate to set $relcond(0; \mathcal{L}, \mathcal{U}) = 0$.

6.4. Other representations. We now present the sensitivity of a simple eigenvalue λ w.r.t. D and L, keeping Δ constant. Assume that ΔT admits triangular factorization $\Delta T = \Delta LDL^T$ and let λ be a simple eigenvalue,

$$\Delta LDL^T \boldsymbol{x} = \boldsymbol{x}\lambda, \quad \boldsymbol{y}^* \Delta LDL^T = \lambda \boldsymbol{y}^*, \quad \lambda \neq 0.$$

Recall that $\boldsymbol{x}$, $\boldsymbol{y}$ and λ may be complex, $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$ and $L = I + \mathring{L}$ with

$$\mathring{L} = \begin{pmatrix} 0 & & & & \\ l_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & l_{n-2} & 0 & \\ & & & l_{n-1} & 0 \end{pmatrix}.$$

When the matrix is balanced $\boldsymbol{y}^*$ is determined by $\boldsymbol{x}$. Transpose $\Delta LDL^T \boldsymbol{x} = \boldsymbol{x}\lambda$ and insert $I = \Delta^2$ to find

$$\left(\boldsymbol{x}^T \Delta\right) \left(\Delta LDL^T\right) = \lambda \left(\boldsymbol{x}^T \Delta\right)$$

and compare with $\boldsymbol{y}^* \Delta LDL^T = \lambda \boldsymbol{y}^*$ to see that $\boldsymbol{y}^* = \boldsymbol{x}^T \Delta$. Recall that $\boldsymbol{x}$ and λ may be complex and, since λ is simple, $0 \neq \boldsymbol{y}^* \boldsymbol{x} = \boldsymbol{x}^T \Delta \boldsymbol{x}$.

Following the analysis of (6.8) but using partial derivatives of λ with respect to $d_1, \ldots, d_n$ and $l_1, \ldots, l_{n-1}$ for this case, we find

$$\begin{aligned} |\lambda| |\boldsymbol{x}^T \Delta \boldsymbol{x}| \| \, \mathit{relgrad}_{L,D}(\lambda) \|_1 &= \sum_{j=1}^{n} \left| d_j [(L^T \boldsymbol{x})_j]^2 \right| + 2 \sum_{j=1}^{n-1} |(\boldsymbol{x}^T \mathring{L})_j (DL^T \boldsymbol{x})_j| \\ &= |L^T \boldsymbol{x}|^T |DL^T \boldsymbol{x}| + 2 |\boldsymbol{x}^T \mathring{L}| |DL^T \boldsymbol{x}| \\ &= \left(|\boldsymbol{x}^T L| + 2 |\boldsymbol{x}^T \mathring{L}| \right) |DL^T \boldsymbol{x}|. \end{aligned} \tag{6.12}$$

As we did in Theorems 6.1 and 6.3, we express formally this result in Theorem 6.4.

THEOREM 6.4. *Let B be a balanced unreduced real tridiagonal matrix that permits a triangular factorization $B = \Delta LDL^T$ with factors as in* (5.3) *and let $\mathring{L} = L - I$. Let $\lambda \neq 0$ be a simple eigenvalue of B with right eigenvector $\boldsymbol{x}$. Then $\mathit{relcond}(\lambda; L, D) := \| \, \mathit{relgrad}_{L,D}(\lambda) \|_1$ is equal to*

$$\begin{aligned} \mathit{relcond}(\lambda; L, D) &= \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta |\lambda|} : (\lambda + \delta\lambda) \text{ is an eigenvalue of } \Delta(L + \delta L)(D + \delta D)(L + \delta L)^T, \right. \\ &\qquad \left. |\delta L| \leq \eta |\mathring{L}|, \ |\delta D| \leq \eta |D| \right\} \\ &= \frac{\left(|\boldsymbol{x}^T L| + 2 |\boldsymbol{x}^T \mathring{L}| \right) |DL^T \boldsymbol{x}|}{|\lambda| \, |\boldsymbol{x}^T \Delta \boldsymbol{x}|}. \end{aligned}$$

Observe that neither the 1's nor the 0's of L and D are changed in the perturbations.

Next we express $\mathit{relcond}(\lambda; L, D)$ in a form that is more convenient for computations. To this purpose, we can extract a factor $|\lambda|$ on the right side of (6.12) using, from $\Delta LDL^T \boldsymbol{x} = \boldsymbol{x}\lambda$,

$$DL^T \boldsymbol{x} = L^{-1} \Delta \boldsymbol{x} \lambda. \tag{6.13}$$

Substitute (6.13) in (6.12) and cancel $|\lambda|$ ($\neq 0$) to obtain

$$|\boldsymbol{x}^T \Delta \boldsymbol{x}| \|\, relgrad_{L,D}(\lambda)\|_1 = \left(|\boldsymbol{x}^T L| + 2|\boldsymbol{x}^T \mathring{L}|\right) |L^{-1}\Delta \boldsymbol{x}|.$$

For the cost of solving one bidiagonal linear system (L is unit bidiagonal)

$$L\boldsymbol{v} = \Delta \boldsymbol{x} \text{ for } \boldsymbol{v},$$

we obtain the following expression of the relative condition number for $\Delta T = \Delta L D L^T$

$$\boxed{relcond(\lambda; L, D) := \frac{\left(|\boldsymbol{x}^T L| + 2|\boldsymbol{x}^T (L-I)|\right)|\boldsymbol{v}|}{|\boldsymbol{x}^T \Delta \boldsymbol{x}|}.} \tag{6.14}$$

Note that, in general, $|\boldsymbol{x}^T L| + 2|\boldsymbol{x}^T \mathring{L}| \neq |\boldsymbol{x}|^T |L + 2\mathring{L}|$ even though $L = I + \mathring{L}$.

Again, the right hand side of (6.14) is a finite nonzero number for a simple eigenvalue $\lambda = 0$, but the perturbations we consider for D preserve the singularity of D and the zero eigenvalue. Therefore, we set $relcond(0; L, D) = 0$.

Closely related to the factorization $T = LDL^T$ is the factorization $T = \bar{L}\Omega\bar{L}^T$ where $\bar{L} = L|D|^{1/2}$ is lower bidiagonal, $D = |D|^{1/2}\Omega|D|^{1/2}$, $\Omega = \mathrm{diag}(\mathrm{sign}(d_i))$ with $\mathrm{sign}(d_n) = 1$ if $d_n = 0$. This factorization is the closest to the Cholesky factorization that we can get. Now our tridiagonal eigenproblem is associated with two independent signature matrices, Δ, Ω. Thus,

$$\Delta\bar{L}\Omega\bar{L}^T\boldsymbol{x} = \boldsymbol{x}\lambda. \tag{6.15}$$

There is a related eigenproblem dual to (6.15):

$$\Omega\bar{L}^T\Delta\bar{L}\boldsymbol{z} = \boldsymbol{z}\lambda \tag{6.16}$$

obtained by taking a LU transform of (6.15). This gives, for us, the most elegant (symmetric) form of our problem as

$$\bar{L}\Omega\bar{L}^T\boldsymbol{x} = \Delta\boldsymbol{x}\lambda, \quad \bar{L}^T\Delta\bar{L}\boldsymbol{z} = \Omega\boldsymbol{z}\lambda,$$

with just a single bidiagonal matrix $\bar{L}$.

Now we may keep Δ and Ω fixed and ask how sensitive λ is to changes in $\bar{L}$. We follow an analysis similar to those in Theorems 6.1, 6.3 and 6.4, but here based on the partial derivatives of λ with respect to the entries $(1,1),\dots,(n,n)$ and $(2,1),\dots,(n,n-1)$ of $\bar{L}$. This allows us to find the *relative condition number* for $\Delta T = \Delta\bar{L}\Omega\bar{L}^T$ in Theorem 6.5.

THEOREM 6.5. *Let B be a balanced unreduced real tridiagonal matrix that permits a triangular factorization $B = \Delta\bar{L}\Omega\bar{L}^T$, where Δ and Ω are diagonal signature matrices and $\bar{L}$ is a lower bidiagonal matrix. Let $\lambda \neq 0$ be a simple eigenvalue of B with right eigenvector $\boldsymbol{x}$. Then $relcond(\lambda; \bar{L}) := \|\, relgrad_{\bar{L}}(\lambda)\|_1$ is equal to*

$$\begin{aligned} relcond(\lambda; \bar{L}) &= \lim_{\eta \to 0} \sup \left\{ \frac{|\delta\lambda|}{\eta|\lambda|} : (\lambda + \delta\lambda) \text{ is an eigenvalue of } \Delta(\bar{L} + \delta\bar{L})\Omega(\bar{L} + \delta\bar{L})^T,\ |\delta\bar{L}| \leq \eta|\bar{L}| \right\} \\ &= \frac{2\,|\boldsymbol{x}^T|\,|\bar{L}|\ |\bar{L}^T\boldsymbol{x}|}{|\lambda|\,|\boldsymbol{x}^T\Delta\boldsymbol{x}|}. \end{aligned}$$

Observe that neither the 1's nor the 0's of $\bar{L}$ are changed in the perturbations.

For computational purposes, it is more convenient to express *relcond*$(\lambda;\bar{L})$ as follows:

$$relcond(\lambda;\bar{L}) := \frac{2|\boldsymbol{x}^T||\bar{L}||\boldsymbol{w}|}{|\boldsymbol{x}^T\Delta\boldsymbol{x}|}, \tag{6.17}$$

where

$$\bar{L}\boldsymbol{w} = \Delta\boldsymbol{x} \text{ for } \boldsymbol{w}. \tag{6.18}$$

The *relative condition number* for $\Omega\bar{L}^T\Delta\bar{L}$ can be obtained in the same manner and is given by

$$relcond(\lambda;\bar{L}^T) := \frac{2|\boldsymbol{z}^T||\bar{L}^T||\boldsymbol{u}|}{|\boldsymbol{z}^T\Omega\boldsymbol{z}|}, \tag{6.19}$$

where

$$\bar{L}^T\boldsymbol{u} = \Omega\boldsymbol{z} \text{ for } \boldsymbol{u}. \tag{6.20}$$

Note that $\boldsymbol{x} = \Delta\bar{L}\boldsymbol{z}$, $\boldsymbol{z}\lambda = \Omega\bar{L}^T\boldsymbol{x}$ and it may be verified that $|\boldsymbol{x}^T\Delta\boldsymbol{x}| = |\boldsymbol{z}^T\Omega\boldsymbol{z}\lambda|$, if $\lambda \neq 0$ is simple. Thus the relative sensitivity of λ in (6.15) and (6.16) is the same,

$$\boxed{relcond(\lambda;\bar{L}) = \frac{2|\boldsymbol{x}^T||\bar{L}||\boldsymbol{z}|}{|\boldsymbol{x}^T\Delta\boldsymbol{x}|}.} \tag{6.21}$$

Note the bidiagonal linear systems (6.18) and (6.20) are consistent when $\bar{L}$ is singular.

Once again the right hand side of (6.21) is a nonzero finite number for a simple eigenvalue $\lambda = 0$ but we set *relcond*$(0;\bar{L}) = 0$, since the zero eigenvalue or, equivalently, the singularity of the matrix B, is preserved by the perturbations that we have considered.

6.5. Harmless element growth. The following example shows that element growth in the factored form can be harmless. For ϵ close to the roundoff unit, the matrix

$$T = \begin{pmatrix} \epsilon & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & \epsilon \end{pmatrix}$$

has eigenvalues $\lambda_1 = \epsilon$, $\lambda_2 \simeq -1$ and $\lambda_3 \simeq 2$. An eigenvector for λ_1 is $\boldsymbol{x} = \begin{pmatrix} 1 & 0 & -1 \end{pmatrix}^T$. Omitting $\mathcal{O}(\epsilon^2)$ terms, the factored form for $T = LDL^T$ is given by

$$T = \begin{pmatrix} 1 & & \\ \epsilon^{-1} & 1 & \\ & -\epsilon & 1 \end{pmatrix} \begin{pmatrix} \epsilon & & \\ & 1-\epsilon^{-1} & \\ & & 2\epsilon \end{pmatrix} \begin{pmatrix} 1 & \epsilon^{-1} & \\ & 1 & -\epsilon \\ & & 1 \end{pmatrix}.$$

The relative condition number *relcond*$(\epsilon; L, D)$ for the smallest eigenvalue $\lambda_1 = \epsilon$ shows that despite element growth this eigenvalue is robust. In fact, solving $L\boldsymbol{v} = \boldsymbol{x}$ for $\boldsymbol{v}$, we obtain $\boldsymbol{v} = \begin{pmatrix} 1 & -\epsilon^{-1} & -2 \end{pmatrix}^T$, $\boldsymbol{x}^T L = \begin{pmatrix} 1 & \epsilon & -1 \end{pmatrix}$, $\boldsymbol{x}^T \mathring{L} = \begin{pmatrix} 0 & \epsilon & 0 \end{pmatrix}$ and

$$relcond(\epsilon; L, D) = \frac{\left(|\boldsymbol{x}^T L| + 2|\boldsymbol{x}^T\mathring{L}|\right)|\boldsymbol{v}|}{\|\boldsymbol{x}\|^2} = \frac{1}{2}\begin{pmatrix} 1 & 3\epsilon & 1 \end{pmatrix}\begin{pmatrix} 1 \\ \epsilon^{-1} \\ 2 \end{pmatrix} = 3.$$

Also, we have the representation $T = \bar{L}\Omega\bar{L}^T$ with

$$\bar{L} = \begin{pmatrix} \epsilon^{1/2} & & \\ \epsilon^{-1/2} & \left|1-\epsilon^{-1}\right|^{1/2} & \\ & -\epsilon\left|1-\epsilon^{-1}\right|^{1/2} & (2\epsilon)^{1/2} \end{pmatrix} \quad \text{and} \quad \Omega = \operatorname{diag}(1,-1,1).$$

Solving $\bar{L}\boldsymbol{v_1} = \boldsymbol{x}$ for $\boldsymbol{v_1}$, we obtain $\boldsymbol{v_1} = \begin{pmatrix} \epsilon^{-1/2} & -\epsilon^{-1}\left|1-\epsilon^{-1}\right|^{-1/2} & -2\big(2\epsilon\big)^{-1/2} \end{pmatrix}^T$, $2|\boldsymbol{x}^T||\bar{L}| = 2\left(\epsilon^{1/2}, \epsilon\left|1-\epsilon^{-1}\right|^{1/2}, (2\epsilon)^{1/2}\right)$ and

$$\mathit{relcond}(\epsilon;\bar{L}) = \frac{2|\boldsymbol{x}^T||\bar{L}||\boldsymbol{v_1}|}{\|\boldsymbol{x}\|^2} = \left(\epsilon^{1/2}, \epsilon\left|1-\epsilon^{-1}\right|^{1/2}, (2\epsilon)^{1/2}\right)\begin{pmatrix} \epsilon^{-1/2} \\ \epsilon^{-1}\left|1-\epsilon^{-1}\right|^{-1/2} \\ 2(2\epsilon)^{-1/2} \end{pmatrix} = 4.$$

In contrast, $\mathit{relcond}(\lambda_2; L, D)$ and $\mathit{relcond}(\lambda_3; L, D)$ are $\mathcal{O}(1/\epsilon)$. The factored form of the given T is bad for computing $(\lambda_2, \boldsymbol{x}_2)$ and $(\lambda_3, \boldsymbol{x}_3)$ but excellent for computing $(\lambda_1, \boldsymbol{x}_1)$. When T is shifted close to λ_2 or λ_3 there is no element growth and the eigenvectors may be computed accurately.

We can destroy symmetry via a diagonal similarity transform $T \to STS^{-1}$. The eigenvectors change in a simple way but the *relcond*'s remain the same.

6.6. Equivalence of $\boldsymbol{\mathit{relcond}}(\lambda;\mathcal{L},\mathcal{U})$, $\boldsymbol{\mathit{relcond}}(\lambda;L,D)$, $\boldsymbol{\mathit{relcond}}(\lambda;\bar{L})$. There is no need to compare numerically the three representations $J = \mathcal{L}\mathcal{U}$, $\Delta T = \Delta LDL^T$ and $\Delta T = \Delta\bar{L}\Omega\bar{L}^T$ since the relative condition numbers are equivalent as shown in the next result.

Lemma 6.6.

$$\mathit{relcond}(\lambda;\mathcal{L},\mathcal{U}) \le \mathit{relcond}(\lambda;L,D) \le 3\,\mathit{relcond}(\lambda;\mathcal{L},\mathcal{U})$$

and

$$\frac{1}{2}\,\mathit{relcond}(\lambda;\bar{L}) \le \mathit{relcond}(\lambda;L,D) \le \frac{3}{2}\,\mathit{relcond}(\lambda;\bar{L}).$$

Proof. For S invertible and diagonal, let $J = SBS^{-1} = S\Delta TS^{-1}$ and consider the triangular factorizations $J = \mathcal{L}\mathcal{U}$ and $B = \Delta T = \Delta LDL^T$. Then we have

$$\mathcal{L}\mathcal{U} = S\Delta LDL^TS^{-1}.$$

Uniqueness of LU factorization guarantees that

$$\mathcal{L} = S\Delta LS^{-1}\Delta \quad \text{and} \quad \mathcal{U} = \Delta SDL^TS^{-1}$$

and

$$\mathring{\mathcal{L}} = S\Delta\mathring{L}S^{-1}\Delta \quad \text{and} \quad \operatorname{diag}(\mathfrak{u}) = \Delta D.$$

The eigenvectors, in the usual notation, satisfy

$$\boldsymbol{x}_B = S^{-1}\boldsymbol{x}_J \quad \text{and} \quad \boldsymbol{x}_B^T\Delta = \boldsymbol{y}_B^* = \boldsymbol{y}_J^*S.$$

Consequently, $|\boldsymbol{y}_J^*\boldsymbol{x}_J| = |\boldsymbol{y}_B^*\boldsymbol{x}_B| = |\boldsymbol{x}_B^T\Delta\boldsymbol{x}_B|$ and, from (6.8),

$$|\lambda||\boldsymbol{y}_J^*\boldsymbol{x}_J|\,\mathit{relcond}(\lambda;\mathcal{L},\mathcal{U}) = |\boldsymbol{y}_J^*\mathcal{L}\operatorname{diag}(\mathbf{u})||\boldsymbol{x}_J| + |\boldsymbol{y}_J|^T|\mathring{\mathcal{L}}\mathcal{U}\boldsymbol{x}_J|,$$

we obtain

$$|\lambda||\boldsymbol{x}_B^T\Delta\boldsymbol{x}_B|\,relcond(\lambda;\mathcal{L},\mathcal{U}) = \left(|\boldsymbol{x}_B|^T + |\boldsymbol{x}_B^T\mathring{L}|\right)|DL^T\boldsymbol{x}_B|.$$

Compared with (6.12) in the derivation of $relcond(\lambda; L, D)$,

$$|\lambda||\boldsymbol{x}_B^T\Delta\boldsymbol{x}_B|\,relcond(\lambda;L,D) = \left(|\boldsymbol{x}_B^T L| + 2|\boldsymbol{x}_B^T\mathring{L}|\right)|DL^T\boldsymbol{x}_B|,$$

to conclude that, since

$$|\boldsymbol{x}_B^T L| = |\boldsymbol{x}_B^T(I+\mathring{L})| \leq |\boldsymbol{x}_B|^T + |\boldsymbol{x}_B^T\mathring{L}|, \tag{6.22}$$

we have

$$|\boldsymbol{x}_B^T L| + 2|\boldsymbol{x}_B^T\mathring{L}| \leq 3\left(|\boldsymbol{x}_B|^T + |\boldsymbol{x}_B^T\mathring{L}|\right)$$

and

$$relcond(\lambda;L,D) \leq 3\,relcond(\lambda;\mathcal{L},\mathcal{U}).$$

Also, since

$$|\boldsymbol{x}_B|^T = |\boldsymbol{x}_B^T(L-\mathring{L})| \leq |\boldsymbol{x}_B^T L| + |\boldsymbol{x}_B^T\mathring{L}|, \tag{6.23}$$

$$relcond(\lambda;\mathcal{L},\mathcal{U}) \leq relcond(\lambda;L,D).$$

The derivation of the second part of the lemma is similar and is omitted. □

7. Preparation for case studies.

7.1. Sampling the perturbations. Let ν denote the number of real parameters in the representation of a matrix ($\nu = 3n-2$ for C, $\nu = 2n-1$ for $\mathcal{LU}$ and ΔLDL^T). For the perturbation level η we change each parameter p in turn to $p(1\pm\eta)$ and calculate the new spectrum. This will give a sample of 2ν perturbations for each λ and one can find the greatest change. The size depends on λ and on η. We define an *empirical sensitivity* function

$$emp(\lambda,\eta) := \frac{\max|\delta\lambda|}{\eta|\lambda|}$$

where $\max|\delta\lambda|$ is the max over our 2ν samples.

We want to know the values of η for which $|\delta\lambda/\lambda|$ is proportional to η, i.e., the perturbations are small enough to make the *h.o.t.* negligible, so we define η^* by

$$\eta^* = \max\{\eta : emp(\lambda,\eta) \text{ has at least one decimal digit in common with } emp(\lambda,\tau), \tau<\eta, \\ \text{until roundoff in computing the perturbed eigenvalues interferes}\}.$$

Finally we define our *relative empirical sensitivity measure* by

$$emp(\lambda) := emp(\lambda,\eta^*) \tag{7.1}$$

and record the perturbation level η^* for each λ, $\eta^*(\lambda)$ in our studies.

7.2. Finding $|\delta\lambda|$. To determine $|\delta\lambda|$ is not as straightforward as it seems at first glance. A big challenge lies in pairing up correctly the new spectrum with the original one. If one new eigenvalue is incorrectly paired then our computed $|\delta\lambda|$ will not give us correct information. What we do is to order the eigenvalues of the original matrix by increasing $relcond(\lambda)$ (i.e. take the easiest cases first). We find the closest new λ to the first old λ and remove each from its list. Then we find the closest remaining new λ to the second old λ and remove each from its list and so on until each list is empty. Easier ordering, by real part or magnitude, do not resolve the difficulties. This can be seen, by eye, in Figure 2.1.

We have to be concerned about perturbations to a complex conjugate pair of eigenvalues giving rise to two real eigenvalues (and the opposite). See Figure 2.1 to see the challenge of pairing even when eigenvalues are distinct. What complex "+" should be paired with the real "·" furthest from the origin? In addition, when the perturbed eigenvalue regions for two distinct eigenvalues overlap then we have to be concerned about some perturbations giving rise to a multiple eigenvalue. At this stage of our studies we avoid such difficult cases.

7.3. Finding η^*. Another difficulty is the size of η^*. Our decision to use Matlab and compute with *roundoff unit* $\varepsilon \simeq 2 \times 10^{-16}$ confined our η values to the range $10^{-1}, 10^{-2}, \ldots, 10^{-14}$ and we try them all. We can not know $\eta^*(\lambda)$ in advance. When examining the other condition numbers that we compute we ignore any values corresponding to $\eta > \eta^*(\lambda)$. Clearly if $\eta^*(\lambda) = 10^{-14}$ then λ is too sensitive to permit a study using Matlab. In the same line of thinking, $\log_{10} relcond(\lambda)$ tells us roughly how many decimal digits of λ are "lost" through finite precision computation. If $relcond(\lambda) = 10^8$ then the best we can hope for from an algorithm executed in Matlab is that the leading $16 - 8 = 8$ will be correct. We are not aware of any previous studies that indicate when perturbations are small enough to neglect higher order terms.

7.4. Checking the eigenvectors. All condition numbers use the eigenvectors and we need both column and row eigenvectors:

$$C\boldsymbol{x} = \boldsymbol{x}\lambda, \qquad \boldsymbol{y}^* C = \lambda \boldsymbol{y}^*.$$

We do not normalize our eigenvectors. Figure 2.1 shows that Matlab often produces a different spectrum for C^T and C. Hence we are reluctant to accept Matlab's eigenvectors for C^T as the partners of the eigenvectors for C. In exact arithmetic the two spectra are the same but Matlab, and any other method based on similarities, returns the eigenvalues of $C + E$ where $\|E\| = \mathcal{O}(\varepsilon\|C\|)$ and the matrix E for C^T differs from the E for C, in general. How do we make sure that the computed right and left eigenvectors are correctly paired?

Instead we have used a code developed by J. Slemons as part of his Ph.D thesis at the University of Washington in Seattle [17]. This code exploits tridiagonal form and automatically produces the same approximations for C and C^T. It accepts an (accurate) eigenvalue approximation and then performs a *Rayleigh quotient iteration* using both column and row eigenvectors. It returns an eigenpair $(\boldsymbol{x}, \boldsymbol{y}^*)$ together with an improved eigenvalue estimate of $\boldsymbol{y}^* C\boldsymbol{x}/\boldsymbol{y}^*\boldsymbol{x}$, the generalized Rayleigh quotient. We compute column and row residual norms relative to the eigenvalue,

$$\frac{\|C\boldsymbol{x} - \boldsymbol{x}\lambda\|}{|\lambda|\|\boldsymbol{x}\|} \quad \text{and} \quad \frac{\|\boldsymbol{y}^* C - \lambda\boldsymbol{y}^*\|}{|\lambda|\|\boldsymbol{y}^*\|}, \tag{7.2}$$

and record them both. This is a much stricter measure than the usual $\dfrac{\|C\boldsymbol{x} - \boldsymbol{x}\lambda\|}{\|\boldsymbol{x}\|\|C\|}$. If either row or column residual norms rises above 10^{-6} we abort this numerical example.

We assemble all the eigenvectors, correctly paired, and compute

$$G = (Row\ eigenvectors) \cdot (Column\ eigenvectors) = Y^*X$$

We checked the diagonality of G by displaying

$$\max_{i \neq j} \left(\frac{|g_{ij}|^2}{|g_{ii}||g_{jj}|} \right).$$

If this exceeds $\sqrt{\varepsilon} \simeq 10^{-8}$ then we abort this numerical example.

7.5. Element growth. Representations $J = \mathcal{L}\mathcal{U}$ and $T = LDL^T$ can break down and, more commonly, can produce factors far greater than J and T. We compare the maximum entry in the factors with the maximum entry in J or T. If this exceeds $1/\varepsilon \simeq 10^{16}$ we abort this case.

It must be stressed that in the symmetric case large element growth does not necessarily prevent small eigenvalues from being relatively robust (See example in Section 6.5). In these cases it is the large eigenvalues that may be more sensitive to the parameters in $\mathcal{L}$ and $\mathcal{U}$ than to the entries in J. In such cases it is better to shift ($J \longrightarrow J - \mu I$) to a point close to the large eigenvalue and then factor once again: $J \longrightarrow J - \mu I = \widetilde{\mathcal{L}}\widetilde{\mathcal{U}}$.

For these reasons we allow a growth tolerance as large as $1/\varepsilon$. Indeed, one reason for our studies is to see if features of eigenvalue computation extend from the symmetric case to the unsymmetric case. They do, in our examples.

8. Numerical studies.

8.1. Questions. Below are the questions that shaped our study.

- Representation 1 (C given by $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$)
  How much does the relative Wilkinson condition number $BGT(\lambda)$ overestimate the sensitivity of an eigenvalue?
- Representation 2 ($J = \mathcal{L}\mathcal{U}$)
  Are the very small eigenvalues less sensitive to relative changes in $\mathcal{L}$, $\mathcal{U}$ than to the same relative changes in the entries of J?
  Does element growth in $\mathcal{L}$ and $\mathcal{U}$ affect the sensitivity with respect to $\mathcal{L}$ and $\mathcal{U}$?
- How can one find out the perturbation level η^* above which the second order effects dominate the first order (linear) effects? Before we began this study we never asked ourselves this question. See section 7.3.

8.2. Study 1 - Graded matrix. This matrix C was created in ΔT form with

$$\begin{aligned} T &= \operatorname{diag}(a_1, \ldots, a_{n-1}, a_n) + \operatorname{diag}_-(b_1, \ldots, b_{n-1}) + \operatorname{diag}_+(c_1, \ldots, c_{n-1}), \\ &\quad a_j = b_j = c_j = 3^{-(j-1)}, \quad j = 1, \ldots, n-1, \\ &\quad a_n = 3^{-(n-1)}, \\ \text{and } \Delta &= \operatorname{diag}(\boldsymbol{\delta}), \delta_j = (-1)^{\lfloor (j+1)/2 \rfloor}, \quad j = 1, \ldots, n. \end{aligned}$$

The result is a balanced matrix with eigenvalues of different magnitude.

Our numerical study has $n = 20$. The eigenvalues range from 10^{-9} to 1.7 in magnitude. There are 6 real eigenvalues. The maximum of the strict relative residual norms, see (7.2), was 10^{-12} and the closeness of the computed Y^*X to diagonal form was 10^{-29}.

Our purpose is to show how misleading the BGT condition number can be. All the eigenvalues are determined to high relative accuracy using both representations C and $J = \mathcal{L}\mathcal{U}$. The smallest value of η^* is 10^{-2} and this indicates the robust nature of the eigenvalues.

Remarks.

1. For the representation C (matrix entries) and for the eigenvalues with minimum and maximum absolute values, λ_{min} and λ_{max}, respectively,

$$\begin{array}{lll} BGT(\lambda_{min}; C) & = 8.4 \times 10^8, & relcond(\lambda_{min}; C) = 14.5 \\ BGT(\lambda_{max}; C) & = 1.05, & relcond(\lambda_{max}; C) = 1.05. \end{array}$$

   For all the other eigenvalues,

$$6.0 \leq BGT(\lambda; C) \leq 7.0 \times 10^8,$$
$$3.8 \leq relcond(\lambda; C) \leq 61.6.$$

2. When C is put into J form the condition number of the similarity is 4×10^{81}. The $\mathcal{L}\mathcal{U}$ factorization incurred no element growth and all $relcond(\lambda; \mathcal{L}, \mathcal{U})$ values dropped below 6.5,

$$1.3 \leq relcond(\lambda; \mathcal{L}, \mathcal{U}) \leq 6.5.$$

3. For both the representations

$$relcond(\lambda) \leq 40\, emp(\lambda),$$

   but in most cases the factor was 2.

Results for $n = 50, 100$ are the same for $C = \Delta T$ but the condition number of the similarity transform to J quickly escalates to overflow and spoils the eigenvector computation thus preventing the calculation of $relcond(\lambda; \mathcal{L}, \mathcal{U})$. However, the computation of $relcond(\lambda; L, D)$ went through without difficulty and the values are small, which, together with Lemma 6.6, guarantees that also the values of $relcond(\lambda; \mathcal{L}, \mathcal{U})$ are small.

8.3. Study 2 - Generalized Bessel matrix. Bessel matrices, associated with generalized Bessel polynomials, are nonsymmetric tridiagonal matrices defined by

$$B_n^{(a,b)} = \mathrm{diag}(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) + \mathrm{diag}_-(\beta_1, \ldots, \beta_{n-1}) + \mathrm{diag}_+(\gamma_1, \ldots, \gamma_{n-1})$$

with

$$\begin{aligned} \alpha_1 &= -\frac{b}{a}, \quad \alpha_j := -b\frac{a-2}{(2j+a-2)(2j+a-4)}, \quad j = 2, \ldots, n, \\ \beta_1 &= \frac{\alpha_1}{a+1}, \quad \beta_j := -b\frac{j}{(2j+a-1)(2j+a-2)}, \quad j = 2, \ldots, n-1, \\ \gamma_1 &= -\alpha_1, \quad \gamma_j := b\frac{j+a-2}{(2j+a-2)(2j+a-3)}, \quad j = 2, \ldots, n-1. \end{aligned}$$

Parameter b is a scaling factor and most authors take $b = 2$ and so do we. The case $a \in \mathbb{R}$ is the most investigated in literature. The eigenvalues of these matrices $B_n^{(a,b)}$ suffer from ill-conditioning that increases with n - close to a defective matrix. In Pasquini [15] it is mentioned that the ill-conditining seems to reach its maximum when a ranges from -8.5 to -4.5.

Our example takes $a = -4.5$, $b = 2$ and $n = 10$, $C = B_{10}^{(-4.5,2)}$. This matrix has well separated complex eigenvalues, far from small (for all λ, $0.19 \leq |\lambda| \leq 0.27$), that are sensitive to small relative changes in the matrix entries. All our representations reflect this sensitivity - but in milder form. The maximum of the eigenvector relative residuals was 10^{-13} and the departure of Y^*X from diagonality 10^{-23}.

In [1] the authors take $n = 18$. Here is where our empirical sensitivity measure is useful. With $n = 18$ the most sensitive eigenvalues (those furthest from the imaginary axis) have $\eta^* \leq 10^{-14}$ and we cannot assume that Matlab function `eig` will yield any correct digits. With $n = 10$, $\eta^* = 10^{-10}$ for the most sensitive eigenvalue and we get at least one correct digit in all the eigenvalues we compute in studying sensitivity. In other words, in addition to $relcond(\lambda)$ it is nice to know for what values of η the relative change $|\delta\lambda/\lambda|$ is majorized by $\eta\, relcond(\lambda)$.

We were disappointed, at first, that the factored representation $\mathcal{LU}$ gave smaller *relcond*'s only by factors between 10 and 100. Fortunately, η^* rises to the range $[10^{-9}, 10^{-6}]$ from $[10^{-10}, 10^{-7}]$ for C, the matrix entries. Relative changes of level 10^{-5} are too large to enjoy a linear response in the eigenvalues.

The BGT is more realistic (since there are no small eigenvalues). For all λ,

$$7.7 \times 10^6 \leq BGT(\lambda; C) \leq 1.8 \times 10^9,$$

as against

$$\begin{gathered} 2.1 \times 10^6 \leq relcond(\lambda; C) \leq 7.0 \times 10^8, \\ 9.8 \times 10^4 \leq relcond(\lambda; \mathcal{L}, \mathcal{U}) \leq 3.8 \times 10^7. \end{gathered}$$

Also, for both representations,

$$relcond(\lambda) \leq 5\, emp(\lambda).$$

When we shift by real parts of λ's none of the condition numbers change dramatically but the factored form is about 100 times more robust.

Now we turn to the usefulness of the factored form. When we shift by σ chosen very close to any eigenvalue λ then $BGT(\lambda - \sigma; C - \sigma I)$ and $relcond(\lambda - \sigma; C - \sigma I)$ increase greatly, as expected. However, with $\widetilde{\mathcal{L}}\widetilde{\mathcal{U}} = C - \sigma I$, $relcond(\lambda; \widetilde{\mathcal{L}}, \widetilde{\mathcal{U}})$ remains approximately the same as before shifting. For example, for $\lambda = 3.20 \times 10^{-3} \pm 1.85 \times 10^{-1} i$ we obtain

$$BGT(\lambda, C) = 7.7 \times 10^6, \quad relcond(\lambda; C) = 2.1 \times 10^6, \quad relcond(\lambda; \mathcal{L}, \mathcal{U}) = 9.8 \times 10^4,$$

and with $\sigma = \lambda(1 + 10^3 * \varepsilon)$ we find

$$BGT(\lambda - \sigma; C - \sigma I) = 3.5 \times 10^{19}, \quad relcond(\lambda - \sigma; C - \sigma I) = 9.3 \times 10^{18}, \quad relcond(\lambda - \sigma; \widetilde{\mathcal{L}}, \widetilde{\mathcal{U}}) = 5.8 \times 10^4.$$

So here is an example that the known benefits of factoring in the symmetric case can extend to the unsymmetric case. This is one of the questions which gave rise to our study.

8.4. Study 3 - Matrix with clusters. At the end of section 7.2 we said that we would avoid examples in which the eigenvalues were too close to each other, to admit unique paring. Yet matrix Test 5 in [1] was designed to have large, tight clusters,

$$\begin{aligned} &C = D^{-1}\left[\mathrm{diag}(\alpha_1, \ldots, \alpha_{n-1}, \alpha_n) + \mathrm{diag}_-(1, \ldots, 1) + \mathrm{diag}_+(1, \ldots, 1)\right], \\ &D = \mathrm{diag}(\beta_1, \ldots, \beta_n), \\ &\alpha_k = 10^{5(-1)^k} \cdot (-1)^{\lfloor k/4 \rfloor}, \quad \beta_k = (-1)^{\lfloor k/3 \rfloor}, \quad k = 1, \ldots, n \end{aligned}$$

and the excuse for discussing it is our surprise at finding that all the *relcond*'s with $n = 100$ are small (two are 586, the rest are less than 12) despite BGT values up to 10^{10} for the eigenvalues near 0. How is this possible?

The eigenvectors supplied the explanation. The matrix has a repetitive structure and the diagonal entries are a good guide to the eigenvalues. For the large real eigenvalues near $\pm 10^5$ the eigenvectors have spikes $(-10^{-5}, -1, 10^{-5})$ (complex conjugate pairs have spikes $(10^{-5}, 1, -10^{-7}, -1, 10^{-5})$), at the appropriate places, and negligible elsewhere. Hence the numerical supports for many eigenvectors are disjoint. The essential structure of the matrix is exhibited with $n = 10$ and we show details for this case in table 8.1. The matrix is

$$C = \begin{pmatrix} 10^{-5} & 1 & & & & & & & & \\ 1 & 10^5 & 1 & & & & & & & \\ & -1 & -10^{-5} & -1 & & & & & & \\ & & -1 & 10^5 & -1 & & & & & \\ & & & -1 & 10^{-5} & -1 & & & & \\ & & & & 1 & -10^5 & 1 & & & \\ & & & & & 1 & -10^{-5} & 1 & & \\ & & & & & & 1 & 10^5 & 1 & \\ & & & & & & & -1 & -10^{-5} & -1 \\ & & & & & & & & -1 & -10^5 \end{pmatrix}$$

and it has 5 eigenvalues near 0, 3 eigenvalues near 10^5 and 2 near -10^5. The diagonal entries are a good guide to the eigenvalues. This accounts for values 1.00 in the table.

λ	η^*	$emp(\lambda)$	$BGT(\lambda, C)$	$relcond(\lambda, C)$
5.76×10^{-6}	10^{-1}	2.12	2.5×10^{10}	12.9
$-1.83 \times 10^{-6} + 1.03 \times 10^{-5}i$	10^{-1}	0.76	1.2×10^{10}	7.5
$-1.83 \times 10^{-6} - 1.03 \times 10^{-5}i$	10^{-1}	0.76	1.2×10^{10}	7.5
$-1.11 \times 10^{-5} + 5.95 \times 10^{-6}i$	10^{-2}	1.58	1.8×10^{10}	14.1
$-1.11 \times 10^{-5} - 5.95 \times 10^{-6}i$	10^{-2}	1.00	1.8×10^{10}	14.1
1.00×10^5	10^{-1}	1.00	1.00	1.00
-1.00×10^5	10^{-1}	1.00	1.00	1.00
1.00×10^5	10^{-1}	1.00	685	685
-1.00×10^5	10^{-1}	1.00	1.00	1.00
1.00×10^5	10^{-1}	1.00	685	685

Table 8.1
Study 3 - details for size $n = 10$

8.5. Study 4 - Modified shifted Wilkinson matrix. This example illustrates again the fact that representations $\mathcal{LU}$ and ΔLDL^T can determine a very tiny eigenvalue to high relative accuracy while the matrix entries deliver large *relcond*'s. Our shift σ is the computed $\lambda_{min}(W_{21}^+) \simeq -1.25$ where W_{21}^+ is the well-known Wilkinson matrix,

$$W_{21}^+ = \mathrm{diag}(10, 9, \ldots, 2, 1, 0, 1, 2, \ldots, 9, 10) + \mathrm{diag}_{-1}(1, 1, \ldots 1, 1) + \mathrm{diag}_{+1}(1, 1, \ldots 1, 1),$$

and our matrix is

$$C = \Delta(W_{21}^{+} - \lambda_{min})$$

where $\Delta = \mathrm{diag}(\boldsymbol{\delta})$, $\delta_j = (-1)^{\lfloor (j+1)/2 \rfloor}$, $j = 1, \ldots, 21$.

The tiny eigenvalue is $\lambda = \lambda(C) = 2.55 \times 10^{-15}$, $\mathcal{LU} = J$ and $\Delta LDL^T = B$. The pleasing result is

$$relcond(\lambda; C) = 2.3 \times 10^{15} \quad \text{and} \quad relcond(\lambda; \mathcal{L}, \mathcal{U}) = relcond(\lambda; L, D) = 20.7.$$

For all the other eigenvalues, all *relcond*'s for these representations are less than 6.5.

9. Conclusion. Many unsymmetric tridiagonal matrices define all their eigenvalues to high relative accuracy. This is surprising. The only counter-example in our test bed are the generalized Bessel matrices. Even a matrix from Bini, Gemignani and Tisseur in [1] with 3 extremely tight clusters had this nice property.

We reported on a procedure for measuring η^*, the relative perturbation level above which higher order terms interfere with standard perturbation theory.

The main goal was to show that, after shifting close to an eigenvalue, the factored form can define the tiny eigenvalues very well and thus allow the possibility of accurate computation of the eigenvectors. It just happened that in none of our examples was there excessive element growth in factoring these nearly singular unsymmetric matrices. In addition, the example discussed in Section 6.5 indicates that element growth may be harmless for the sensitivity of tiny eigenvalues in most cases. The usefulness of our new relative condition numbers with respect to factored forms is evident in our examples.

REFERENCES

[1] D. A. Bini, L. Gemignani and F. Tisseur. *The Ehrlich-Aberth method for the nonsymmetric tridiagonal eigenvalue problem*. SIAM J. Matrix Analysis and its Aplications, 27(1):153-175, 2005.

[2] P. A. Clement. *A class of triple-diagonal matrices for test purposes*. SIAM Review, 1:50-52, 1959.

[3] C. Ferreira and B. Parlett. *Convergence of LR algorithm for a one-point spectrum tridiagonal matrix*. Numerische Mathematic, 113(3):417-431, 2009.

[4] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore and London, 3rd ed., 1996.

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

[6] D. J. Higham and N. J. Higham. *Structured backward error and condition of generalized eigenvalue problems*. SIAM J. Matrix Analysis and its Aplications, 20:493-512, 1998.

[7] I. C. F. Ipsen. *Relative perturbation results for matrix eigenvalues and singular values*. Acta Numerica, 151-201, 1998.

[8] M. Karow, D. Kressner and F. Tisseur. *Structured eigenvalue condition numbers*. SIAM J. Matrix Analysis and its Aplications, 28(4):1052-1068, 2006.

[9] H. W. J. Lenferink and M.N. Spijker. *On the use of stability regions in the numerical analysis of initial value problems*. Mathematics of Computation, 57(195): 221-237, 1991.

[10] R.-C. Li. *Relative perturbation theory. III. More bounds on eigenvalue variations*. Linear Algebra and Its Aplications, 266:337-345, 1997.

[11] Z. S. Liu. *On the extended HR algorithm*. Technical Report PAM-564, Center for Pure and Applied Mathematics, University of California, Berkeley, CA, USA, 1992.

[12] S. Noschese and L. Pasquini. *Eigenvalue condition numbers: zero-structured versus traditional*. Journal of Computational and Applied Mathematics, 185: 174-189, 2006.

[13] B. N. Parlett. *Spectral sensitivity of products of bidiagonals*. Linear Algebra and Its Aplications, 275-276:417-431, 1998.

[14] B. N. Parlett and C. Reinsch. *Balancing a matrix for calculation of Eigenvalues and Eigenvectors*. Numerische Mathematik, 13:292-304, 1969.

[15] L. Pasquini. *Accurate computation of the zeros of the generalized Bessel polynomials*. Numerische Mathematik, 86:507-538, 2000.

[16] J. Rice. *A theory of condition*. SIAM Journal on Numerical Analysis, 3(2):287-310, 1966.

[17] J. Slemons. *Toward the solution of the eigenproblem: nonsymmetric tridiagonal matrices*. Ph.D thesis, University of Washington, Seattle, 2008.

[18] G. W. Stewart and Ji-guang Sun. *Matrix Perturbation Theory*. Academic Press, INC., 1990.

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