RELATIVE PERTURBATION RESULTS FOR EIGENVALUES AND EIGENVECTORS OF DIAGONALISABLE MATRICES

STANLEY C. EISENSTAT* AND ILSE C. F. IPSEN†

Abstract. Let $\hat{\lambda}$ and $\hat{x}$ be a perturbed eigenpair of a diagonalisable matrix A. The problem is to bound the error in $\hat{\lambda}$ and $\hat{x}$. We present one absolute perturbation bound and two relative perturbation bounds.

The absolute perturbation bound implies that the condition number for $\hat{x}$ is the norm of an orthogonal projection of the reduced resolvent at $\hat{\lambda}$. This condition number can be a lot less pessimistic than the traditional one, which is derived from a first-order analysis. A further upper bound leads to an extension of Davis and Kahan's $\sin\theta$ Theorem from Hermitian to diagonalisable matrices.

The two relative perturbation bounds assume that $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a perturbed matrix D_1AD_2, where D_1 and D_2 are non-singular, but D_1AD_2 is not necessarily diagonalisable. We derive a bound on the relative error in $\hat{\lambda}$ and a $\sin\theta$ theorem based on a relative eigenvalue separation. The perturbation bounds contain both the deviation of D_1 and D_2 from similarity and the deviation of D_2 from identity.

Key words. eigenvalue, eigenvector, condition number, relative error, diagonalisable matrix, reduced resolvent, angle between subspaces

AMS subject classification. 15A18, 15A42, 15A60, 65F15, 65F35, 65G99

1. Introduction.

The protagonist is a complex, diagonalisable matrix A with eigendecomposition

$$A = X\Lambda X^{-1},$$

where Λ is a diagonal matrix whose diagonal elements λ_i are the eigenvalues of A. The complex number $\hat{\lambda}$, the 'perturbed eigenvalue', approximates an eigenvalue of A; and the unit vector $\hat{x}$, the 'perturbed eigenvector', approximates an eigenvector. The problem is to bound the error in $\hat{\lambda}$ and $\hat{x}$. Our results extend the work in [19, 18] and are stronger and simpler than the results in [24, 25].

First we derive an upper bound on $\sin\theta$, where θ is the angle between $\hat{x}$ and the eigenspace corresponding to those eigenvalues closest to $\hat{\lambda}$ (Theorem 4.1). Our 'condition number' is the norm of an orthogonal projection of the reduced resolvent, which can be arbitrarily much smaller than the norm of the reduced resolvent that appears in a first-order analysis [32, Definition 3.3.2]. An upper bound on this condition number leads to an extension of Davis and Kahan's $\sin\theta$ Theorem [9, §6], [10, §2] from Hermitian to diagonalisable matrices (Corollary 4.3). Compared to the Hermitian case, the new bound contains the additional factor κ_2, the condition number of a subset of the left eigenvectors. This is an exact rather than a first-order bound and holds without any assumptions on the perturbations.

Next we derive two relative perturbation bounds based on a multiplicative perturbation model. We assume that $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a perturbed matrix D_1AD_2, where D_1 and D_2 are non-singular matrices and D_1AD_2 need not be diagonalisable. We prove a bound on the relative error in $\hat{\lambda}$ that contains the

* Department of Computer Science, Yale University, P. O. Box 208285, New Haven, CT 06520-8285 (eisenstat-stan@cs.yale.edu). The research of this author was supported in part by NSF grant CCR-9400921.

† Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205 (ipsen@math.ncsu.edu). The research of this author was supported in part by NSF grants CCR-9102853 and CCR-9400921.

deviation of D_1 and D_2 from similarity (Theorem 6.1). The bound is tight when the perturbation is a similarity transformation $D_1 = D_2^{-1}$ or when $\hat{\lambda} = 0$.

From the absolute $\sin\theta$ theorem described above we derive a relative bound on $\sin\theta$ that consists of two summands: the deviation of D_2 from identity; and the deviation of D_1 and D_2 from similarity, amplified by κ_2 and a relative eigenvalue separation (Theorem 7.2). We conclude that the eigenvectors of diagonalisable matrices are well-conditioned when the perturbation is caused by a similarity transformation; and that the null vectors of diagonalisable matrices are well-conditioned when the perturbation is multiplicative (Corollary 7.3).

Our relative perturbation bounds are no stronger than the traditional absolute perturbation bounds because we use the traditional bounds to derive them.

Notation. $\|\cdot\|$ denotes the two-norm. If V is a matrix of full column rank, then

$$V^{\dagger} \equiv (V^*V)^{-1}V^*$$

is the Moore Penrose-inverse of V and

$$\kappa(V) \equiv \|V\|\,\|V^{\dagger}\|$$

is its two-norm condition number.

2. Related Work. In 1959 Ostrowski gave the first relative perturbation bounds for eigenvalues by considering multiplicative perturbations D^*AD of Hermitian matrices A [28],[22, Theorem 4.5.9]. He bounded the ratio of exact and perturbed eigenvalues in terms of the smallest and largest eigenvalues of D^*D,

$$\lambda_{min}(D^*D)\,\lambda_i(A) \leq \lambda_i(D^*AD) \leq \lambda_{max}(D^*D)\,\lambda_i(A).$$

About thirty years later bounds on the ratio of exact and perturbed eigenvalues appeared for different types of structured matrices, but without any reference to Ostrowski's bounds. Inspired by Kahan's work in 1966 [23], Demmel and Kahan [12] and Deift, Demmel, Li, and Tomei [11] derive relative perturbation bounds for eigenvalues of real symmetric tridiagonal matrices with zero diagonal. Demmel and Gragg [14] extend these bounds to real symmetric positive-definite matrices with acyclic graphs. Deift et al. [11] also derive bounds on the error angles of eigenvectors in terms of a relative eigenvalue separation.

Subsequent work on perturbation theory has extended these results to larger classes of matrices. These bounds are derived for two different perturbation models: component-wise relative perturbations and multiplicative perturbations.

In the context of (a superset of) component-wise relative perturbations, Barlow and Demmel [1] derive bounds for real symmetric scaled diagonally dominant matrices; and Demmel and Veselić [13] derive bounds for real symmetric positive-definite matrices. Mathias [26] unifies the results by Demmel and Veselić [13] with the results by Barlow and Demmel [1] for positive-definite matrices. Veselić and Slapničar [33, 37] extend these bounds to indefinite Hermitian matrices. In addition they prove a relative Gerschgorin theorem for diagonalisable matrices. Pietzsch [29] adapts the eigenvalue bounds in [1, 13, 33, 37] to real skew-symmetric matrices $A = -A^T$ by observing that $iA = -iA^T$ is Hermitian. Drmač [15, 16] gives residual bounds for Hermitian matrices.

In the context of multiplicative perturbations of Hermitian matrices, Eisenstat and Ipsen [19, 18] bound the relative error in the eigenvalues and the error angle between subspaces. Similarly, Li [24, 25] derives bounds for eigenvalues of diagonalisable matrices and for eigenvectors of Hermitian matrices. About ten years earlier, Price [30] had already recommended multiplicative perturbations in the form of matrix exponentials to bound the absolute error in matrix transformations. Applications include the computation of eigenvalues of symmetric matrices by Jacobi's method [30, §3] and the solution of bordered systems of linear equations [21].

Finally, Drmač and Hari [15, 16, 17] derive relative perturbation bounds for Hermitian matrices when the perturbed eigenvalues are Ritz values.

3. Setting the Stage. To establish a correspondence between perturbed and exact quantities, we partition the eigenvalues of A so that Λ_1 contains all eigenvalues closest to $\hat{\lambda}$ and Λ_2 contains the remaining eigenvalues; i.e.,

$$\Lambda = \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix}$$

where

$$\|\Lambda_1 - \hat{\lambda} I\| = \min_i |\lambda_i - \hat{\lambda}|, \qquad \text{and} \qquad \|\Lambda_1 - \hat{\lambda} I\| < 1/\|(\Lambda_2 - \hat{\lambda} I)^{-1}\|.$$

Thus $\|\Lambda_1 - \hat{\lambda} I\|$ represents the absolute error in $\hat{\lambda}$.

The 'absolute gap'

$$\text{absgap} \equiv 1/\|(\Lambda_2 - \hat{\lambda} I)^{-1}\| = \min_i |(\Lambda_2)_{ii} - \hat{\lambda}|$$

is the absolute separation of $\hat{\lambda}$ from the eigenvalues of Λ_2. It approximates the absolute separation of the desired eigenvalues Λ_1 from the remaining eigenvalues Λ_2.

We partition the eigenvectors conformally with Λ,

$$X = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \qquad \text{and} \qquad X^{-1} = \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix},$$

where * denotes the conjugate transpose. Then $\hat{x}$ can be regarded as approximating a unit vector in range(X_1). Since Λ_1 consists of all eigenvalues on a circle of minimal radius around $\hat{\lambda}$, in general range(X_1) is associated with several distinct, possibly multiple eigenvalues.

To determine how well the unit vector $\hat{x}$ approximates a vector in the right eigenspace range(X_1), we measure the angle θ between $\hat{x}$ and range(X_1). More precisely, let $0 \leq \theta \leq \pi/2$ be the largest principal angle between range$(\hat{x})$ and range(X_1) [20, §12.4.3]. Then $\sin\theta$ is obtained from the orthogonal projection

$$P \equiv I - (X_1^\dagger)^* X_1^*.$$

of $\hat{x}$ onto range$(X_1)^-$, the orthogonal complement of range(X_1).

THEOREM 3.1 ([9, §6], [10, P. 10], [36, (2.3)]).

$$\sin\theta = \|P\hat{x}\|.$$

When A is normal then range(X_2) = range$(X_1)^-$ and X_2 has orthonormal columns. Hence

$$P = X_2 X_2^* \qquad \text{and} \qquad \sin\theta = \|X_2 X_2^* \hat{x}\| = \|X_2^* \hat{x}\|.$$

When A is non-normal, X_2 may not be the orthogonal complement of X_1, and the columns of X_2 may not be orthonormal. Our perturbation bounds for eigenvectors of diagonalisable matrices are derived by bounding $\sin\theta$ from above.

The following result expresses P in terms of a basis Y_2 for the left eigenspace range(Y_2).

LEMMA 3.2. *The orthogonal projector P satisfies*

$$P = (Y_2^{\dagger})^* Y_2^*.$$

Proof. $X^{-1}X = I$ implies range$(Y_2) = $ range$(X_1)^{-}$. Thus

$$P = I - (X_1^{\dagger})^* X_1^* = (Y_2^{\dagger})^* Y_2^*. \qquad \square$$

In general the orthogonal projector P in Lemma 3.2 is not identical to the (oblique) spectral projector

$$Z \equiv X_2 Y_2^*.$$

Although both projectors share the same null space, range(X_1), their ranges are different. One projects on the left eigenspace, while the other projects on the right one:

$$\text{range}(P) = \text{range}(Y_2) \qquad \text{and} \qquad \text{range}(Z) = \text{range}(X_2).$$

4. A $\sin\boldsymbol{\theta}$ Theorem for Diagonalisable Matrices. The traditional perturbation bound for the eigenvectors of a Hermitian matrix is Davis and Kahan's $\sin\theta$ Theorem [9, §6], [10, §2],

$$\sin\theta \le \|r\|/\text{absgap},$$

where $r \equiv (A - \hat\lambda I)\hat x$ is the residual. That is, for a Hermitian matrix the condition number of the perturbed eigenvector is inversely proportional to the separation between $\hat\lambda$ and the remaining eigenvalues Λ_2. We extend this result to diagonalisable matrices by first deriving a $\sin\theta$ theorem whose condition number is an orthogonal projection of the 'reduced resolvent at $\hat\lambda$',

$$S(\hat\lambda) \equiv X_2(\Lambda_2 - \hat\lambda I)^{-1} Y_2^*$$

(see [7, §2.2], [32, §III.3.2]).

THEOREM 4.1. *The angle θ satisfies*

$$\sin\theta \le \|PS(\hat\lambda)\|\,\|r\|.$$

Proof. As our goal is a residual bound for the angle, we write

$$r = (A - \hat\lambda I)\hat x = X \begin{pmatrix} \Lambda_1 - \hat\lambda I & \\ & \Lambda_2 - \hat\lambda I \end{pmatrix} X^{-1}\hat x = X \begin{pmatrix} \Lambda_1 - \hat\lambda I & \\ & \Lambda_2 - \hat\lambda I \end{pmatrix} \begin{pmatrix} \hat c \\ \hat s \end{pmatrix},$$

where $\hat c \equiv Y_1^* \hat x$ and $\hat s \equiv Y_2^* \hat x$. Multiplying both sides by $(\Lambda_2 - \hat\lambda I)^{-1} Y_2^*$ gives

$$\hat s = (\Lambda_2 - \hat\lambda I)^{-1}\, Y_2^* r.$$

Then Lemma 3.2 and $Y_2^* X_2 = I$ imply

$$P\hat{x} = (Y_2^\dagger)^* Y_2^* \hat{x} = (Y_2^\dagger)^* \hat{s} = (Y_2^\dagger)^* Y_2^* X_2 (\Lambda_2 - \hat{\lambda} I)^{-1}\, Y_2^* r = PS(\hat{\lambda}) r.$$

Finally, Theorem 3.1 implies

$$\sin\theta = \|P\hat{x}\| = \|PS(\hat{\lambda})r\| \le \|PS(\hat{\lambda})\|\, \|r\|. \qquad \square$$

When A is normal then $P = X_2 X_2^*$ and X_2 has orthonormal columns. Hence

$$\|PS(\hat{\lambda})\| = \|S(\hat{\lambda})\| = 1/\text{absgap},$$

and Theorem 4.1 reduces to Davis and Kahan's $\sin\theta$ Theorem. Moreover, it is not difficult to extend Theorem 4.1 to eigenspaces of larger dimension.

Let's compare $\|PS(\hat{\lambda})\|$ to a similar condition number: A first-order perturbation analysis for arbitary matrices yields as condition number for $\hat{x}$ the norm of the reduced resolvent, $\|S(\hat{\lambda})\|$ [32, Definition III.3.2]. When A is diagonalisable and $\hat{\lambda}$ is simple, an expansion of $S(\hat{\lambda})$ gives information about the sensitivity of the eigenspace corresponding to $\hat{x}$ [20, §7.2.4], [32, §III.3.2]. The following example shows that our condition number $\|PS(\hat{\lambda})\|$ can be much less pessimistic than $\|S(\hat{\lambda})\|$.

Remark. $\|PS(\hat{\lambda})\|$ can be much smaller than $\|S(\hat{\lambda})\|$.

Since P is an orthogonal projector, $\|P\| \le 1$ and $\|PS(\hat{\lambda})\| \le \|S(\hat{\lambda})\|$. Consider the case $\Lambda_2 = \lambda_2 I$. Then the reduced resolvent is a multiple of the spectral projector Z so that

$$S(\hat{\lambda}) = \frac{1}{\lambda_2 - \hat{\lambda}}\, Z \qquad \text{and} \qquad PS(\hat{\lambda}) = \frac{1}{\lambda_2 - \hat{\lambda}}\, P.$$

Thus

$$\|S(\hat{\lambda})\| = \frac{1}{|\lambda_2 - \hat{\lambda}|}\, \|Z\| \qquad \text{and} \qquad \|PS(\hat{\lambda})\| = \frac{1}{|\lambda_2 - \hat{\lambda}|}.$$

Let $X = QL$ be a QL decomposition of X, where Q is unitary and L is lower triangular. Partitioning Q and L conformally with X,

$$Q = (\, Q_1 \quad Q_2 \,) \qquad \text{and} \qquad L = \begin{pmatrix} L_{11} & \\ L_{21} & L_{22} \end{pmatrix},$$

gives

$$X_2 = Q_2 L_{22} \qquad \text{and} \qquad Y_2^* = (\, -L_{22}^{-1} L_{21} L_{11}^{-1} \quad L_{22}^{-1} \,)\, Q^*.$$

Thus

$$Z = X_2 Y_2^* = Q_2\ (\, -L_{21} L_{11}^{-1} \quad I \,)\, Q^*$$

and

$$\|Z\| = \|\, (\, -L_{21} L_{11}^{-1} \quad I \,)\, \| \ge \|L_{21} L_{11}^{-1}\|.$$

Choosing L_{11} and L_{21} appropriately makes $\|Z\|$, and hence $\|S(\hat{\lambda})\|$, arbitrarily large while $\|PS(\hat{\lambda})\|$ remains constant.

To extract the eigenvalue separation from $\|PS(\hat{\lambda})\|$, we derive an upper bound.

COROLLARY 4.2. *The orthogonal projection of the reduced resolvent satisfies*

$$\|PS(\hat{\lambda})\| \leq \kappa_2/\text{absgap}$$

where $\kappa_2 \equiv \kappa(Y_2) \geq 1$.

Proof. The result follows from

$$\begin{aligned} \|PS(\hat{\lambda})\| &= \|(Y_2^\dagger)^* Y_2^* X_2 (\Lambda_2 - \hat{\lambda} I)^{-1} \, Y_2^*\| \\ &= \|(Y_2^\dagger)^* (\Lambda_2 - \hat{\lambda} I)^{-1} \, Y_2^*\| \\ &\leq \|(Y_2^\dagger)^*\| \, \|(\Lambda_2 - \hat{\lambda} I)^{-1}\| \, \|Y_2^*\| \end{aligned}$$

since $Y_2^* X_2 = I$. ☐

COROLLARY 4.3. *The angle* θ *satisfies*

$$\sin\theta \leq \kappa_2 \, \|r\|/\text{absgap}.$$

The worse conditioned the left eigenvectors Y_2 and hence X^{-1}, the larger κ_2 is likely to be. If Y_2 has orthonormal columns, which occurs when A is normal, then $\kappa_2 = 1$ and Theorem 4.3 reduces to Davis and Kahan's $\sin\theta$ Theorem. Although Davis and Kahan considered only Hermitian matrices, the ideas for this extension to diagonalisable matrices are already present in [10, Theorem 6.1] (essentially their factor $1/\epsilon$ corresponds to our κ_2 above).

Related Work. Bhatia, Davis and McIntosh [4] prove $\sin\theta$ theorems in the more general context of normal operators on Hilbert spaces. Their bounds are of the same form as Corollary 4.3,

$$\sin\theta \leq c \, \|A\hat{X} - \hat{X}B\|/\text{absgap},$$

where A and B are normal operators and $\hat{X}$ represents a perturbed subspace of any dimension. Determining the value of the positive constant c amounts to solving a minimisation problem for functions in $\mathcal{L}_1$ [3].

The existing results cited below hold in the general situation when the dimension of the perturbed subspace is arbitrary. Here we restrict the discussion to our context, which means exact and the perturbed eigenspace have the same dimension. Hence X_1 consists of a single unit vector, say x_1.

For diagonalisable matrices Varah [36, Theorem 2.2] shows that if $\|r\|$ is sufficiently small then

$$\sin\theta \leq \kappa(X) \, \|r\|/\text{absgap}.$$

For general, possibly defective matrices, Stewart [34, Theorem 4.1], [35, Theorem 4.11] derives a $\tan\theta$ bound for the case when $\hat{\lambda}$ is a Rayleigh quotient and $\|r\|$ is sufficiently small. And Ruhe [31, Corollary 1] bounds the sine of the angle between $\hat{x}$ and a singular vector associated with the smallest singular value of A (here absgap is replaced by the gap $\sqrt{\sigma_{n-1}^2 - \sigma_n^2}$ between the two smallest singular values of A).

Instead of $\sin\theta$ one can also bound $\|x_1 - \hat{x}\|$. Such bounds are slightly stronger because they imply $\sin\theta$ theorems:

$$\|x_1 - \hat{x}\| = 2\sin\frac{\theta}{2} \geq 2\cos\frac{\theta}{2}\sin\frac{\theta}{2} = \sin\theta.$$

For instance, Bohte's bound [5, (11)] and its *a posteriori* version [6, (2.11)] are of the form

$$\|x_1 - \hat{x}\| \leq 2\kappa(X)\,\|r\|/\mathrm{absgap}.$$

For normal matrices Wilkinson [38, (3.54.14)] shows that

$$\|x_1 - \alpha\hat{x}\| \leq \frac{\|r\|}{\mathrm{absgap}}\sqrt{1 + \left(\frac{\|r\|}{\mathrm{absgap}}\right)^2},$$

where α is some scalar with $|\alpha| = 1$. Since $\kappa_2 = 1$ for normal matrices, Corollary 4.3 is stronger than Wilkinson's bound.

5. Multiplicative Perturbations. Assume that the perturbed quantities $\hat{\lambda}$ and $\hat{x}$ are an exact eigenvalue and eigenvector of a matrix D_1AD_2, where D_1 and D_2 are non-singular matrices; i.e.,

$$(D_1AD_2)\,\hat{x} = \hat{\lambda}\hat{x}.$$

The matrix D_1AD_2 is not required to be diagonalisable. When $D_1 = D_2^{-1}$ then D_1AD_2 is a similarity transformation of A, which means that A and D_1AD_2 have the same eigenvalues. When $D_1 = D_2^*$ then D_1AD_2 is a congruence transformation of A, which means for Hermitian A that A and D_1AD_2 have the same inertia.

The derivation of relative perturbation bounds for such *multiplicative perturbations* is based on the following idea. From $(D_1AD_2)\hat{x} = \hat{\lambda}\hat{x}$ follows

$$A(D_2\hat{x}) = \hat{\lambda}D_1^{-1}D_2^{-1}(D_2\hat{x}).$$

Setting $z \equiv D_2\hat{x}/\|D_2\hat{x}\|$, we have

$$Az = \hat{\lambda}D_1^{-1}D_2^{-1}z.$$

Thus the residual for $\hat{\lambda}$ and z is

$$f \equiv Az - \hat{\lambda}z = \hat{\lambda}\,(D_1^{-1}D_2^{-1} - I)z = \hat{\lambda}\,(D_1^{-1} - D_2)\hat{x}/\|D_2\hat{x}\|.$$

The upper bounds

$$\|f\| \leq |\hat{\lambda}|\,\|I - D_1^{-1}D_2^{-1}\| \qquad \text{and} \qquad \|f\| \leq |\hat{\lambda}|\,\|D_1^{-1} - D_2\|/\|D_2\hat{x}\|$$

consist of two parts: the factor $|\hat{\lambda}|$, which is responsible for the perturbation bound being relative, and the deviation of D_1 and D_2 from similarity. The factor $\|D_1^{-1} - D_2\|$ is an absolute deviation from similarity, while the factor $\|I - D_1^{-1}D_2^{-1}\|$ constitutes a relative deviation because

$$I - D_1^{-1}D_2^{-1} = (D_2 - D_1^{-1})D_2^{-1}$$

is a difference relative to D_2.

6. Eigenvalues of Multiplicative Perturbations. We now bound the relative error in $\hat{\lambda}$ in terms of the relative deviation from similarity of the matrices D_1 and D_2.

THEOREM 6.1. *The perturbed eigenvalue $\hat{\lambda}$ satisfies*

$$\|\Lambda_1 - \hat{\lambda} I\| \leq |\hat{\lambda}|\, \kappa(X)\, \|I - D_1^{-1} D_2^{-1}\|.$$

Proof. Apply the Bauer-Fike Theorem with residual bound [2, Theorem IIIa], [27, §3.2.7] to the residual f for $\hat{\lambda}$ and z to get

$$\|\Lambda_1 - \hat{\lambda} I\| = \min_i |\lambda_i - \hat{\lambda}| \leq \kappa(X)\, \|f\|$$

and bound $\|f\|$ as above. □

In the case of a similarity transformation $D_1 = D_2^{-1}$, the eigenvalues are preserved and the bound above is zero, which is tight. When $\hat{\lambda} = 0$, A itself must be singular and the bound is zero, which is again tight.

It is not clear how to extend Theorem 6.1 to general, possibly defective matrices and end up with pleasant bounds. For instance, the generalisation of the Bauer-Fike Theorem [8, Theorem 3A] gives an upper bound on the relative error that contains a factor $1/|\hat{\lambda}|^{(p-1)/p}$, where p is the index of nilpotency of the strictly upper triangular matrix in a Schur decomposition of A.

Related Work. Veselić and Slapničar [37] prove a relative Gerschgorin theorem for perturbed matrices $A + \delta A$ with $\|\delta A x\| \leq \eta\, \|Ax\|$ for any x. They show that the eigenvalues $\hat{\lambda}$ of $A + \delta A$ lie in the union of disks [37, Theorem 3.17]

$$|\lambda_i - \hat{\lambda}| \leq |\lambda_i|\, \kappa(X)\, \eta.$$

Li [24] establishes a one-to-one pairing between the exact and perturbed eigenvalues and bounds the errors in all eigenvalue pairs simultaneously:

THEOREM 6.2 (LI [24, THEOREMS 6.1 AND 6.4]). *Let $D_1 A D_2$ be diagonalisable with eigenvalues $\hat{\lambda}_i$ eigenvector matrix $\hat{X}$.*

Then there exists a permutation τ such that

$$\sqrt{\sum_i \frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|^2}{\lambda_i^2 + \hat{\lambda}_{\tau(i)}^2}} \;\leq\; \kappa(X)\, \kappa(\hat{X}) \min\left\{ \sqrt{\|I - D_1\|_F^2 + \|I - D_2^{-1}\|_F^2},\; \sqrt{\|I - D_1^{-1}\|_F^2 + \|I - D_2\|_F^2} \right\}.$$

If in addition all λ_i and all $\hat{\lambda}_i$ are nonnegative then for any p-norm, $1 \leq p \leq \infty$,

$$\max_i \frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|}{\sqrt[q]{\lambda_i^q + \hat{\lambda}_{\tau(i)}^q}} \;\leq\; \kappa(X)\, \kappa(\hat{X}) \min\left\{ \sqrt[s]{\|I - D_1\|_p^s + \|I - D_2^{-1}\|_p^s},\; \sqrt[s]{\|I - D_1^{-1}\|_p^s + \|I - D_2\|_p^s} \right\},$$

where $s = q/(q-1)$.

In contrast to Theorem 6.1, these bounds are non-zero for nontrivial similarity transformations $D_1 = D_2^{-1} \neq I$.

In the special case $D_1 = I$ or $D_2 = I$, Li [24, Theorem 6.6] proves a variation of Theorem 6.1 where the upper bound contains the exact eigenvalue instead of the perturbed one.

The following two bounds can be interpreted as a relative version of Weyl's theorem [22, Theorem 4.3.1]. They follow from the relative bounds by Ostrowski [28], [22, Theorem 4.5.9] mentioned in the beginning of §2.

THEOREM 6.3 (EISENSTAT AND IPSEN [19, THEOREM 2.1]). *Let A be Hermitian with λ_i in increasing order, let D be non-singular, and let $\hat{\lambda}_i$ be the eigenvalues of D^*AD in increasing order. Then*

$$|\hat{\lambda}_i - \lambda_i| \le |\lambda_i| \, \|I - D^*D\| \qquad \text{for all } i.$$

COROLLARY 6.4 (LI [24, THEOREM 7.2]). *Under the assumptions of Theorem 6.3,*

$$|\lambda_i - \hat{\lambda}_i| \le |\hat{\lambda}_i| \, \|I - D^{-*}D^{-1}\| \qquad \text{for all } i.$$

Proof. Reverse the rôles of A and D^*AD in Theorem 6.3.

This bound is stronger than the specialisation of Theorem 6.1 to Hermitian matrices because it also establishes a one-to-one correspondence between perturbed and exact eigenvalues.

7. Eigenvectors of Multiplicative Perturbations. We use the absolute perturbation bound in Corollary 4.3 to derive bounds on $\sin\theta$ in terms of a 'relative gap', i.e., the relative separation of $\hat{\lambda}$ from the remaining eigenvalues Λ_2. We define the relative gap via its inverse,

$$1/\text{relgap} \equiv |\hat{\lambda}|/\text{absgap} \ge 0,$$

to ensure that the definition is valid even when $\hat{\lambda} = 0$.

Since multiplicative perturbations lead to a residual for $z = D_2\hat{x}/\|D_2\hat{x}\|$ rather than for $\hat{x}$, a straightforward application of Corollary 4.3 results in a bound on the angle ϕ between z and $\text{range}(X_1)$ rather than the desired bound on θ. The following lemma shows how to adjust this bound.

LEMMA 7.1. *Let $0 \le \phi \le \pi/2$ be the largest principal angle between z and* $\text{range}(X_1)$. *Then*

$$\sin\theta \le \|D_2\hat{x}\| \, \sin\phi + \|D_2 - I\|,$$

and

$$\sin\theta \le \sin\phi + \|D_2 - I\|.$$

Proof. To derive the first bound write

$$P\hat{x} = PD_2\hat{x} \; + \; P(I - D_2)\hat{x} = \|D_2\hat{x}\| \, Pz \; + \; P(I - D_2)\hat{x},$$

and note that

$$\sin\theta = \|P\hat{x}\| \qquad \text{and} \qquad \sin\phi = \|Pz\|.$$

The second bound follows from the first if $\|D_2\hat{x}\| \le 1$, so assume that $\|D_2\hat{x}\| > 1$. Then

$$\sin\theta = \|P\hat{x}\| \le \|Pz\| \;+\; \|P(z-\hat{x})\| \le \sin\phi \;+\; \|z-\hat{x}\|.$$

Note that

$$\begin{aligned}\|(D_2 - I)\hat{x}\|^2 &= \|(D_2\hat{x} - z) + (z - \hat{x})\|^2 \\ &= \|z-\hat{x}\|^2 + \|D_2\hat{x} - z\|^2 + 2\,\mathrm{Re}\,(D_2\hat{x}-z)^*(z-\hat{x}).\end{aligned}$$

Since $\hat{x}$ and z are unit vectors,

$$2\,\mathrm{Re}\,(D_2\hat{x}-z)^*(z-\hat{x}) = 2\,(\|D_2\hat{x}\| - 1)\,\mathrm{Re}\,(1 - z^*\hat{x}) \ge 0.$$

Thus

$$\|(D_2 - I)\hat{x}\|^2 \ge \|z-\hat{x}\|^2,$$

and the result follows from the inequality $\|(D_2 - I)\hat{x}\| \le \|D_2 - I\|$. $\square$

In general any bound on $\sin\theta$ derived from Lemma 7.1 must consist of two summands. The theorem below bounds the first summand by the (absolute or relative) deviation of D_1 and D_2 from a similarity transformation, amplified by κ_2 and by the relative eigenvalue separation; and bounds the second summand by the (absolute and relative) deviation of the similarity transformation from the identity.

THEOREM 7.2. *Let*

$$\alpha_1 \equiv \|D_1^{-1} - D_2\|, \qquad \alpha_2 \equiv \|I - D_1^{-1}D_2^{-1}\|, \qquad \textit{and} \qquad \kappa_2 \equiv \kappa(Y_2).$$

Then

$$\sin\theta \le \kappa_2\,\frac{\min\{\alpha_1,\ \alpha_2\}}{\mathrm{relgap}} \;+\; \|I - D_2\|,$$

where $\alpha_1 \le \alpha_2$ *if* $\|D_2^{-1}\| \le 1$ *and* $\alpha_2 \le \alpha_1$ *if* $\|D_2\| \le 1$.

Proof. Apply Corollary 4.3 to the residual f for $\hat{\lambda}$ and z to get

$$\sin\phi \le \kappa_2\,\|f\|/\mathrm{absgap}.$$

To derive the bound with $\alpha_1 \le \alpha_2$, bound $\|f\|$ by

$$\|f\| \le |\hat{\lambda}|\,\|D_1^{-1} - D_2\|/\|D_2\hat{x}\| = |\hat{\lambda}|\,\alpha_1/\|D_2\hat{x}\|$$

and use the first bound in Lemma 7.1. To derive the bound with $\alpha_2 \le \alpha_1$, bound $\|f\|$ by

$$\|f\| \le |\hat{\lambda}|\,\|I - D_1^{-1}D_2^{-1}\| = |\hat{\lambda}|\,\alpha_2$$

and use the second bound in Lemma 7.1. Finally, $\|D_2\| \le 1$ implies

$$\alpha_1 = \|D_2 - D_1^{-1}\| \le \|I - D_1^{-1}D_2^{-1}\;\|D_2\|\| \le \alpha_2,$$

while $\|D_2^{-1}\| \le 1$ implies

$$\alpha_2 = \|I - D_1^{-1}D_2^{-1}\| \le \|D_2 - D_1^{-1}\|\;\|D_2^{-1}\| \le \alpha_1. \qquad \square$$

In some cases the first summand can be omitted.

COROLLARY 7.3. *If $D_1 = D_2^{-1}$ or $\hat\lambda = 0$, then*

$$\sin\theta = \|P(I-D_2)\hat x\| \le \|I-D_2\|.$$

Proof. Consider the case $D_1 = D_2^{-1}$. Then $D_1AD_2\hat x = \hat\lambda\hat x$ implies $AD_2\hat x = \hat\lambda D_2\hat x$, i.e., $\hat\lambda$ and $D_2\hat x$ are an exact eigenpair of A. According to the partitioning defined in §3, we must have $\Lambda_1 = \hat\lambda I$, so $D_2\hat x$ must be in $\text{range}(X_1) = \text{range}(Y_2)^-$. Thus $PD_2\hat x = 0$ and we have

$$\sin\theta = \|P\hat x\| = \|P(D_2\hat x - \hat x)\| \le \|I-D_2\|.$$

Next consider the case $\hat\lambda = 0$. Then $D_1AD_2\hat x = 0\hat x$ implies that $D_2^{-1}AD_2\hat x = 0\,\hat x$, since D_1 and D_2 are non-singular. Hence $\hat\lambda$ and $\hat x$ are an exact eigenpair of a similarity transformation of A, and we are back to the first case. ☐

For similarity transformations $D_1 = D_2^{-1}$, the bounds in Corollary 7.3 are essentially tight. The error angle is bounded by the relative deviation of D_2 from identity, without any amplification by κ_2 or by a relative gap. Therefore we can say that the eigenvectors of diagonalisable matrices are well-conditioned when the perturbation is caused by a similarity transformation.

Moreover, when $\hat\lambda = 0$, the bounds in Corollary 7.3 are again essentially tight. This means that the null vectors of diagonalisable matrices are also well-conditioned when the perturbation is multiplicative.

Related Work. Eisenstat and Ipsen [19, 18] bound $\sin\theta$ for Hermitian matrices. The result below corresponds to a specialisation of Theorem 7.2.

COROLLARY 7.4 (EISENSTAT AND IPSEN [18, THEOREM 2.1]). *Let A be Hermitian with λ_i in increasing order, let D be non-singular, and let $\hat\lambda_i$ be the eigenvalues of D^*AD in increasing order. Also let $0 \le \theta_i \le \pi/2$ be the largest principal angle between an eigenvector associated with $\hat\lambda_i$ and the eigenspace associated with λ_i. Then*

$$\sin\theta_i \le \frac{\|I-D^{-*}D^{-1}\|}{\text{relgap}_i} \;+\; \|I-D\|,$$

where

$$\frac{1}{\text{relgap}_i} \equiv \frac{|\hat\lambda_i|}{\min_{\lambda_j\ne\lambda_i}|\lambda_j-\hat\lambda_i|} = \max_{\lambda_j\ne\lambda_i}\frac{|\hat\lambda_i|}{|\lambda_j-\hat\lambda_i|}.$$

When the perturbed eigenspace $\hat X$ and the exact eigenspace X_1 have the same dimension $k \ge 1$, a similar bound holds for the largest principal angle $\|\sin\Theta\|$ between the two spaces.

THEOREM 7.5 (EISENSTAT AND IPSEN [18, VARIANT OF THEOREM 3.1]). *Let A be Hermitian, let D be non-singular, and let the k columns of $\hat X$ be an orthonormal basis for the eigenspace associated with the eigenvalues $\hat\lambda_1 \le \cdots \le \hat\lambda_k$ of D^*AD. Let the eigenvalues of A be ordered such that λ_i is the l_ith eigenvalue of A whenever $\hat\lambda_i$ is the l_ith eigenvalue of D^*AD for $1 \le i \le k$. Then*

$$\|\sin\Theta\| \le \sqrt{k}\,\left(\frac{\|I-D^{-*}D^{-1}\|}{\text{relgap}^*} \;+\; \|I-D\|\right),$$

where

$$\sin\Theta \equiv P\hat{X},$$

and

$$\frac{1}{\mathrm{relgap}^*} \equiv \max_{1\le i\le k}\max_{j>k}\frac{|\hat\lambda_i|}{|\lambda_j-\hat\lambda_i|}.$$

Li [25] derives the following bound on eigenspace perturbations for Hermitian matrices.

THEOREM 7.6 (LI [25, THEOREM 3.1]). *Under the assumptions of Theorem 7.5,*

$$\|\sin\Theta\|_F \le \frac{\|(D^*-D^{-1})X_1\|_F}{\mathrm{relgap}'} + \|(I-D^*)X_1\|_F,$$

where

$$\frac{1}{\mathrm{relgap}'} \equiv \max_{1\le i\le k}\max_{j>k}\frac{|\hat\lambda_j|}{|\lambda_i-\hat\lambda_j|}.$$

It is not clear what the relation between Theorems 7.5 or 7.6 is, and whether one is stronger than the other.

Acknowledgement. We thank Axel Ruhe and Krešimir Veselić for helpful discussions, and Ren-Cang Li and the referees for suggestions that improved the paper.

REFERENCES

[1] J. BARLOW AND J. W. DEMMEL, *Computing accurate eigensystems of scaled diagonally dominant matrices*, SIAM J. Numer. Anal., 27 (1990), pp. 762–91.
[2] F. BAUER AND C. FIKE, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–41.
[3] R. BHATIA, C. DAVIS, AND P. KOOSIS, *An extremal problem in Fourier analysis with applications in operator theory*, J. Funct. Anal., 82 (1989), pp. 138–50.
[4] R. BHATIA, C. DAVIS, AND A. MCINTOSH, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl., 52/53 (1983), pp. 45–57.
[5] Z. BOHTE, *A posteriori error bounds for eigensystems of matrices*, in Numerical Methods and Approximation Theory III, University of Niš, Yugoslavia, 1988, pp. 45–67.
[6] ———, *Computable error bounds for approximate eigensystems of matrices*, in VII Conference on Applied Mathematics, R. Scitovski, ed., University of Osijek, Croatia, 1990, pp. 29–37.
[7] F. CHATELIN, *Valeurs Propres de Matrices*, Masson, Paris, 1986.
[8] K. CHU, *Generalization of the Bauer-Fike theorem*, Numer. Math., 49 (1986), pp. 685–91.
[9] C. DAVIS AND W. KAHAN, *Some new bounds on perturbation of subspaces*, Bull. Amer. Math. Soc., 75 (1969), pp. 863–8.
[10] ———, *The rotation of eigenvectors by a perturbation, III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.
[11] P. DEIFT, J. W. DEMMEL, L.-C. LI, AND C. TOMEI, *The bidiagonal singular value decomposition and Hamiltonian mechanics*, SIAM J. Numer. Anal., 28 (1991), pp. 1463–516.
[12] J. DEMMEL AND W. KAHAN, *Accurate singular values of bidiagonal matrices*, SIAM J. Sci. Stat. Comput., 11 (1990), pp. 873–912.
[13] J. DEMMEL AND K. VESELIĆ, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–45.
[14] J. W. DEMMEL AND W. GRAGG, *On computing accurate singular values and eigenvalues of matrices with acyclic graphs*, Lin. Alg. Applics., 185 (1993), pp. 203–17.
[15] Z. DRMAČ, *Computing the Singular and the Generalized Singular Values*, PhD thesis, Fachbereich Mathematik, Fernuniversität Gesamthochschule Hagen, Germany, 1994.

[16] ———, *On relative residual bounds for the eigenvalues of a Hermitian matrix*, Linear Algebra Appl., 244 (1996), pp. 155–64.

[17] Z. Drmač and V. Hari, *Relative a posteriori residual bounds for the eigenvalues of a Hermitian matrix*, tech. rep., Department of Mathematics, University of Hagen, Germany, 1994.

[18] S. Eisenstat and I. Ipsen, *Relative perturbation bounds for eigenspaces and singular vector subspaces*, in Applied Linear Algebra, SIAM, 1994, pp. 62–5.

[19] ———, *Relative perturbation techniques for singular value problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1972–88.

[20] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins Press, Baltimore, second ed., 1989.

[21] W. Govaerts and J. Pryce, *Block elimination with one refinement solves bordered linear systems accurately*, BIT, 30 (1990), pp. 490–507.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[23] W. Kahan, *Accurate eigenvalues of a symmetric tridiagonal matrix*, Technical Report No. CS-41, Department of Computer Science, Stanford University, July 1966 (revised June 1968).

[24] R. Li, *Relative perturbation theory: (I), eigenvalue variations*, LAPACK working note 84, revised, Computer Science Department, University of Tennessee, Knoxville, January 1996.

[25] ———, *Relative perturbation theory: (II), eigenspace variations*, LAPACK working note 85, revised, Computer Science Department, University of Tennessee, Knoxville, April 1996.

[26] R. Mathias, *Spectral perturbation bounds for graded positive definite matrices*, technical report, Department of Mathematics, College of William and Mary, Williamsburg, VA, May 1994.

[27] J. Ortega, *Numerical Analysis, A Second Course*, Academic Press, New York, 1972.

[28] A. Ostrowski, *A quantitative formulation of Sylvester's law of inertia*, Proc. Nat. Acad,. Sci., 45 (1959), pp. 740–4.

[29] E. Pietzsch, *Genaue Eigenwertberechnung Nichtsingulärer Schiefsymmetrischer Matrizen*, PhD thesis, Fachbereich Mathematik, Fernuniversität Gesamthochschule Hagen, Germany, 1993.

[30] J. Pryce, *Multiplicative error analysis of matrix transformation algorithms*, IMA Journal of Numerical Analysis, 5 (1985), pp. 437–45.

[31] A. Ruhe, *Perturbation bounds for means of eigenvalues and invariant subspaces*, Bit, 10 (1970), pp. 343–54.

[32] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, New York, 1992.

[33] I. Slapničar, *Accurate Symmetric Eigenreduction by a Jacobi Method*, PhD thesis, Fachbereich Mathematik, Fernuniversität Gesamthochschule Hagen, Germany, 1992.

[34] G. Stewart, *Error bounds for approximate invariant subspaces, of closed linear operators*, SIAM J. Numer. Anal., 8 (1971), pp. 796–808.

[35] ———, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review, 15 (1973), pp. 727–64.

[36] J. Varah, *Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned*, Math. Comp., 24 (1970), pp. 137–49.

[37] K. Veselić and I. Slapničar, *Floating-point perturbations of Hermitian matrices*, Linear Algebra Appl., 195 (1993), pp. 81–116.

[38] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.