Relative perturbation theory: IV. sin 2θ theorems[☆]

Ren-Cang Li

Department of Mathematics, University of Kentucky, 715 Patterson Office Tower, Lexington, KY 40506, USA

Abstract

The double angle theorems of Davis and Kahan bound the change in an invariant subspace when a Hermitian matrix A is subject to an additive perturbation $A \rightarrow \tilde{A} = A + \Delta A$. This paper supplies analogous results when A is subject to a congruential, or multiplicative, perturbation $A \rightarrow \tilde{A} = D^* A D$. The relative gaps that appear in the bounds involve the spectrum of only one matrix, either A or $\tilde{A}$, in contrast to the gaps that appear in the single angle bounds. The double angle theorems do not directly bound the difference between the old invariant subspace $\mathscr{S}$ and the new one $\tilde{\mathscr{S}}$ but instead bound the difference between $\tilde{\mathscr{S}}$ and its reflection $J\tilde{\mathscr{S}}$ where the mirror is $\mathscr{S}$ and J reverses $\mathscr{S}^\perp$, the orthogonal complement of $\mathscr{S}$. The double angle bounds are proportional to the departure from the identity and from orthogonality of the matrix $\tilde{D} \stackrel{\text{def}}{=} D^{-1} J D J$. Note that $\tilde{D}$ is invariant under the transformation $D \rightarrow D/\alpha$ for $\alpha \neq 0$, whereas the single angle theorems give bounds proportional to D's departure from the identity and from orthogonality. The corresponding results for the singular value problem when a (nonsquare) matrix B is perturbed to $\tilde{B} = D_1^* B D_2$ are also presented. © 2000 Elsevier Science Inc. All rights reserved.

1. Introduction

Eigenvalue and singular value computations to high relative accuracy have been attracting lots of attention over the last 10 years or so. Tremendous progress has

been made both in theoretical understanding and numerical algorithms, see [1,4,7–14,18,25–28] and references therein. On the algorithmic side there are Demmel–Kahan QR methods for bidiagonal singular value computations [8], (two-sided) Jacobi methods for the eigenvalue problems of positive definite matrices and for the singular value computations [9,25,28], bisection method for scaled diagonally dominant matrices [1] and for matrices with acyclic graphs [7,17], new implementations of the qd method [14,27], and Demmel's algorithms for structured matrices [6], and more recently [10] showed how to compute singular value decompositions (SVDs) to high relative accuracy for matrices that can be factored accurately as $B = X\Gamma Y^*$ where Γ is diagonal and X and Y are any well-conditioned matrices; on the theoretical side, analogous results to many celebrated theorems for absolute perturbations $A \rightarrow \tilde{A} = A + \Delta A$ are obtained for perturbations that are multiplicative $A \rightarrow \tilde{A} = D^*AE$ ($E = D$ when A is Hermitian) [12,13,16,18,20–22], though exceptions remain.

This paper presents analogues to the double angle theorems of Davis and Kahan [3] in the case of multiplicative perturbations. For one-dimensional eigenspace, Demmel [5, Theorem 5.7, p. 208] obtained an analogue, but his approach does not seem to be easily adaptable to eigenspaces of higher dimensions. Our new double angle theorems that work for eigenspaces of any arbitrary dimension have two advantages over the existing single angle theorems. Consider the Hermitian eigenvalue problem for A and $\tilde{A} = D^*AD$, where D is nonsingular. The first advantage, also presented in Davis and Kahan sin 2θ theorems, is that (relative) gaps are defined using exclusively eigenvalues of either A or $\tilde{A}$ but not both. We observe that if $D = \alpha I$, a multiple of the identity, A and $\tilde{A} = |\alpha|^2 A$ share the same eigenspaces, but the existing bounds, e.g., [21, Theorem 3.1], do not reflect this. In fact, as long as D is *close* to some multiple of the identity, the eigenspaces of A and $\tilde{A}$, when properly matched, are close. The new sin 2θ theorems provide upper bounds that are invariant under rescaling $D \rightarrow D/\alpha$ for $\alpha \neq 0$. This is the second advantage.

The rest of this paper is organized as follows. Section 2 derives relative sin 2θ theorems for the Hermitian eigenvalue problem. Section 3 develops relative sin 2θ theorems for singular value problem.

Notation. We shall follow the notation set forth in the first two parts of this series [20,21]. For convenience, we spell out some of them here. For *relative distances* we use, besides the *classical* one $|\alpha - \tilde{\alpha}|/|\alpha|$,

$$\varrho_p(\alpha, \tilde{\alpha}) = \frac{|\alpha - \tilde{\alpha}|}{\sqrt[p]{|\alpha|^p + |\tilde{\alpha}|^p}} \quad \text{for } 1 \leqslant p \leqslant \infty \text{ and } \chi(\alpha, \tilde{\alpha}) = \frac{|\alpha - \tilde{\alpha}|}{\sqrt{|\alpha\tilde{\alpha}|}},$$

with convention $0/0 = 0$ for convenience. It was proved in [20] that ϱ_p is indeed a metric on the set of real numbers and recently Barrlund [2] went further to show that it is a metric on the set of complex numbers, also. χ fails to satisfy the triangle inequality and thus is not a metric. Nevertheless all those relative distances are topologically equivalent [20], and thus for the purpose of bounding relative errors,

any relative metric is just as good as others. $\|X\|_2$ and $\|X\|_F$ denote the spectral and Frobenius norms of matrix X, respectively. $\lambda(X)$ is the set of the eigenvalues of X, and $\sigma(X)$ is the set of the singular values of X. X^* is the conjugate transpose. I_n denotes the $n \times n$ identity matrix (we may simply write I instead if no confusion).

2. Relative sin 2θ theorems for eigenspace variations

Let A and $\tilde{A}$ be two Hermitian matrices whose eigendecompositions are

$$A = (U_1 \, U_2) \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix} \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix},$$

$$\tilde{A} = (\tilde{U}_1 \, \tilde{U}_2) \begin{pmatrix} \tilde{\Lambda}_1 & \\ & \tilde{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \tilde{U}_1^* \\ \tilde{U}_2^* \end{pmatrix}, \tag{2.1}$$

where $(\overset{k}{U_1} \, \overset{n-k}{U_2})$ and $(\overset{k}{\tilde{U}_1} \, \overset{n-k}{\tilde{U}_2})$ are unitary, and

$$\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_k), \quad \Lambda_2 = \mathrm{diag}(\lambda_{k+1}, \ldots, \lambda_n), \tag{2.2}$$

$$\tilde{\Lambda}_1 = \mathrm{diag}(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_k), \quad \tilde{\Lambda}_2 = \mathrm{diag}(\tilde{\lambda}_{k+1}, \ldots, \tilde{\lambda}_n). \tag{2.3}$$

We shall treat $\tilde{A}$ as a perturbed matrix of A, and derive bounds on the changes in subspace $\mathscr{S} \overset{\text{def}}{=} \mathrm{span}(U_1)$, A's invariant subspace spanned by U_1's columns. We do this by bounding the sines of the double canonical angles between $\mathscr{S}$ and $\tilde{\mathscr{S}} \overset{\text{def}}{=} \mathrm{span}(\tilde{U}_1)$. Define

$$J \overset{\text{def}}{=} (U_1 \, U_2) \begin{pmatrix} I_k & \\ & -I_{n-k} \end{pmatrix} \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix}, \quad \hat{A} \overset{\text{def}}{=} J\tilde{A}J. \tag{2.4}$$

The matrix J was implicitly used but not explicitly formed by Davis and Kahan [3] in deriving double angle theorems. It can be verified that

$$J^* = J, \quad J^2 = I_n, \quad J^{-1} = J, \quad JAJ = A.$$

So, J is unitary, and $\tilde{A}$ and $\hat{A}$ are unitarily similar and thus have the same eigenvalues. In fact, a complete eigendecomposition of $\hat{A}$ is

$$\hat{A} = (\hat{U}_1 \, \hat{U}_2) \begin{pmatrix} \tilde{\Lambda}_1 & \\ & \tilde{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \hat{U}_1^* \\ \hat{U}_2^* \end{pmatrix}, \tag{2.5}$$

where $\hat{U}_i = J\tilde{U}_i$ for $i = 1, 2$. Geometrically,[1] $J\tilde{\mathscr{S}}$ is a reflection of $\tilde{\mathscr{S}}$, where the mirror for J is $\mathscr{S}$ and J reverses $\mathscr{S}^\perp$, the orthogonal complement of $\mathscr{S}$. This explains the following lemma that relates $\Theta(\tilde{U}_1 \, \hat{U}_1)$ to $\Theta(U_1 \, \tilde{U}_1)$.

---

[1] I am grateful to Professor B.N. Parlett who pointed out this geometric interpretation to me.

Lemma 2.1 (Davis–Kahan [3]). *We have* $\sigma(\sin \Theta(\tilde{U}_1, \hat{U}_1)) = \sigma(\sin 2\Theta(U_1, \tilde{U}_1))$.
This is equivalent to say that for all unitarily invariant norms $||| \cdot |||$

$$||| \sin \Theta(\tilde{U}_1, \hat{U}_1)||| = ||| \sin 2\Theta(U_1, \tilde{U}_1)|||.$$

Proof. It is essential in [3], and more explicitly embedded in the proofs in [19,29] with the help of Van Loan [30, Theorem 2]. $\square$

Write $\tilde{A} = A + H$, then $\hat{A} = A + JHJ$, one of which will be considered as a perturbed one of the other. What we have so far is due to Davis and Kahan [3] who then continued to combine $\sin\theta$ theorems already proved and that $\hat{A} - \tilde{A} = JHJ - H$ easily bounded in term of norms of H. But such a combination does not work for us. We need to interpret the change from $\tilde{A}$ to $\hat{A}$ as caused by some multiplicative perturbation that is close to the identity. Although $\hat{A} = J\tilde{A}J$ by definition, this J, as a multiplicative perturbation, is too far away from a multiple of the identity αI since

$$\|J - \alpha I\|_2 = \max\{|1 - \alpha|, |1 + \alpha|\} \geqslant 1$$

always for any α unless $k = n$. So we have to do something different.

2.1. Multiplicatively perturbed A to $\tilde{A} = D^* A D$

Notice $JAJ = A$ and thus

$$\hat{A} = J\tilde{A}J = JD^*ADJ = J^*D^*JAJDJ = J^*D^*JD^{-*}\tilde{A}D^{-1}JDJ.$$

Therefore

$$\hat{A} = \tilde{D}^*\tilde{A}\tilde{D}, \quad \tilde{D} \stackrel{\text{def}}{=} D^{-1}JDJ. \tag{2.6}$$

This $\tilde{D}$ is close to the identity if D is close to some multiple of the identity, and $\tilde{D}$ is close to some unitary matrix if D is close to some multiple of a unitary matrix. We shall return to this later in this section. (2.6) is the key to our success. With it and Lemma 2.1, double angle theorems follow from the existing single angle theorems. To keep this paper fairly short, we provide a detailed account of only one double angle theorem with full discussion while briefly stating others.

The following single angle theorem is in [21, Theorem 3.1], where the subscript in η_c is an indication of it being defined with the *c*lassical relative measurement.

Theorem 2.1 (Li [21]). *Let A and* $\tilde{A} = D^*AD$ *be two* $n \times n$ *Hermitian matrices with eigendecompositions* (2.1)–(2.3), *where D is nonsingular. If* $\lambda(\Lambda_1) \cap \lambda(\tilde{\Lambda}_2) = \emptyset$, *then*

$$|| \sin \Theta(U_1, \tilde{U}_1)||_{\mathrm{F}} \leqslant \frac{\sqrt{\|(I - D^{-1})U_1\|_{\mathrm{F}}^2 + \|(I - D^*)U_1\|_{\mathrm{F}}^2}}{\eta_2}, \tag{2.7}$$

$$|| \sin \Theta(U_1, \tilde{U}_1)||_F \leqslant \|(I - D^*)U_1\|_F + \frac{\|(D^* - D^{-1})U_1\|_F}{\eta_c}, \tag{2.8}$$

where

$$\eta_2 \overset{\text{def}}{=} \min_{\mu \in \lambda(\Lambda_1), \tilde{\mu} \in \lambda(\tilde{\Lambda}_2)} \varrho_2(\mu, \tilde{\mu}) \quad and \quad \eta_c \overset{def}{=} \min_{\mu \in \lambda(\Lambda_1), \tilde{\mu} \in \lambda(\tilde{\Lambda}_2)} \frac{|\mu - \tilde{\mu}|}{|\tilde{\mu}|}.$$

Our first double angle theorem is a consequence of Theorem 2.1.

Theorem 2.2. *Let A and $\tilde{A} = D^*AD$ be two $n \times n$ Hermitian matrices with eigendecompositions* (2.1)–(2.3), *where D is nonsingular. If $\lambda(\tilde{\Lambda}_1) \cap \lambda(\tilde{\Lambda}_2) = \emptyset$, then*

$$|| \sin 2\Theta(U_1, \tilde{U}_1)||_F \leqslant \frac{\sqrt{\|(I - \tilde{D}^{-1})\tilde{U}_1\|_F^2 + \|(I - \tilde{D}^*)\tilde{U}_1\|_F^2}}{\tilde{\eta}_2}, \tag{2.9}$$

$$|| \sin 2\Theta(U_1, \tilde{U}_1)||_F \leqslant \|(I - \tilde{D}^*)\tilde{U}_1\|_F + \frac{\|(\tilde{D}^* - \tilde{D}^{-1})\tilde{U}_1\|_F}{\tilde{\eta}_c}, \tag{2.10}$$

where $\tilde{D}$ is defined in (2.6),

$$\tilde{\eta}_2 \overset{\text{def}}{=} \min_{\mu \in \lambda(\tilde{\Lambda}_1), \nu \in \lambda(\tilde{\Lambda}_2)} \varrho_2(\mu, \nu), \qquad \tilde{\eta}_c \overset{\text{def}}{=} \min_{\mu \in \lambda(\tilde{\Lambda}_1), \nu \in \lambda(\tilde{\Lambda}_2)} \frac{|\mu - \nu|}{|\nu|}.$$

Proof. Bear (2.1) and (2.5) in mind, and then apply Theorem 2.1 to $\tilde{A}$ and $\hat{A}$ to get

$$\| \sin \Theta(\tilde{U}_1, \hat{U}_1)\|_F \leqslant \text{[RHS of (2.9)]},$$
$$\| \sin \Theta(\tilde{U}_1, \hat{U}_1)\|_F \leqslant \text{[RHS of (2.10)]}.$$

Combining them with Lemma 2.1 completes the proof.   □

Remark 2.1. As is noted in [21, footnote 3, p. 478], a bound slightly different from (2.10) is

$$|| \sin 2\Theta(U_1, \tilde{U}_1)||_F \leqslant ||(I - \tilde{D}^{-1})\tilde{U}_1||_F + \frac{\|(\tilde{D}^* - \tilde{D}^{-1})\tilde{U}_1\|_F}{\tilde{\eta}_c'}, \tag{2.11}$$

where $\tilde{\eta}_c' \overset{\text{def}}{=} \min_{\mu \in \lambda(\tilde{\Lambda}_1), \nu \in \lambda(\tilde{\Lambda}_2)} |\mu - \nu|/|\mu|$. This and those relative gaps in Theorem 2.2 are defined in terms of eigenvalues of $\tilde{A}$ only, in contrast to Theorem 2.1 and other theorems in [12,21] which use gaps defined in terms of eigenvalues of both A and $\tilde{A}$. This feature coincides with Davis–Kahan $\sin 2\theta$ theorems.

The upper bounds in Theorem 2.2 have an interesting invariant property that the existing single angle theorems, e.g., Theorem 2.1, lack. Notice that as long as D is

close to some multiple of the identity, the eigenspaces of A and $\tilde{A}$, when properly matched, are close, but existing bounds do not yield small error bounds in this case. The latter can be cured by considering A and $(D/\alpha)^*A(D/\alpha)$ for a judiciously chosen α to make D/α close to the identity, e.g., take $\alpha = \mathrm{e}^{\mathrm{i}\psi}\|D\|_2$ for some ψ or determine it by optimizing final bounds with α as a free parameter as Li and Stewart [24] did to the singular value problem. Even though this blemish is curable, it is still nice to have bounds like those in Theorem 2.2 that are automatically immune to the drawback since $\tilde{D} = D^{-1}JDJ \equiv J(D/\alpha)^{-1}J(D/\alpha)J$ for any $\alpha \neq 0$.

Next we show how to bound $I - \tilde{D}^{-1}$, $I - \tilde{D}^*$, and $\tilde{D}^* - \tilde{D}^{-1}$ in terms of the deviations of D from the identity or orthogonality (if necessary, D should be rescaled[2]). The following identities are easy to verify:

$$I - \tilde{D}^{-1} = J(I - D^{-1})J + JD^{-1}J(I - D) \tag{2.12}$$
$$= (I - D) + J(I - D^{-1})JD \tag{2.13}$$
$$= JD^{-1}(DJ - JD) \tag{2.14}$$
$$= (D^{-1}J - JD^{-1})JD \tag{2.15}$$
$$I - \tilde{D} = (I - D^{-1}) + D^{-1}J(I - D)J \tag{2.16}$$
$$= J(I - D)J + (I - D^{-1})JDJ \tag{2.17}$$
$$= D^{-1}J(JD - DJ) \tag{2.18}$$
$$= (JD^{-1} - D^{-1}J)DJ \tag{2.19}$$
$$\tilde{D}^* - \tilde{D}^{-1} = JD^*J(D^{-*} - D) + J(D^* - D^{-1})JD \tag{2.20}$$
$$= J(D^* - D^{-1})JD^{-*} + JD^{-1}J(D^{-*} - D). \tag{2.21}$$

An immediate consequence of (2.14) and (2.18) is that[3] $\sin 2\Theta(U_1, \tilde{U}_1) \equiv 0$ if $JD = DJ$. These identities make it possible to bound the right-hand sides of (2.9) and (2.10) in Theorem 2.2 and those of (2.26) and (2.27) in Theorem 2.3 below by norms of $I - D$, $I - D^{-1}$, or $D^* - D^{-1}$. We present here the following corollary as an example. The reader may derive some other variations depending on his/her needs.

Corollary 2.1. *Under the conditions and notation of Theorem* 2.2, *we have*

$$\frac{1}{2}\|\sin 2\Theta(U_1, \tilde{U}_1)\|_F \leqslant \frac{\sqrt{\|I - D\|_F^2 + \|D\|_2^2\|I - D^{-1}\|_F^2}}{\tilde{\eta}_2}, \tag{2.22}$$

$$\frac{1}{2}\|\sin 2\Theta(U_1, \tilde{U}_1)\|_F \leqslant \|D^{-1}\|_2\|I - D\|_F + \frac{\|D\|_2\|D^* - D^{-1}\|_F}{\tilde{\eta}_c}. \tag{2.23}$$

---

[2] For example, instead of (2.12) we would use $I - \tilde{D}^{-1} = J(I - \alpha D^{-1})J + J\alpha D^{-1}J(I - D/\alpha)$.

[3] In fact if $JD = DJ$, then $D = U \operatorname{diag}(D_1, D_2)U^*$ where D_1 is $k \times k$, and then

$$D^*AD = U \operatorname{diag}(D_1^*\Lambda_1 D_1, D_2^*\Lambda_2 D_2)U^*,$$

so U_1 and $\tilde{U}_1$ span the same subspace.

Proof. Use the above identities to get

$$\|(I - \tilde{D}^{-1})\tilde{U}_1\|_{\mathrm{F}}^2 \leqslant \|I - \tilde{D}^{-1}\|_{\mathrm{F}}^2$$

$$\leqslant \left(\|I - D\|_{\mathrm{F}} + \|J(I - D^{-1})JD\|_{\mathrm{F}}\right)^2 \qquad \text{(by (2.13))}$$

$$\leqslant 2\|I - D\|_{\mathrm{F}}^2 + 2\|D\|_2^2\|I - D^{-1}\|_{\mathrm{F}}^2,$$

$$\|(I - \tilde{D}^*)\tilde{U}_1\|_{\mathrm{F}}^2 \leqslant \|I - \tilde{D}\|_{\mathrm{F}}^2$$

$$\leqslant \left(\|J(I - D)J\|_{\mathrm{F}} + \|(I - D^{-1})JDJ\|_{\mathrm{F}}\right)^2 \qquad \text{(by (2.17))}$$

$$\leqslant 2\|I - D\|_{\mathrm{F}}^2 + 2\|D\|_2^2\|I - D^{-1}\|_{\mathrm{F}}^2.$$

Inequality (2.22) now follows from (2.9). To derive (2.23) from (2.10), we observe from (2.16) that $I - \tilde{D} = D^{-1}(D - I) + D^{-1}J(I - D)J$ and thus

$$\|(I - \tilde{D}^*)\tilde{U}_1\|_{\mathrm{F}} \leqslant 2\|D^{-1}\|_2\|I - D\|_{\mathrm{F}},$$

$$\|(\tilde{D}^* - \tilde{D}^{-1})\tilde{U}_1\|_{\mathrm{F}} \leqslant \|JD^*J(D^{-*} - D) + J(D^* - D^{-1})JD\|_{\mathrm{F}} \quad \text{(by (2.20))}$$

$$\leqslant 2\|D\|_2\|D^* - D^{-1}\|_{\mathrm{F}}. \qquad \square$$

Remark 2.2. In deriving Corollary 2.1, we have traded some sharpness for the comparative simplicity in (2.22) and (2.23), as can be seen from our proof. For example when $k = 1$, all the $\|\cdot\|_{\mathrm{F}}$'s in Theorem 2.2, including the $\||\cdot\||$'s in Theorem 2.3, are effectively $\|\cdot\|_2$, and consequently all the $\|\cdot\|_{\mathrm{F}}$'s in this corollary can be replaced by $\|\cdot\|_2$'s. This presents an improvement since $\|\cdot\|_2 \leqslant \|\cdot\|_{\mathrm{F}}$ always. There is a way to deal with this sudden discontinuous jump by introducing a norm $\|\cdot\|_{2,\ell}$ defined as

$$\|X\|_{2,\ell} \stackrel{\text{def}}{=} \sqrt{\sum_{j=1}^{\ell} [\sigma_j(X)]^2},$$

where $\sigma_1(X) \geqslant \sigma_2(X) \geqslant \cdots$ are the singular values of X. $\|X\|_{2,\ell}$ called *Ky Fan 2-l norms* [15, Problem 3, p. 199] are unitarily invariant. For an $m \times n$ matrix X, $\|X\|_{2,1} = \|X\|_2$ and $\|X\|_{2,m} = \|X\|_{2,n} = \|X\|_{\mathrm{F}}$. It can be seen that *all inequalities in this paper and in Part* II [21] *of this series are still valid with all the* $\|\cdot\|_{\mathrm{F}}$'s *replaced by* $\|\cdot\|_{2,k}$. Analogous claim holds for inequalities involving a general unitarily invariant norm, but we shall not dwell on this too much.

For $k = 1$, Demmel [5, Theorem 5.7, p. 208] proved

$$\frac{1}{2}\sin 2\Theta(U_1, \tilde{U}_1) \leqslant \frac{\epsilon_1}{1 - \epsilon_1} \cdot \frac{1}{\tilde{\eta}_{\mathrm{c}}'} + \epsilon_2, \tag{2.24}$$

where $\epsilon_1 = \|I - D^{-*}D^{-1}\|_2$, $\epsilon_2 = \|I - D\|_2$ and $\tilde{\eta}'_c$ is defined as in Remark 2.1. For D close to the identity (2.24) and ours are comparable. Bearing in mind the argument in Remark 2.2, analogously to Corollary 2.1, we get for $k = 1$

$$\frac{1}{2}\|\sin 2\Theta(U_1, \tilde{U}_1)\|_2 \leqslant \|D^{-1}\|_2 \|I - D\|_2 + \frac{\|D\|_2 \|D^* - D^{-1}\|_2}{\tilde{\eta}'_c}. \quad (2.25)$$

It can be proved that the ratio of the right-hand sides of (2.24) and (2.25) is $1 + \mathcal{O}(\|I - D\|_2)$. Demmel's (2.24) does not enjoy the invariant property with respect to scaling $D \to D/\alpha$.

The next theorem provides bounds for all unitarily invariant norms at the price of a more severe restriction (as in Fig. 1) on how $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ are separated, as in the $\sin\theta$ theorems for all unitarily invariant norms in [21].

Theorem 2.3. *Let A and $\tilde{A} = D^* A D$ be two $n \times n$ Hermitian matrices with eigendecompositions* (2.1)–(2.3), *where D is nonsingular. Assume that the spectra of $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ distribute as in Fig.* 1. *Then for any unitarily invariant norm $|||\cdot|||$*

$$|||\sin 2\Theta(U_1, \tilde{U}_1)||| \leqslant \frac{\sqrt[q]{|||(I - \tilde{D}^{-1})\tilde{U}_1|||^q + |||(I - \tilde{D}^*)\tilde{U}_1|||^q}}{\underline{\tilde{\eta}}_p}, \quad (2.26)$$

$$|||\sin 2\Theta(U_1, \tilde{U}_1)||| \leqslant |||(I - \tilde{D}^*)\tilde{U}_1||| + \frac{|||(\tilde{D}^* - \tilde{D}^{-1})\tilde{U}_1|||}{\underline{\tilde{\eta}}_c}, \quad (2.27)$$

where $\tilde{D}$ is defined in (2.6), *q is defined by $1/p + 1/q = 1$, and*

$$\underline{\tilde{\eta}}_p \overset{\text{def}}{=} \varrho_p(\alpha, \alpha + \delta), \quad \underline{\tilde{\eta}}_c \overset{\text{def}}{=} \begin{cases} \delta/(\alpha + \delta) & \text{if Fig. 1(a)}, \\ \delta/\alpha & \text{if Fig. 1(b)}. \end{cases}$$

Proof. It is a consequence of Li [21, Theorem 3.2] applied to $\tilde{A}$ and $\hat{A}$ and Lemma 2.1. $\square$

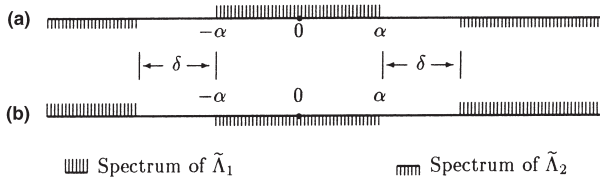Theorem 2.3 has a corollary similar to Corollary 2.1.

Fig. 1. The spectrum of $\tilde{\Lambda}_1$ and that of $\tilde{\Lambda}_2$ are separated by two intervals, and one of the spectra scatters around the origin.

2.2. Nonnegative-definite matrices scalably well-conditioned

In what follows we show how the previous ideas can be applied to a more realistic situation when A can be scaled to improve its condition number. Consider $n \times n$ *nonnegative-definite* Hermitian matrix $A = S^* H S$ which is perturbed in a special way to $\tilde{A} = S^* \tilde{H} S$, where S is a scaling matrix and usually diagonal. But this is not necessary to the theorems below. The elements of S can vary wildly. H is nonsingular and usually better-conditioned than A itself. Set $\Delta H \stackrel{\text{def}}{=} \tilde{H} - H$. As in [21, pp. 481–482], we have

$$ A = BB^*, \quad \tilde{A} = \tilde{B}\tilde{B}^*, $$

where

$$ B = S^* H^{1/2}, \quad \tilde{B} = BD, \quad D = \left(I + H^{-1/2}(\Delta H)H^{-1/2} \right)^{1/2}. \tag{2.28} $$

Given the eigendecompositions of A and $\tilde{A}$ as in (2.1)–(2.3), we define J and $\hat{A}$ as in (2.4). Set

$$ Q = B^{-1} J B \quad \text{a unitary matrix}, $$

since $QQ^* = B^{-1} J BB^* J B^{-*} = B^{-1} BB^* B^{-*} = I$, where we have used $JAJ = A$. Use $JB = BQ$ to get

$$ \hat{A} = J\tilde{A}J = JBDD^*B^*J $$

$$ = BQDD^*Q^*B^* = BDD^{-1}QDD^*Q^*D^{-*}D^*B^* = (\tilde{B}\tilde{D})(\tilde{B}\tilde{D})^*, $$

where

$$ \tilde{D} = D^{-1}QD. \tag{2.29} $$

$\tilde{D}$ is nearly unitary if D is and D depends on H, not A. The proof outlined in [21, pp. 481–482] for [21, Theorems 3.3 and 3.4] and Lemma 2.1 yield the following theorems.

Theorem 2.4. *Let $A = S^* H S$ and $\tilde{A} = S^* \tilde{H} S$ be two $n \times n$ Hermitian matrices with eigendecompositions (2.1)–(2.3). H is positive definite and $\|H^{-1}\|_2 \|\Delta H\|_2 < 1$. If $\tilde{\eta}_\chi \stackrel{\text{def}}{=} \min\limits_{\mu \in \lambda(\tilde{A}_1), \nu \in \lambda(\tilde{A}_2)} \chi(\mu, \nu) > 0$, then*

$$ \|\sin 2\Theta(U_1, \tilde{U}_1)\|_{\mathrm{F}} \leqslant \frac{\|\tilde{D}^* - \tilde{D}^{-1}\|_{\mathrm{F}}}{\tilde{\eta}_\chi}, \tag{2.30} $$

where D and $\tilde{D}$ are as in (2.28) and (2.29).

Theorem 2.5. *Let $A = S^* H S$ and $\tilde{A} = S^* \tilde{H} S$ be two $n \times n$ Hermitian matrices with eigendecompositions (2.1)–(2.3). H is positive definite and $\|H^{-1}\|_2 \|\Delta H\|_2 <$*

1. *Assume that the spectra $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ distribute as in Fig. 1. Then for any unitarily invariant norm $||| \cdot |||$*

$$||| \sin 2\Theta(U_1, \tilde{U}_1)||| \leqslant \frac{|||\tilde{D}^* - \tilde{D}^{-1}|||}{\underline{\tilde{\eta}}_\chi}, \tag{2.31}$$

where $\underline{\tilde{\eta}}_\chi \overset{\text{def}}{=} \chi(\alpha, \alpha + \delta)$, and D and $\tilde{D}$ are as in (2.28) and (2.29).

Using the technique of Li [23], we can even obtain a bound in any unitarily invariant norm on $\sin 2\Theta(U_1, \tilde{U}_1)$ under the conditions of Theorem 2.4, in contrast to the stronger conditions of Theorem 2.5.

We now show how to bound the right-hand sides of (2.30) and (2.31) in terms of $D - D^{-1}$ and ΔH, instead of $\tilde{D}$. Notice that $D^* = D$. We observe

$$\tilde{D}^* - \tilde{D}^{-1} = D^{-1}(D^2 Q^* - Q^* D^2)D^{-1} \tag{2.32}$$

$$= (D - D^{-1})Q^* D^{-1} + D^{-1}Q^*(D^{-1} - D) \tag{2.33}$$

$$= DQ^*(D^{-1} - D) + (D - D^{-1})Q^* D. \tag{2.34}$$

An immediate consequence of (2.32) is that[4] $\sin 2\Theta(U_1, \tilde{U}_1) \equiv 0$ if $D^2 Q = QD^2$.

Corollary 2.2. *Under the conditions and notation of Theorem* 2.4, *we have*

$$\frac{1}{2}\| \sin 2\Theta(U, \tilde{U}_1)\|_{\mathrm{F}} \leqslant \frac{\|D\|_2 \|D - D^{-1}\|_{\mathrm{F}}}{\tilde{\eta}_\chi}, \tag{2.35}$$

$$\leqslant \frac{\sqrt{1 + \|H^{-1}\|_2 \Delta H\|_2}}{\sqrt{1 - \|H^{-1}\|_2 \Delta H\|_2}} \frac{\|H^{-1}\|_2 \|\Delta H\|_{\mathrm{F}}}{\eta_\chi}. \tag{2.36}$$

Proof. (2.30) yields (2.35), and (2.36) follows from (2.35) and the bound on $D - D^{-1}$ in [21, p. 482]. □

We note in passing that (2.36) is still valid with $\|H^{-1}\|_2 \|\Delta H\|_p$ replaced by

$$\|H^{-1/2}(\Delta H)H^{-1/2}\|_p \leqslant \|H^{-1}(\Delta H)\|_p, \quad p = 2, \mathrm{F},$$

see [21, p. 482] for details. Theorem 2.5 has a similar corollary.

Remark 2.3. Our approach may be extended to diagonalizable matrices. Suppose that both A and $\tilde{A}$ are diagonalizable and let

---

[4] In fact if $D^2 Q = QD^2$, then $BD^2B^{-1}J = JBD^2B^{-1}$, thus $BD^2B^{-1} = U \operatorname{diag}(D_1, D_2)U^*$ where D_1 is $k \times k$, and then

$$\tilde{A} = BD^2B^* = BD^2B^{-1}BB^* = U \operatorname{diag}(D_1\Lambda_1, D_2\Lambda_2)U^*,$$

so U_1 and $\tilde{U}_1$ span the same subspace.

$$A(X_1 \ X_2) = (X_1 \ X_2) \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix}, \quad \tilde{A}(\tilde{X}_1 \ \tilde{X}_2) = (\tilde{X}_1 \ \tilde{X}_2) \begin{pmatrix} \tilde{\Lambda}_1 & \\ & \tilde{\Lambda}_2 \end{pmatrix},$$

where $(\overset{k}{X_1} \ \overset{n-k}{X_2})$ and $(\overset{k}{\tilde{X}_1} \ \overset{n-k}{\tilde{X}_2})$ are nonsingular, and Λ_i and $\tilde{\Lambda}_j$ are defined as in (2.2) and (2.3) with λ_i's and $\tilde{\lambda}_j$'s possibly complex. Partition

$$(X_1 \ X_2)^{-1} = \begin{matrix} k \\ n-k \end{matrix} \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix}, \quad (\tilde{X}_1 \ \tilde{X}_2)^{-1} = \begin{matrix} k \\ n-k \end{matrix} \begin{pmatrix} \tilde{Y}_1^* \\ \tilde{Y}_2^* \end{pmatrix}.$$

Define

$$J_X \overset{\text{def}}{=} (X_1 \ X_2) \begin{pmatrix} I_k & \\ & -I_{n-k} \end{pmatrix} \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix}, \quad \hat{A} \overset{\text{def}}{=} J_X \tilde{A} J_X. \tag{2.37}$$

It can be verified that $J_X^2 = I$, $\|J_X\|_2 \leqslant \kappa(X) \equiv \|X\|_2 \|X^{-1}\|_2$. So $\tilde{A}$ and $\hat{A}$ are similar and thus have the same eigenvalues. In fact, a complete eigendecomposition of $\hat{A}$ is

$$\hat{A} = (\hat{X}_1 \ \hat{X}_2) \begin{pmatrix} \tilde{\Lambda}_1 & \\ & \tilde{\Lambda}_2 \end{pmatrix} \begin{pmatrix} \hat{Y}_1^* \\ \hat{Y}_2^* \end{pmatrix},$$

where $\hat{X}_i = J_X \tilde{X}_i$ and $\hat{Y}_i = J_X^* \tilde{Y}_i$ for $i = 1, 2$. As before, we shall now work with $\tilde{A}$ and $\hat{A}$ instead.

Lemma 2.2 (Sun [29]). *We have*

$$\kappa(X) ||| \sin \Theta(\tilde{X}_1, \hat{X}_1) ||| \geqslant ||| \sin 2\Theta(X_1, \tilde{X}_1) |||$$
$$- 2\omega \|X_1\|_2 \|Y_1\|_2 ||| \sin \Theta(X_1, \tilde{X}_1) |||^2,$$

where $\omega = \|(Y_1^* Y_1)^{-1/2} Y_1^* Y_2 (Y_2^* Y_2)^{-1/2}\|_2$.

Proof. It is essential in [29], see also [19, pp. 256–258]. $\square$

The argument so far is borrowed from [29] who extended the treatment of Davis and Kahan for double angle theorems to the generalized eigenvalue problem of a definite matrix pair. Now if $\tilde{A} = D_1^* A D_2$, we write

$$\hat{A} = J_X \tilde{A} J_X = J_X D_1^* A D_2 J_X = J_X D_1^* J_X A J_X D_2 J_X$$
$$= J_X D_1^* J_X D_1^{-*} \tilde{A} D_2^{-1} J_X D_2 J_X.$$

Therefore

$$\hat{A} = \tilde{D}_1^* \tilde{A} \tilde{D}_2, \quad \tilde{D}_1 \overset{\text{def}}{=} D_1^{-1} J_X^* D_1 J_X^*, \quad \tilde{D}_2 \overset{\text{def}}{=} D_2^{-1} J_X D_2 J_X.$$

$\tilde{D}_i$ is close to the identity if D_i is close to some multiple of the identity. Now following the outline given in [21, Remark 3.3] on $\tilde{A}$ and $\hat{A}$ yields $\sin 2\theta$ theorems for diagonalizable matrices.

3. Relative $\sin 2\theta$ theorems for singular subspace variation

Let B and $\tilde{B}$ be two $m \times n$ $(m \geqslant n)$ (complex) matrices with SVDs

$$
B = (U_1 \; U_2) \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix}
$$

and

$$
\tilde{B} = (\tilde{U}_1 \tilde{U}_2) \begin{pmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{V}_1^* \\ \tilde{V}_2^* \end{pmatrix}, \tag{3.1}
$$

where $U = (\overset{k}{U_1} \; \overset{m-k}{U_2})$ and $\tilde{U} = (\overset{k}{\tilde{U}_1} \; \overset{m-k}{\tilde{U}_2})$ are $m \times m$ unitary, and $V = (\overset{k}{V_1} \; \overset{n-k}{V_2})$ and $\tilde{V} = (\overset{k}{\tilde{V}_1} \; \overset{n-k}{\tilde{V}_2})$ are $n \times n$ unitary, $1 \leqslant k < n$, and

$$
\Sigma_1 = \mathrm{diag}(\sigma_1, \ldots, \sigma_k), \quad \Sigma_2 = \mathrm{diag}(\sigma_{k+1}, \ldots, \sigma_n), \tag{3.2}
$$

$$
\tilde{\Sigma}_1 = \mathrm{diag}(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_k), \quad \tilde{\Sigma}_2 = \mathrm{diag}(\tilde{\sigma}_{k+1}, \ldots, \tilde{\sigma}_n). \tag{3.3}
$$

Define

$$
J_U \overset{\text{def}}{=} (U_1 \; U_2) \begin{pmatrix} I_k & \\ & -I_{m-k} \end{pmatrix} \begin{pmatrix} U_1^* \\ U_2^* \end{pmatrix},
$$

$$
J_V \overset{\text{def}}{=} (V_1 \; V_2) \begin{pmatrix} I_k & \\ & -I_{n-k} \end{pmatrix} \begin{pmatrix} V_1^* \\ V_2^* \end{pmatrix}, \quad \hat{B} \overset{\text{def}}{=} J_U \tilde{B} J_V. \tag{3.4}
$$

It can be verified that

$$
J^* = J, \quad J^2 = I, \quad J^{-1} = J \qquad \text{for } J = J_U, J_V \text{ and } J_U B J_V = B.
$$

So both J_U and J_V are unitary. In fact, the SVD of $\hat{B}$ is

$$
\hat{B} = (\hat{U}_1 \; \hat{U}_2) \begin{pmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & \tilde{\Sigma}_2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{V}_1^* \\ \hat{V}_2^* \end{pmatrix},
$$

where $\hat{U}_i = J_U \tilde{U}_i$ and $\hat{V}_i = J_V^* \tilde{V}_i$ for $i = 1, 2$.

Remark 3.1. Write $\tilde{B} = B + H$, then $\hat{B} = B + J_U H J_V$. Wedin $\sin\theta$ theorems [31] applied to $\tilde{B}$ and $\hat{B}$ lead to absolute $\sin 2\theta$ theorems for the singular value problem, e.g.,

$$
\sqrt{\| \sin 2\Theta(U_1, \tilde{U}_1) \|_{\mathrm{F}}^2 + \| \sin 2\Theta(V_1, \tilde{V}_1) \|_{\mathrm{F}}^2} \leqslant \frac{\sqrt{2} \| H \|_{\mathrm{F}}}{\delta}, \tag{3.5}
$$

where $\delta \overset{\text{def}}{=} \min_{\mu \in \sigma(\tilde{\Sigma}_1), \nu \in \sigma_{\mathrm{ext}}(\tilde{\Sigma}_2)} |\mu - \nu|$, and σ_{ext} is defined in Theorem 3.1. But to the best of my knowledge, this has not been done.

We would like to transform the perturbations that force $\tilde{B}$ to $\hat{B}$ into multiplicative ones. Since $J_U B J_V = B$, we have

$$\hat{B} = J_U \tilde{B} J_V = J_U D_1^* B D_2 J_V = J_U D_1^* J_U B J_V D_2 J_V$$

$$= J_U D_1^* J_U D_1^{-*} \tilde{B} D_2^{-1} J_V D_2 J_V.$$

Therefore

$$\hat{B} = \tilde{D}_1^* \tilde{B} \tilde{D}_2, \quad \tilde{D}_1 \stackrel{\text{def}}{=} D_1^{-1} J_U D_1 J_U, \quad \tilde{D}_2 \stackrel{\text{def}}{=} D_2^{-1} J_V D_2 J_V. \tag{3.6}$$

$\tilde{D}_i$ is close to the identity (some unitary matrix) if D_i is close to some multiple of the identity (some unitary matrix).

Theorem 3.1. *Let B and $\tilde{B} = D_1^* B D_2$ be two $m \times n$ ($m \geqslant n$) (complex) matrices with SVDs* (3.1)–(3.3), *where D_1 and D_2 are nonsingular. Let*

$$\tilde{\eta}_2 \stackrel{\text{def}}{=} \min_{\mu \in \sigma(\tilde{\Sigma}_1), \nu \in \sigma_{\text{ext}}(\tilde{\Sigma}_2)} \varrho_2(\mu, \nu) \quad \text{and} \quad \tilde{\eta}_c \stackrel{\text{def}}{=} \min_{\mu \in \sigma(\Sigma_1), \nu \in \sigma_{\text{ext}}(\tilde{\Sigma}_2)} \frac{|\mu - \nu|}{|\nu|}, \tag{3.7}$$

where $\sigma_{\text{ext}}(\tilde{\Sigma}_2) \equiv \sigma(\tilde{\Sigma}_2) \cup \{0\}$ if $m > n$, and $\sigma_{\text{ext}}(\tilde{\Sigma}_2) \equiv \sigma(\tilde{\Sigma}_2)$ otherwise. If $\tilde{\eta}_c, \tilde{\eta}_2 > 0$, then

$$\sqrt{\| \sin 2\Theta(U_1, \tilde{U}_1) \|_F^2 + \| \sin 2\Theta(V_1, \tilde{V}_1) \|_F^2}$$

$$\leqslant \frac{\sqrt{\|(I - \tilde{D}_1^*)\tilde{U}_1\|_F^2 + \|(I - \tilde{D}_1^{-1}\tilde{U}_1\|_F^2 + \|(I - \tilde{D}_2^*)\tilde{V}_1\|_F^2 + \|(I - \tilde{D}_2^{-1})\tilde{V}_1\|_F^2}}{\tilde{\eta}_2},$$

$$\tag{3.8}$$

$$\sqrt{\| \sin 2\Theta(U_1, \tilde{U}_1) \|_F^2 + \| \sin 2\Theta(V_1, \tilde{V}_1) \|_F^2}$$

$$\leqslant \sqrt{\|(I - \tilde{D}_1^*)\tilde{U}_1\|_F^2 + \|(I - \tilde{D}_2^*)\tilde{V}_1\|_F^2}$$

$$+ \frac{1}{\tilde{\eta}_c} \sqrt{\|(\tilde{D}_1^* - \tilde{D}_1^{-1})\tilde{U}_1\|_F^2 + \|(\tilde{D}_2^* - \tilde{D}_2^{-1})\tilde{V}_1\|_F^2},$$

$$\tag{3.9}$$

where $\tilde{D}_i$'s are defined in (3.6).

Proof. It follows from Lemma 2.1 and [21, Theorem 4.1] applied to $\tilde{B}$ and $\hat{B}$. □

Theorem 3.1 has a corollary similar to Corollary 2.1 that yield bounds in terms of the deviations of D_i from the identity or orthogonality. Bounds in any unitarily invariant norm under stronger assumption on the separation of $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$ than that in Theorem 3.1 can also be obtained with the help of Lemma 2.1 and [21, Theorem 4.2].

Remark 3.2. The above theorem applies to a more realistic situation when B can be scaled to improve its condition number. Consider $B = GS$ and $\tilde{B} = \tilde{G}S$ are $m \times n$ ($m \geqslant n$); S is a scaling matrix and both G and $\tilde{G}$ are $m \times n$; G has full column rank. Let $G^\dagger = (G^*G)^{-1}G^*$ the pseudo-inverse of G. Notice that $G^\dagger G = I$. Then

$$\tilde{B} = \tilde{G}S = (G + \Delta G)S = (I + (\Delta G)G^\dagger)GS = (I + (\Delta G)G^\dagger)B.$$

If $\|(\Delta G)G^\dagger\|_2 \leqslant \|G^\dagger\|_2 \|\Delta G\|_2 < 1$, $\tilde{G}$ has full column rank, too. We see $\tilde{B} = \tilde{G}S = [I + (\Delta G)G^\dagger]GS = D_1^* B D_2$, where $D_1^* = I + (\Delta G)G^\dagger$ and $D_2 = I$. Theorem 3.1 can now be applied to B and $\tilde{B}$. We omit explicitly stating them.

Bounds in Theorem 3.1 are invariant under rescaling $D_i \to D_i/\alpha_i$, unlike the existing bounds, e.g., in [12,21]. This provided one of the motivations that led to a recent paper by Stewart and the current author [24]. The other motivation for [24] is to derive bounds that reflect the intrinsic differences in how left and right multiplicative perturbations affect left and right singular subspaces, e.g, when D_2 is unitary it does not affect the left singular subspaces at all. Such bounds can easily obtained when the technique of this paper is combined with the main result of Li and Stewart [24]. For example, we have

Theorem 3.2. *Let the conditions of Theorem* 3.1 *hold, and let* $\tilde{Q}_i$ *be the unitary polar factor of* $\tilde{D}_i$. *Then*

$$\| \sin 2\Theta(U_1, \tilde{U}_1)\|_{\mathrm{F}} \leqslant \frac{\epsilon}{\tilde{\eta}_2} + \|(I - \tilde{Q}_1^*)\tilde{U}_1\|_{\mathrm{F}}, \tag{3.10}$$

$$\| \sin 2\Theta(V_1, \tilde{V}_1)\|_{\mathrm{F}} \leqslant \frac{\epsilon}{\tilde{\eta}_2} + \|(I - \tilde{Q}_2^*)\tilde{V}_1\|_{\mathrm{F}}, \tag{3.11}$$

where $\tilde{D}_i$'s *are defined as in* (3.6), *and*

$$\epsilon^2 = \|(\tilde{D}_1 - \tilde{Q}_1)^*\tilde{U}_1\|_{\mathrm{F}}^2 + \|[\tilde{D}_1^{-1} - \tilde{Q}_1^*]\tilde{U}_1\|_{\mathrm{F}}^2$$
$$+ \|(\tilde{D}_2 - \tilde{Q}_2)^*\tilde{V}_1\|_{\mathrm{F}}^2 + \|(\tilde{D}_2^{-1} - \tilde{Q}_2^*)\tilde{V}_1\|_{\mathrm{F}}^2.$$

It can be seen that the deviations of some multiples of D_i's from orthogonality transform into the deviations of $\tilde{D}_i$'s from orthogonality, and hence in this theorem D_i's contribute to ϵ by the deviations of their multiples from orthogonality. Therefore D_1 affects span(V_1) only by the deviation of its some multiple from orthogonality rather than the identity and similar argument holds for D_2 and span(U_1).

References

[1] J. Barlow, J. Demmel, Computing accurate eigensystems of scaled diagonally dominant matrices, SIAM J. Numer. Anal. 27 (1990) 762–791.

[2] A. Barrlund, The p-relative distance is a metric, SIAM J. Matrix Anal. Appl. (1998), to appear.

[3] C. Davis, W. Kahan, The rotation of eigenvectors by a perturbation. III, SIAM J. Numer. Anal. 7 (1970) 1–46.

[4] P. Deift, J. Demmel, L.-C. Li, C. Tomei, The bidiagonal singular value decomposition and Hamiltonian mechanics, SIAM J. Numer. Anal. 28 (1991) 1463–1516.

[5] J. Demmel, Applied Numerical Linear Algebra, SIAM, Philadelphia, PA, 1997.

[6] J. Demmel, Accurate SVDs of structure matrices, SIAM J. Matrix Anal. Appl. 21 (2000) 562–580.

[7] J. Demmel, W. Gragg, On computing accurate singular values and eigenvalues of matrices with acyclic graphs, Linear Algebra Appl. 185 (1993) 203–217.

[8] J. Demmel, W. Kahan, Accurate singular values of bidiagonal matrices, SIAM J. Sci. Statist. Comput. 11 (1990) 873–912.

[9] J. Demmel, K. Veselić, Jacobi's method is more accurate than QR, SIAM J. Matrix Anal. Appl. 13 (1992) 1204–1245.

[10] J.W. Demmel, M. Gu, S.C. Eisenstat, I. Slapničar, K. Veselić, Z. Drmač, Computing the singular value decomposition with high relative accuracy, Linear Algebra Appl. 299 (1999) 21–80.

[11] G. DiLena, R.I. Peluso, G. Piazza, Results on the relative perturbation of the singular values of a matrix, BIT 33 (1993) 647–653.

[12] S.C. Eisenstat, I.C.F. Ipsen, Relative perturbation bounds for eigenspaces and singular vector subspaces, in: J.G. Lewis (Ed.), Proceedings of the Fifth SIAM Conference on Applied Linear Algebra, SIAM, Philadelphia, PA, 1994, pp. 62–66.

[13] S.C. Eisenstat, I.C.F. Ipsen, Relative perturbation techniques for singular value problems, SIAM J. Numer. Anal. 32 (1995) 1972–1988.

[14] K.V. Fernando, B.N. Parlett, Accurate singular values and differential qd algorithms, Numer. Math. 67 (1994) 191–229.

[15] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[16] I.C.F. Ipsen, Relative perturbation results for matrix eigenvalues and singular values, Acta Numerica 7 (1998) 151–201.

[17] W. Kahan, Accurate eigenvalues of a symmetric tridiagonal matrix, Technical Report CS41, Computer Science Department, Stanford University, Stanford, CA, 1966 (revised June 1968).

[18] C.-K. Li, R. Mathias, The Lidskii–Mirsky–Wielandt theorem – additive and multiplicative versions, Numer. Math. 81 (1999) 377–413.

[19] R.-C. Li, On perturbations of matrix pencils with real spectra, Math. Comput. 62 (1994) 231–265.

[20] R.-C. Li, Relative perturbation theory: I. Eigenvalue and singular value variations, SIAM J. Matrix Anal. Appl. 19 (1998) 956–982.

[21] R.-C. Li, Relative perturbation theory: II. Eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl. 20 (1999) 471–492.

[22] R.-C. Li, Relative perturbation theory: (III) more bounds on eigenvalue variation, Linear Algebra Appl. 266 (1997) 337–345.

[23] R.-C. Li, A bound on the solution to a structured Sylvester equation with an application to relative perturbation theory, SIAM J. Matrix Anal. Appl. 21 (2000) 440–445.

[24] R.-C. Li, G.W. Stewart, A new relative perturbation theorem for singular subspaces, Linear Algebra Appl. (2000), to appear.

[25] R. Mathias, Accurate eigensystem computations by Jacobi methods, SIAM J. Matrix Anal. Appl. 16 (1995) 977–1003.

[26] R. Mathias, Spectral perturbation bounds for positive definite matrices, SIAM J. Matrix Anal. Appl. 18 (1997) 959–980.

[27] B.N. Parlett, O.A. Marques, An implementation of the dqds algorithm (positive case), submitted for publication.

[28] I. Slapničar, Accurate symmetric eigenreduction by a Jacobi method, Ph.D. Thesis, Fernuniversität – Gesamthochschule – Hagen, Facbereich Mathematik, 1992.

[29] J.-G. Sun, Perturbation bounds for eigenspaces of a definite matrix pair, Numer. Math. 41 (1983) 321–343.

[30] C.F. Van Loan, Generalizing the singular value decomposition, SIAM J. Numer. Anal. 13 (1976) 76–83.

[31] P.-Å. Wedin, Perturbation bounds in connection with singular value decomposition, BIT 12 (1972) 99–111.