IF A MATRIX HAS ONLY A SINGLE EIGENVALUE HOW SENSITIVE IS THIS EIGENVALUE?

GRACE E. CHO* AND ILSE C.F. IPSEN†

Abstract. For matrices with a single eigenvalue we analyse the sensitivity of the eigenvalue to perturbations in the matrix. We derive a closed form result that is similar in spirit to an analytical result by Lidskii; improve a bound by Henrici; and express the resolvent norm in terms of the nonnormal part N of a Schur decomposition. In particular we show that the resolvent norm can increase as a power of $\|N\|$, independent of the eigenvalue error.

Key words. eigenvalue, defectiveness, nonnormality, Jordan decomposition, Schur decomposition, resolvent

AMS subject classification. 15A18, 15A42, 15A60, 65F15, 65F35

1. Introduction.

We study the eigenvalue sensitivity of a complex square matrix A that has only a single eigenvalue λ.

Why do we restrict ourselves to this simple class of matrices? First, we want to exclude the possibility of coupling among different eigenvalues [3, §10]. We can only deal with coupling once we understand how an eigenvalue behaves on its own, without interference from other eigenvalues. Second, we wanted to find a closed form perturbation bound that has the same spirit as an analytical perturbation result by Lidskii [16], which is presented for matrices with a single eigenvalue. Lidskii's result was brought to public attention by Moro, Burke and Overton [19].

A matrix with a single eigenvalue has a Jordan decomposition $A = XJX^{-1}$, where all diagonal elements of J are equal to λ. Two factors are usually blamed for the sensitivity of λ to perturbations in A, the defectiveness of λ [6, p 259] and the size of the *Jordan condition number* $\|X\|\,\|X^{-1}\|$ [17, 22].

1.1. Motivation.

The following well-known example [20, §3.1] illustrates that defectiveness can make an eigenvalue extremely sensitive to perturbations.

Let J_0 be a Jordan block of order n and $J_0 + E$ be a perturbation,

$$J_0 = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}, \qquad J_0 + E = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \epsilon & & & \lambda \end{pmatrix},$$

where $0 < \epsilon < 1$. The perturbation in any eigenvalue μ of $J_0 + E$ equals

$$|\lambda - \mu| = \epsilon^{1/n}.$$

One can change the situation slightly and replace all the ones on the superdiagonal

*Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (ehcho@eos.ncsu.edu).

†Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (ipsen@math.ncsu.edu). This research was supported in part by NSF grant CCR-9400921.

of J_0 by ν, where $\nu > 0$:

$$A = \begin{pmatrix} \lambda & \nu & & \\ & \ddots & \ddots & \\ & & \ddots & \nu \\ & & & \lambda \end{pmatrix}, \qquad A + E = \begin{pmatrix} \lambda & \nu & & \\ & \ddots & \ddots & \\ & & \ddots & \nu \\ \epsilon & & & \lambda \end{pmatrix}.$$

Now the perturbation in any eigenvalue μ of $A + E$ equals

$$|\lambda - \mu| = (\nu^{n-1}\,\epsilon)^{1/n}.$$

If $\nu \leq \epsilon$ then

$$|\lambda - \mu| \leq \epsilon.$$

This means for small enough superdiagonal elements the eigenvalue perturbation is bounded by the norm of the perturbation matrix. Therefore the eigenvalue λ is insensitive to these particular perturbations although it is maximally defective.

Although A is not a Jordan matrix, it is diagonally similar to one. That is, $A = XJ_0X^{-1}$ where

$$X = \begin{pmatrix} 1 & & & \\ & 1/\nu & & \\ & & \ddots & \\ & & & 1/\nu^{n-1} \end{pmatrix}, \qquad 0 < \nu \leq 1.$$

One could argue, as Wilkinson does [25, §1.8], that the ones on the superdiagonal of a Jordan matrix have no special significance and could be replaced with arbitrary non-zero elements via a diagonal similarity, which is how we got from J_0 to A. However such a similarity can severely affect the sensitivity of the eigenvalues. The eigenvalues of J_0 are highly sensitivity while those of A for $\nu \leq \epsilon$ are insensitive. The Jordan condition number $\|X\|\|X^{-1}\| = 1/\nu^{n-1}$ does not recognise this. As ν becomes smaller the eigenvalue becomes less sensitive to perturbations while the Jordan condition number goes to infinity. Already Wilkinson[1] was aware of this [25, §2.26], and a similar example for diagonalisable matrices appears in [18, §1]. Therefore the Jordan condition number can severely overestimate the eigenvalue sensitivity. The culprit is the discontinuity of the Jordan decomposition.

In effect, the above diagonal similarity transformation amounts to changing the departure of the matrix from normality. The departure of J_0 from normality (in the two-norm) is equal to one, while for A with $\nu \leq \epsilon$ it is ϵ. The much smaller departure from normality of A results in reduced eigenvalue sensitivity. This example motivated us to use a Schur decomposition, instead of a Jordan decomposition, and to estimate the eigenvalue perturbation in terms of non-normality. A similar approach is pursued in [2].

1.2. Overview. In §2 we review the traditional error bounds for eigenvalues based on a Jordan decomposition, as well as Lidskii's result which is a power series expansion of the eigenvalue error. In §3 we derive a closed form bound that is similar in spirit to an analytical result by Lidskii. In §4 we extend this bound, and in §5

[1]We thank Pete Stewart for bringing this to our attention.

we derive from this extension a tighter version of a bound by Henrici (for the special case of matrices with a single eigenvalue). In §6 we show that eigenvalues of weakly non-normal matrices are insensitive to perturbations. In §7 we express the resolvent norm in terms of the non-normal part and show that the resolvent norm can increase as a power of the non-normality. We conclude with a summary in §8.

1.3. Notation. The identity matrix of order n is

$$I = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix}$$

with columns e_i. The norm is the two-norm $\|x\| \equiv \sqrt{x^* x}$, where the superscript * denotes the conjugate transpose.

2. Existing Results. We briefly review existing perturbation results based on a Jordan decomposition.

Let A be a matrix with a single eigenvalue λ, and μ be an eigenvalue of the perturbed matrix $A+E$. Let $A = XJX^{-1}$ be a Jordan decomposition of A where all diagonal elements of J are equal to λ, and s is the index of λ, that is the order of a largest Jordan block in J.

The following bound [15, Theorem 2], [23, Theorem IV.1.12] extends the Bauer-Fike Theorem for diagonalisable matrices [1, Theorem IIIa] to defective matrices,

$$\frac{|\lambda - \mu|^s}{1 + |\lambda - \mu| + \ldots + |\lambda - \mu|^{s-1}} \le \|X^{-1} E X\|. \tag{2.1}$$

The bound suggests that the sensitivity of an eigenvalue increases with the Jordan condition number $\|X\|\,\|X^{-1}\|$ and the size s of a largest Jordan block. Slightly weaker bounds appear in [4, Proposition 1.12.4] and [14, Theorem 8].

Lidskii's analytical perturbation result [16, Theorem 1] is a fine-tuned version of bounds like (2.1). It suggests that, to first order, it is only eigenvectors that determine the sensitivity of an eigenvalue; no generalised eigenvectors of higher grade are involved. We state Lidskii's result only for matrices A that are similar to a Jordan block. Let A be a matrix of order n with $A = XJ_0X^{-1}$, where

$$J_0 = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}.$$

Distinguish the first column of X and the last row of X^{-1},

$$x \equiv X e_1, \qquad y^* \equiv e_n^T X^{-1},$$

which are respective right and left eigenvectors of A. Then [16, Theorem 1]

$$|\mu - \lambda| = |y^* E x|^{1/n} + o\left(\|E\|^{1/n}\right).$$

For instance, consider a Jordan block J_0 of order n, where $x = e_1$ and $y = e_n$, and perturb element $(n, 1)$ of J_0 by ϵ, as in §1.1. Application of Lidskii's result gives $y^* E x = \epsilon$ and the (correct) first order term

$$|\lambda - \mu| = \epsilon^{1/n} + o(\epsilon^{1/n}).$$

3. A Result Similar in Spirit to Lidskii's. We derive an expression for the eigenvalue error that involves only eigenvectors, but no generalised eigenvectors. In contrast to Lidskii's result our's is not a power series expansion.

We need a unit-norm eigenvector z of μ,

$$(A+E)z = \mu z, \qquad \|z\| = 1,$$

with residual

$$r \equiv (A-\mu I)z = -Ez.$$

THEOREM 3.1. *Let A have a single eigenvalue λ, and let P be the orthogonal projector onto the left eigenspace of A. Then*

$$|\lambda-\mu|\ \|Pz\| = \|Pr\|.$$

Proof. Since the columns of the orthogonal projector P are left eigenvectors, $r=(A-\mu I)z$ implies

$$Pr = P(A-\mu I)z = (\lambda-\mu)Pz.$$

□

Theorem 3.1 shows that the eigenvalue perturbation can be expressed in terms of left eigenvectors of A. No generalised eigenvectors enter the picture. Since the projector onto the left eigenspace is orthogonal, the eigenvalue perturbation is independent of the conditioning of the eigenbasis.

Theorem 3.1 leads to a useful bound when $Pz \neq 0$, that is, when the perturbed eigenvector has a contribution in the left eigenspace of A.

COROLLARY 3.2. *If $Pz \neq 0$ in Theorem 3.1 then*

$$|\lambda-\mu| \leq \frac{\|r\|}{\|Pz\|}.$$

The amplification factor $1/\|Pz\|$ in Corollary 3.2 bounds the sensitivity of the eigenvalue λ to perturbations. The following example illustrates that Corollary 3.2 can be tight.

Let J_0 and J_0+E be as in §1.1. Then J_0+E has a perturbed eigenvalue $\mu = \lambda + \epsilon^{1/n}$ with unnormalised eigenvector

$$z \equiv \begin{pmatrix} 1 \\ \epsilon^{1/n} \\ \epsilon^{2/n} \\ \vdots \\ \epsilon^{(n-1)/n} \end{pmatrix}.$$

The orthogonal projector onto the left eigenspace of J_0 is $P = e_n e_n^T$. Hence

$$Pz = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \epsilon^{(n-1)/n} \end{pmatrix}, \qquad r = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\epsilon \end{pmatrix}.$$

This implies

$$\frac{\|r\|}{\|Pz\|} = \epsilon^{1/n},$$

and the bound Corollary 3.2 is equal to the eigenvalue error.

$$|\lambda - \mu| \leq \|r\|.$$

Our numerical experiments have indicated that Corollary 3.2 can be very pessimistic. That's because it uses only a small part of the perturbed eigenvector, namely the part in the unperturbed left eigenspace (note that the equality in Theorem 3.1 can offset a small $\|Pz\|$ by a much smaller $\|Pr\|$, thereby producing a reasonably small eigenvalue error). In the next section we use more information in z to obtain a tighter bound.

4. Extension. We derive an expression for the eigenvalue error that gives useful information for any perturbed eigenvector, even if it has only little contribution in the left eigenspace. This expression represents an extension of Theorem 3.1.

We switch from Jordan to Schur decomposition. Let $A = Q(\lambda I - N)Q^*$ be a Schur decomposition of A, where Q is unitary and N is nilpotent. Although N is not uniquely determined, its Frobenius norm $\|N\|_F$ is [11, §1.2]. Here we choose a particular Schur decomposition that reflects the Jordan structure of A [10, §10], [13, 21]. Let s be the index of λ. Then one can find Q and N such that $\lambda I - N$ is block upper triangular of order s,

$$\lambda I - N = \begin{pmatrix} \lambda I_1 & * & \dots & * \\ & \lambda I_2 & & \vdots \\ & & \ddots & * \\ & & & \lambda I_s \end{pmatrix}. \tag{4.1}$$

Here I_k is the identity matrix of order n_k, where n_k is the number of linearly independent generalised eigenvectors of A of grade $s-k+1$. For instance, n_s is the geometric multiplicity of λ; and n_{s-1} is the number of linearly independent eigenvectors of grade 2. In the special case when A is nondefective (semisimple) $s = 1$ and $A = \lambda I$; and when A is maximally defective (nonderogatory) $s = n$ and $n_k = 1$ for $1 \leq k \leq n$.

Partition Q conformally with $\lambda I - N$,

$$Q = (\, Q_1 \quad \dots \quad Q_s \,)$$

where Q_k is $n \times n_k$. The columns of Q_{s-k+1} are an orthonormal basis for the generalised left eigenspace of grade k. For instance, the columns of Q_s represent an orthonormal basis for the left eigenspace of A. The orthogonal projector onto the invariant subspace consisting of the generalised left eigenspaces of grades 1 through k is

$$P_k \equiv (\, Q_{s-k+1} \quad \dots \quad Q_s \,)(\, Q_{s-k+1} \quad \dots \quad Q_s \,)^*.$$

For instance, P in Theorem 3.1 is here P_1, the orthogonal projector onto the left eigenspace, while P_s is the identity matrix.

THEOREM 4.1. *Let A have a single eigenvalue λ with index s, and let μ and z be a perturbed eigenpair with $Az = \mu z + r$. If P_k is the orthogonal projector onto the generalised left eigenspaces of grades 1 through k then*

$$(\lambda-\mu)^k \, P_k z = P_k Q \left(\sum_{i=0}^{k-1} (\lambda-\mu)^{k-1-i} N^i \right) Q^* r, \qquad 1 \le k \le s.$$

Proof. Partitioning the identity matrix conformally with $\lambda I - N$ gives as its kth block column

$$U_k \equiv \begin{pmatrix} 0 \\ \vdots \\ I_k \\ \vdots \\ 0 \end{pmatrix}, \qquad 1 \le k \le s,$$

where I_k is the identity matrix of order n_k. Below we use the fact that $Q_k = QU_k$ and

$$P_k Q = \begin{pmatrix} Q_{s-k+1} & \dots & Q_s \end{pmatrix} \begin{pmatrix} U_{s-k+1} & \dots & U_s \end{pmatrix}^* . \tag{4.2}$$

Write the residual in terms of the Schur decomposition (4.1)

$$Q^* r = ((\lambda-\mu) I - N) \, Q^* z, \tag{4.3}$$

and distinguish the two cases $\mu = \lambda$ and $\mu \ne \lambda$.

When $\mu = \lambda$, the left-hand side of Theorem 4.1 is zero. To determine the value of the right-hand side note that (4.3) implies $Q^* r = -NQ^* z$. Substituting this in the right-hand side of Theorem 4.1 gives

$$\begin{aligned} P_k Q N^{k-1} Q^* r &= -P_k Q N^k Q^* z \\ &= -\begin{pmatrix} Q_{s-k+1} & \dots & Q_s \end{pmatrix} \begin{pmatrix} U_{s-k+1} & \dots & U_s \end{pmatrix}^* N^k \, Q^* z, \end{aligned}$$

where the second equality follows from (4.2). Since N^k has a zero block diagonal and $k-1$ zero block superdiagonals, $U_{s-j+1}^* N^k = 0$ for $1 \le j \le k$. Hence the right-hand side of Theorem 4.1 is zero and the theorem holds for $\mu = \lambda$.

When $\mu \ne \lambda$ then $(\lambda-\mu)I - N$ in (4.3) is non-singular, and [9, Lemma 2.3.3] implies

$$Q^* z = ((\lambda-\mu)I - N)^{-1} \; Q^* r = \sum_{i=0}^{\infty} \frac{1}{(\lambda-\mu)^{i+1}} N^i \, Q^* r.$$

Multiplying by U_{s-j+1}^* gives

$$Q_{s-j+1}^* z = \sum_{i=0}^{j-1} \frac{1}{(\lambda-\mu)^{i+1}} U_{s-j+1}^* N^i \, Q^* r, \qquad 1 \le j \le k,$$

since, as above, $U_{s-j+1}^* N^i = 0$ for $i \ge j$. Pull U_{s-j+1}^* in front and include a few more zero terms in the sum,

$$Q_{s-j+1}^* z = U_{s-j+1}^* \sum_{i=0}^{k-1} \frac{1}{(\lambda-\mu)^{i+1}} N^i \, Q^* r, \qquad 1 \le j \le k.$$

Now stack the k parts on top of each other,

$$\begin{pmatrix} Q_{s-k+1}^* \\ \vdots \\ Q_s^* \end{pmatrix} z = \begin{pmatrix} U_{s-k+1}^* \\ \vdots \\ U_s^* \end{pmatrix} \sum_{i=0}^{k-1} (\lambda-\mu)^{-i-1} N^i \, Q^* r.$$

Finally, multiply both sides by $(\lambda-\mu)^k \begin{pmatrix} Q_{s-k+1} & \dots & Q_s \end{pmatrix}$ and apply (4.2). □

Therefore, projecting the perturbed eigenvector onto eigenspaces of higher grades leads to an expression for a power of the eigenvalue error. The power of the error increases with the grades of the eigenspaces. Likewise the powers of the non-normal part N on the right-hand side in Theorem 4.1 increase with the grades of the eigenspaces.

When A is a normal matrix with a single eigenvalue Theorem 4.1 shows that the eigenvalue error is equal to the residual norm,

$$|\lambda-\mu| = \|r\|, \tag{4.4}$$

because $N = 0$, $s = 1$, and $P_1 = I$.

In the special case when $k = 1$ Theorem 4.1 reduces to Theorem 3.1. The expression below shows what happens when $k = s$.

COROLLARY 4.2. *If A has a single eigenvalue with index s then*

$$|\lambda-\mu|^s = \left\| \left(\sum_{i=0}^{s-1} (\lambda-\mu)^{s-1-i} N^i \right) Q^* r \right\|.$$

Proof. Apply Theorem 4.1 with $k = s$ and use the fact that $P_s = I$ and $\|z\| = 1$. □

The proofs of [11, Theorem 4] and [23, Theorem IV.1.9] contain statements that imply Corollary 4.2 immediately.

Corollary 4.2 suggests that an eigenvalue with small index s can be sensitive to perturbations if the elements of N are large in magnitude. The following example illustrates this. Let $A = \lambda I - N$ be of order n with $N = \nu e_1 e_2^T$ and $\nu > 0$. Since $N^2 = 0$ the index of λ is $s = 2$. Let $E = \epsilon e_2 e_1^T$ and $\epsilon > 0$. Then the error in the eigenvalues μ of $A + E$ is

$$|\lambda-\mu| = \sqrt{\epsilon\,\nu}.$$

The sensitivity of λ increases as $\nu \to \infty$. Therefore a matrix of order n can have an eigenvalue of index 2 that is arbitrarily sensitive to perturbations.

5. Improvement of Henrici's Bound.

Corollary 4.2 immediately leads to an improvement of Henrici's bound in the special case of matrices with a single eigenvalue.

COROLLARY 5.1. *If A has a single eigenvalue λ with index s then*

$$\frac{|\lambda-\mu|^s}{\|N^{s-1}\| + |\lambda-\mu|\,\|N^{s-2}\| + \ldots + |\lambda-\mu|^{s-1}} \le \|r\|.$$

Corollary 5.1 implies a bound for the special case of matrices with a single eigenvalue which Henrici derived for general square matrices and a larger class of norms [11, Theorem 4]. Below is Stewart and Sun's formulation of Henrici's bound [23, Theorem IV.1.9]. If $\eta \equiv |\lambda-\mu|/\|N\|$ then

$$\frac{\eta^n}{1+\eta+\ldots+\eta^{n-1}} \le \frac{\|E\|}{\|N\|}.$$

Similar bounds by Elsner [7, Theorem 4], [8, Theorem 4.1] and Veseliĉ [18, §2] are also expressed in terms of N. Like Henrici's bound they do not provide a smooth transition between normal and non-normal matrices. The bounds in [9, Theorem 7.2.3] and [5] provide a smooth transition but they are weaker.

When A is a Jordan block Corollary 5.1 implies the extended Bauer-Fike theorem (2.1). Although, like (2.1), Corollary 5.1 depends on the size s of a largest Jordan block, the non-normality can mitigate this dependence. Suppose that the non-normal part N is small, e.g. $\|N\| \leq \epsilon < 1$. Then $\|N^i\| \leq \epsilon$ for all i, and the left side in Corollary 5.1 is at least

$$\frac{|\lambda-\mu|^s}{\epsilon(|\lambda-\mu|+\ldots+|\lambda-\mu|^{s-2})+\ |\lambda-\mu|^{s-1}}.$$

The smaller ϵ the closer we are to bounding the quantity we are really interested in, namely $|\lambda-\mu|$. If ϵ is so tiny that the term containing ϵ becomes negligible then Corollary 5.1 bounds essentially the eigenvalue error $|\lambda-\mu|$ proper rather than a power of the error. Therefore a small non-normal part dampens the impact of the index on the eigenvalue sensitivity (this is confirmed in §6).

On the other extreme, suppose that the powers of the non-normal part N are large. Then the large $\|N^i\|$ drag the left-hand side in Corollary 5.1 even further away from the quantity we really want to bound, namely $|\lambda-\mu|$. Therefore a large non-normal part can amplify the impact of the index on the eigenvalue sensitivity. The example in §1.1 with $\nu \gg 1$ illustrates this.

6. Weakly Non-normal Matrices. We call a matrix *weakly non-normal* if its non-normal part N is on the order of machine precision. We show that eigenvalues of weakly non-normal matrices are as insensitive as eigenvalues of normal matrices.

The following bound was already derived for general square matrices in [7, Theorem 3].

THEOREM 6.1. *If A has a single eigenvalue then*

$$|\lambda-\mu| \leq \|r\| + \|N\|.$$

Proof. Using a Schur decomposition in the residual gives

$$r = Az - \mu z = (\lambda-\mu)z - QNQ^*z,$$

so $(\lambda-\mu)z = r + QNQ^*z$. □

If the matrix is weakly non-normal so that $\|N\| = O(\epsilon_{\mathrm{mach}})$, where ϵ_{mach} is the machine precision, then

$$|\lambda-\mu| \leq \|r\| + O(\epsilon_{\mathrm{mach}}).$$

Hence the eigenvalue error is not much larger than the residual norm. This implies that the eigenvalues of a weakly non-normal matrix are no more sensitive to perturbations than the eigenvalues of a normal matrix.

For instance, consider the matrix A from §1.1, now with $\nu = \epsilon$,

$$A = \begin{pmatrix} \lambda & \epsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \epsilon \\ & & & \lambda \end{pmatrix},$$

and assume that E is an arbitrary perturbation with $\|E\| = \epsilon$. Theorem 6.1 bounds the error in any eigenvalue μ of $A+E$ by

$$|\lambda - \mu| \leq 2\epsilon.$$

Therefore a highly defective eigenvalue with a large Jordan condition number can be insensitive to perturbations.

7. The Resolvent. For normal matrices with a single eigenvalue the eigenvalue error is always equal to the residual norm, see (4.4). Unfortunately this is not true for non-normal matrices. We show that the resolvent norm of a non-normal matrix can increase as a power of the non-normality. But the inverse of the resolvent norm is the smallest possible residual norm. Hence the smallest possible residual norm can decrease as a power of the non-normality, independent of the eigenvalue error.

Given a matrix A and a number μ, the norm of the smallest possible residual associated with μ is the inverse of the resolvent norm [10, §15], [12, §2.4], [24, §2]. In our case this means, if A has a single eigenvalue λ and $\mu \neq \lambda$, that

$$\min_{\|z\|=1} \|(A - \mu I)z\| = \frac{1}{\|(A-\mu I)^{-1}\|}.$$

Now express the the resolvent norm in terms of the non-normal part N of a Schur decomposition and use the fact that the index of nilpotency of N is equal to the index of λ. To see this let $A = Q(\lambda I - N)Q^*$ be a Schur decomposition, and let s be the index of λ. Then

$$(A - \lambda I)^{s-1} \neq 0, \qquad (A - \lambda I)^s = 0,$$

if and only if s is the smallest number such that $N^s = 0$.

THEOREM 7.1. *Let A be a matrix with a single eigenvalue λ and let $\mu \neq \lambda$. Then*

$$\|(A - \mu I)^{-1}\| = \frac{1}{|\lambda - \mu|^s} \, \|\sum_{i=0}^{s-1} (\lambda - \mu)^{s-1-i} N^i\|,$$

where s is the index of λ.

Proof. First use a Schur decomposition,

$$A - \mu I = (\lambda - \mu)I - QNQ^* = (\lambda - \mu)(I - \widetilde{N}),$$

where $\widetilde{N} \equiv \frac{1}{\lambda - \mu} QNQ^*$. Then use the fact that the index of nilpotency of $\widetilde{N}$ is s,

$$(A - \mu I)^{-1} = \frac{1}{(\lambda - \mu)} (I - \widetilde{N})^{-1} = \frac{1}{(\lambda - \mu)} \sum_{i=0}^{s-1} \widetilde{N}^i.$$

Hence

$$\|(A - \mu I)^{-1}\| = \frac{1}{|\lambda - \mu|^s} \, \|\sum_{i=0}^{s-1} (\lambda - \mu)^{s-1-i} N^i\|.$$

□

There are two cases when the resolvent norm is large. In the good case the denominator in Theorem 7.1 is small, which means that the perturbed eigenvalue is a good approximation to λ. In the bad case the numerator is large, which means that N has large elements and possibly that s is large. For the residual this implies that a residual norm significantly smaller than machine precision is a sign that the matrix is highly non-normal. For instance, one iteration of inverse iteration often produces an abnormally small residual [12, §5.4]. In this case we cannot conclude that the computed eigenvalue and eigenvector are accurate, we can only conclude that the matrix is highly non-normal.

The following example [12, §5.4] illustrates that the resolvent norm can increase as a power of the nonnormality – independent of the eigenvalue error. Consider again a matrix A like the one in §1.1 but now with $-\nu$ on the off-diagonal, where $\nu > 0$ and $\lambda - \mu > 0$. Then

$$A - \mu I = (\lambda - \mu) \begin{pmatrix} 1 & -\nu/(\lambda-\mu) & & \\ & \ddots & \ddots & \\ & & 1 & -\nu/(\lambda-\mu) \\ & & & 1 \end{pmatrix},$$

and $(A - \mu I)^{-1}$ is equal to

$$\frac{1}{(\lambda-\mu)^n} \begin{pmatrix} (\lambda-\mu)^{n-1} & (\lambda-\mu)^{n-2}\nu & \dots & \nu^{n-1} \\ & \ddots & \ddots & \vdots \\ & & (\lambda-\mu)^{n-1} & (\lambda-\mu)^{n-2}\nu \\ & & & (\lambda-\mu)^{n-1} \end{pmatrix}.$$

Because $\|(A - \mu I)^{-1}\|$ is bounded below by the magnitude of its $(1, n)$ element, we get a lower bound for the resolvent norm

$$\|(A - \mu I)^{-1}\| \geq \frac{\nu^{n-1}}{|\lambda - \mu|^n}.$$

In this example the resolvent norm increases as a power of $\|N\|$ as $\|N\| \to \infty$, independent of the eigenvalue error $|\lambda - \mu|$, while the minimal possible residual norm decreases as a power of $\|N\|$.

8. Summary. For matrices with a single eigenvalue we have analysed the sensitivity of the eigenvalue to perturbations in the matrix.

First we derived an expression for the eigenvalue error that consists of projections onto the left eigenspace; no generalised eigenspaces of higher grade are involved. This expression for the eigenvalue error does not depend on the conditioning of the left eigenbasis.

We extended this result by deriving an expression for the kth power of the eigenvalue error that involves a projection onto an eigenspaces of grade 1 through k, as well as powers of the non-normal part N in a Schur decomposition. We illustrated that an eigenvalue with small index can be highly sensitive to perturbations when $\|N\|$ is large. Vice versa, a highly defective eigenvalue is insensitive if $\|N\|$ is small.

From this more general expression for the eigenvalue error we derived a tighter version (for the special case of matrices with a single eigenvalue) of a bound by Henrici. The bound suggests that a small non-normality $\|N\|$ dampens the impact of the index on eigenvalue sensitivity, while a large $\|N\|$ can amplify the impact of the index on the eigenvalue sensitivity.

We defined weakly non-normal matrices as those matrices where $\|N\|$ is on the order of machine precision. We showed that their eigenvalues are no more sensitive than eigenvalues of normal matrices. In particular this means that a highly defective eigenvalue with large Jordan condition number can be insensitive to perturbations.

At last, we illustrated that the resolvent norm can increase as a power of $\|N\|$ – independent of the eigenvalue error. If the residual is significantly smaller than machine precision, one cannot conclude that the computed eigenvalue and eigenvector are accurate. One can only conclude that the matrix is highly non-normal.

Acknowledgement. We thank Stan Eisenstat for suggestions that significantly improved the paper and Nick Trefethen for correcting several errors, and a referee for simplifying the proof Theorem 3.1.

REFERENCES

[1] F. Bauer and C. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–141.
[2] T. Braconnier and Y. Saad, *Eigenvalue bounds from the Schur form*, tech. rep., University of Minnesota, Department of Computer Science and Engineering, February 1998.
[3] F. Chaitin-Chatelin and V. Frayssé, *Lectures on Finite Precision Computations*, SIAM, Philadelphia, 1996.
[4] F. Chatelin, *Valeurs Propres de Matrices*, Masson, 1986.
[5] K. Chu, *Generalization of the Bauer-Fike theorem*, Numer. Math, 49 (1986), pp. 685–91.
[6] J. Demmel, *On condition numbers and the distance to the nearest ill-posed problem*, Numer. Math., 51 (1987), pp. 251–89.
[7] L. Elsner, *Einschliessungssätze für Eigenwerte nichtnormaler Matrizen*, Funktionalanalysis, Approximationstheorie, Numerische Mathematik, 7 (1967), pp. 185–93.
[8] ———, *Perturbation theorems for the joint spectrum of commuting matrices: A conservative approach*, Linear Algebra Appl., 208/209 (1983), pp. 83–95.
[9] G. Golub and C. van Loan, *Matrix Computations*, Johns Hopkins, Baltimore, 1996.
[10] G. Golub and J. Wilkinson, *Ill-conditioned eigensystems and the computation of the Jordan canonical form*, SIAM Review, 18 (1976), pp. 578–619.
[11] P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, Numer. Math., 14 (1962), pp. 24–40.
[12] I. Ipsen, *Computing an eigenvector with inverse iteration*, SIAM Review, 39 (1997), pp. 254–291.
[13] B. Kågström and A. Ruhe, *An algorithm for numerical compuatation of the Jordan normal form of a complex matrix*, ACM Trans. Math. Software, 6 (1980), pp. 398–419.
[14] W. Kahan, B. Parlett, and E. Jiang, *Residual bounds on approximate eigensystems of nonnormal matrices*, SIAM J. Numer.Anal., 19 (1982), pp. 470–484.
[15] R. Li, *On perturbation bounds for eigenvalues of a matrix.* Manuscript, Computing Center, Academia Sinica, Beijing, P.R. China, 1985.
[16] V. Lidskii, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Maths. Math. Phys., 1 (1966), pp. 73–85. (Zh. vȳchisl. Mat. mat. Fiz. 6, 1, 52-60, 1965).
[17] G. Loizou, *Nonnormality and Jordan condition numbers of matrices*, J. Assoc. Comput. Mach., 16 (1969), pp. 580–4.
[18] D. Meyer and K. Veseliċ, *On some new inclusion theorems for the eigenvalues of partitioned matrices*, Numer. Math., 34 (1980), pp. 431–7.
[19] J. Moro, J. Burke, and M. Overton, *On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (to appear), p. 4.
[20] J. Ortega, *Numerical Analysis, A Second Course*, Academic Press, New York, 1972.
[21] A. Ruhe, *An algorithm for numerical determination of the structure of a general matrix*, BIT, 10 (1970), pp. 196–216.
[22] R. Smith, *The condition numbers of the matrix eigenvalue problem*, Numer. Math., 10 (1967), pp. 232–40.
[23] G. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.
[24] J. Varah, *On the separation of two matrices*, SIAM J. Numer. Anal., 16 (1979), pp. 216–222.
[25] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.