MATHEMATICS OF COMPUTATION
Volume 65, Number 214
April 1996, Pages 637–645

ON PERTURBATIONS OF MATRIX PENCILS WITH REAL SPECTRA. II

RAJENDRA BHATIA AND REN-CANG LI

ABSTRACT. A well-known result on spectral variation of a Hermitian matrix due to Mirsky is the following: *Let A and $\widetilde{A}$ be two $n\times n$ Hermitian matrices, and let $\lambda_1,\dots,\lambda_n$ and $\widetilde{\lambda}_1,\dots,\widetilde{\lambda}_n$ be their eigenvalues arranged in ascending order. Then $|||\mathrm{diag}\,(\lambda_1-\widetilde{\lambda}_1,\dots,\lambda_n-\widetilde{\lambda}_n)||| \le |||A-\widetilde{A}|||$ for any unitarily invariant norm $|||\cdot|||$.* In this paper, we generalize this to the perturbation theory for diagonalizable matrix pencils with real spectra. The much studied case of definite pencils is included in this.

1. INTRODUCTION

In the perturbation theory of the eigenvalue problem $Ax=\lambda x$, a major chapter is devoted to obtaining perturbation bounds for eigenvalues of Hermitian matrices in all unitarily invariant norms (ref. [1, 16]). Here the theory is in a satisfactory and finished form. We have the following result: *Let A and $\widetilde{A}$ be two $n\times n$ Hermitian matrices, and let $\lambda_1,\dots,\lambda_n$ and $\widetilde{\lambda}_1,\dots,\widetilde{\lambda}_n$ respectively be their eigenvalues arranged in ascending order. Then for any unitarily invariant norm $|||\cdot|||$*

$$|||\mathrm{diag}\,(\lambda_1-\widetilde{\lambda}_1,\dots,\lambda_n-\widetilde{\lambda}_n)||| \le |||A-\widetilde{A}|||. \tag{1.1}$$

This was proved by Weyl [21] for the spectral norm and by Loewner [13] for the Frobenius norm. Also, for the Frobenius norm it is a corollary of a theorem by Hoffman and Wielandt [8], who established the theorem for normal matrices. For all unitarily invariant norms the inequality (1.1) was proved by Mirsky [14]. He derived it from a theorem of Wielandt [22] and Lidskii [12].

For the generalized eigenvalue problem $Ax=\lambda Bx$ the corresponding perturbation theory is of more recent origin and is in a less finished form. For the case of definite matrix pencils (the counterpart of the Hermitian case in the standard eigenvalue problem) an analog of (1.1) for the spectral norm was obtained by Stewart [15], Sun [17] and Li [9], and for the Frobenius norm by Sun [19] and Li [9]. Most notably, Li [9] considered the more general case of diagonalizable pencils with real spectra. Somewhat complicated and preliminary results were also obtained for all unitarily invariant norms in [11].

In this paper, which is a continuation of [9], we obtain better and simpler results for the abovementioned problem for all unitarily invariant norms. Our principal

Received by the editor April 13, 1994 and, in revised form, August 4, 1994.
1991 *Mathematics Subject Classification.* Primary 15A22, 15A42, 65F15.
Key words and phrases. Diagonalizable matrix pencil, definite pencil, real spectrum, unitarily invariant norm, perturbation bound.
The first author thanks NSERC Canada for financial support.

observation is that using certain perturbation identities proved in [9] we can reduce the problem of finding perturbation bounds for diagonalizable matrix pencils with real spectra to that of finding perturbation bounds for (the standard eigenvalue problem of) matrices similar to unitary matrices. The latter problem has been solved in [3] for all unitarily invariant norms and in [10] for some other norms.

2. Preliminaries

Throughout the paper, capital letters are used for matrices, lower-case Latin letters for column vectors or scalars and lower-case Greek letters for scalars; $\mathbb{C}^{m\times n}$ is the set of $m\times n$ complex matrices, $\mathbb{U}_n\subset\mathbb{C}^{n\times n}$ the set of $n\times n$ unitary matrices, $\mathbb{C}^m=\mathbb{C}^{m\times 1}$, $\mathbb{C}=\mathbb{C}^1$; and $\mathbb{R}$ is the set of real numbers. The symbol I stands for identity matrices of suitable dimensions (which should be clear from the context). A^H and A^+ denote the conjugate transpose, and the Moore-Penrose inverse of A, respectively. P_X is the orthogonal projection onto the column space of X. It is easy to verify that [16]

$$P_X=XX^+,\quad P_{X^H}=X^+X.$$

We will consider *unitarily invariant norms* $|||\cdot|||$ of matrices. In this we follow Mirsky [14] and Stewart and Sun [16]. To say that the norm is *unitarily invariant* on $\mathbb{C}^{m\times n}$ means that it satisfies, besides the usual properties of any norm, also

(1) $|||UAV|||=|||A|||$ for any $U\in\mathbb{U}_m$ and $V\in\mathbb{U}_n$;
(2) $|||A|||=\|A\|_2$ for any $A\in\mathbb{C}^{m\times n}$, $\operatorname{rank}A=1$.

Two unitarily invariant norms used frequently are the *spectral norm* $\|\cdot\|_2$ and the *Frobenius norm* $\|\cdot\|_F$.

Consider the pencil $A-\lambda B$ with A, $B\in\mathbb{C}^{n\times n}$. The pencil is said to be *regular* if $\det(A-\lambda B)\not\equiv 0$. Let

$$\mathbb{G}_{1,2}=\{(\alpha,\beta)\neq(0,0)\ :\ \alpha,\beta\in\mathbb{C}\}.$$

In what follows, to avoid ambiguity for our purpose, all points $(\xi\alpha,\xi\beta)$ $(\xi\neq 0)$ will be identified as the same. An element (α,β) of $\mathbb{G}_{1,2}$ is called *a generalized eigenvalue* of a regular pencil $A-\lambda B$ if $\det(\beta A-\alpha B)=0$. An element (α,β) of $\mathbb{G}_{1,2}$ is said to be *real* if there exists $0\neq\xi\in\mathbb{C}$ such that $\xi\alpha$, $\xi\beta\in\mathbb{R}$. (For instance, (i,i) is real.) The spectrum of a regular pencil $A-\lambda B$ consists of all its generalized eigenvalues (counted according to their algebraic multiplicities), and is denoted by $\lambda(A,B)$.

Definition 2.1. A regular matrix pencil $A-\lambda B$ of order n is *diagonalizable*, or *normalizable*, if there exist invertible matrices X, $Y\in\mathbb{C}^{n\times n}$ such that

$$Y^HAX=\operatorname{diag}(\alpha_1,\dots,\alpha_n),\quad Y^HBX=\operatorname{diag}(\beta_1,\dots,\beta_n).$$

Let (α,β), $(\gamma,\delta)\in\mathbb{G}_{1,2}$. The *chordal metric*

$$\rho\big((\alpha,\beta),(\gamma,\delta)\big)\stackrel{\text{def}}{=}\frac{|\delta\alpha-\gamma\beta|}{\sqrt{|\alpha|^2+|\beta|^2}\sqrt{|\gamma|^2+|\delta|^2}}\tag{2.1}$$

will be used to measure the distance between these two points. To measure the distance between two regular pencils $A-\lambda B$ and $\widetilde{A}-\lambda\widetilde{B}$ of order n, we use

$$\left|\left|\left|P_{Z^H}-P_{\widetilde{Z}^H}\right|\right|\right|\quad\text{and}\quad\left|\left|\left|Z-\widetilde{Z}\right|\right|\right|,\tag{2.2}$$

where

$$Z = (A, B) \quad \text{and} \quad \widetilde{Z} = (\widetilde{A}, \widetilde{B}). \tag{2.3}$$

3. Main results

Henceforth, $A - \lambda B$ and $\widetilde{A} - \lambda\widetilde{B}$ will always be two regular pencils which are diagonalizable and admit decompositions

$$\left\{\begin{array}{rcl} Y^H A X &=& \Lambda, \\ Y^H B X &=& \Omega, \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{rcl} \widetilde{Y}^H \widetilde{A} \widetilde{X} &=& \widetilde{\Lambda}, \\ \widetilde{Y}^H \widetilde{B} \widetilde{X} &=& \widetilde{\Omega}, \end{array}\right. \tag{3.1}$$

where X, Y, $\widetilde{X}$, $\widetilde{Y} \in \mathbb{C}^{n\times n}$ are nonsingular matrices,

$$\left\{\begin{array}{rcl} \Lambda &=& \operatorname{diag}(\alpha_1, \ldots, \alpha_n), \\ \Omega &=& \operatorname{diag}(\beta_1, \ldots, \beta_n), \end{array}\right. \quad \text{and} \quad \left\{\begin{array}{rcl} \widetilde{\Lambda} &=& \operatorname{diag}(\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_n), \\ \widetilde{\Omega} &=& \operatorname{diag}(\widetilde{\beta}_1, \ldots, \widetilde{\beta}_n), \end{array}\right. \tag{3.2}$$

and α_i, β_i, $\widetilde{\alpha}_j$, $\widetilde{\beta}_j \in \mathbb{R}$, $i, j = 1, \ldots, n$. Clearly,

$$\begin{array}{rcl} \lambda(A, B) &=& \{(\alpha_i, \beta_i),\ i = 1, \ldots, n\}, \\ \lambda(\widetilde{A}, \widetilde{B}) &=& \{(\widetilde{\alpha}_j, \widetilde{\beta}_j),\ j = 1, \ldots, n\}. \end{array}$$

Also, we define Z and $\widetilde{Z}$ as in (2.3). We then have

Theorem 3.1. *There exists a permutation σ of $\{1, \ldots, n\}$ such that for any unitarily invariant norm $|||\cdot|||$*

$$\left|\left|\left| \begin{pmatrix} & \Sigma \\ \Sigma & \end{pmatrix} \right|\right|\right| \le \frac{\pi}{2} \widehat{\kappa}(X) \widehat{\kappa}(\widetilde{X}) \left|\left|\left| P_{Z^H} - P_{\widetilde{Z}^H} \right|\right|\right|, \tag{3.3}$$

where

$$\begin{array}{rcl} \widehat{\kappa}(X) &=& \max\{\|X^{-1}\|_2, \|X\|_2\}^2, \\ \widehat{\kappa}(\widetilde{X}) &=& \max\{\|\widetilde{X}^{-1}\|_2, \|\widetilde{X}\|_2\}^2, \\ \Sigma &=& \operatorname{diag}\left(\rho\left((\alpha_1, \beta_1), (\widetilde{\alpha}_{\sigma(1)}, \widetilde{\beta}_{\sigma(1)})\right), \ldots, \rho\left((\alpha_n, \beta_n), (\widetilde{\alpha}_{\sigma(n)}, \widetilde{\beta}_{\sigma(n)})\right)\right). \end{array} \tag{3.4}$$

In Theorem 3.1, the distance between the two pencils $A - \lambda B$ and $\widetilde{A} - \lambda\widetilde{B}$ is measured via $\left|\left|\left| P_{Z^H} - P_{\widetilde{Z}^H} \right|\right|\right|$. In the following theorem, it is measured more directly by $\left|\left|\left| Z - \widetilde{Z} \right|\right|\right|$.

Theorem 3.2. *In the decompositions* (3.1) *and* (3.2), *if*

$$\alpha_i^2 + \beta_i^2 = \widetilde{\alpha}_j^2 + \widetilde{\beta}_j^2 = 1, \quad i, j = 1, \ldots, n, \tag{3.5}$$

then there exists a permutation σ of $\{1, \ldots, n\}$ such that

$$|||\Sigma||| \le \frac{\pi}{2} \min\{\kappa(X)\|\widetilde{X}\|_2\|\widetilde{Y}^H\|_2,\ \kappa(\widetilde{X})\|X\|_2\|Y^H\|_2\} \left|\left|\left| Z - \widetilde{Z} \right|\right|\right|, \tag{3.6}$$

where Σ is defined as in (3.4), *and*

$$\kappa(X) = \|X\|_2\|X^{-1}\|_2, \quad \kappa(\widetilde{X}) = \|\widetilde{X}\|_2\|\widetilde{X}^{-1}\|_2.$$

A remark regarding the inequalities (3.3) and (3.6) is in order: (3.3) holds as long as we have the decompositions (3.1) and (3.2); while in order for (3.6) to be true, these decompositions have to be adjusted so that the normalization assumption (3.5) holds. To some readers this normalization assumption could be annoying. However, we can get rid of it with the help of the following lemma.

Lemma 3.1. *If $A - \lambda B$ has the decomposition* (3.1) *and* (3.2) *with $\alpha_i^2 + \beta_i^2 = 1$, $i = 1, \ldots, n$, then $\|Y^H\|_2 \le \|Z^+\|_2 \|X^{-1}\|_2$.*

Theorem 3.3. *Under the conditions of Theorem* 3.1, *there exists a permutation σ of $\{1, \ldots, n\}$ such that*

$$\|\!|\Sigma|\!\| \le \frac{\pi}{2} \kappa(X)\kappa(\widetilde{X}) \min\{\|Z^+\|_2, \|\widetilde{Z}^+\|_2\} \left|\!\left|\!\left| Z - \widetilde{Z} \right|\!\right|\!\right|, \tag{3.7}$$

where Σ is defined as in (3.4).

Like (3.3), the inequality (3.7) holds without assuming (3.5) for the decompositions (3.1) and (3.2). The above theorems are directly applicable to *definite pencils.*

Definition 3.1. Let A, $B \in \mathbb{C}^{n\times n}$ be Hermitian; $A - \lambda B$ is said to be *a definite pencil of order n*, if

$$c(A, B) \stackrel{\text{def}}{=} \min\{|x^H(A + iB)x| \ : \ \|x\|_2 = 1\} > 0. \tag{3.8}$$

The quantity $c(A, B)$ is called the *Crawford number* of the definite pencil $A - \lambda B$.

Lemma 3.2. *Let $A - \lambda B$ be a definite pencil of order n. Then there is a nonsingular matrix $X \in \mathbb{C}^{n\times n}$ such that*

$$X^H A X = \operatorname{diag}(\alpha_1, \ldots, \alpha_n), \quad X^H B X = \operatorname{diag}(\beta_1, \ldots, \beta_n). \tag{3.9}$$

In Lemma 3.2, it is easily verified that α_i, $\beta_i \in \mathbb{R}$, and by appropriate choice of X, we can make $\alpha_i^2 + \beta_i^2 = 1$.

Lemma 3.3. *In* (3.9) *of Lemma* 3.2, *if $\alpha_i^2 + \beta_i^2 = 1$, $i = 1, \ldots, n$, then*

$$\|X\|_2 \le \frac{1}{\sqrt{c(A, B)}}, \quad \|X^{-1}\|_2 \le \frac{\|Z\|_2}{\sqrt{c(A, B)}}. \tag{3.10}$$

This lemma is due to Elsner and Sun [7]. Using this and Theorem 3.1 we can prove

Theorem 3.4. *Let $A - \lambda B$ and $\widetilde{A} - \lambda \widetilde{B}$ be two definite pencils of order n. Then there exists a permutation σ of $\{1, \ldots, n\}$ such that for any unitarily invariant norm $|\!|\!| \cdot |\!|\!|$*

(3.11)
$$\left|\!\left|\!\left| \begin{pmatrix} \Sigma & \\ & \Sigma \end{pmatrix} \right|\!\right|\!\right| \le \frac{\pi}{2} \frac{\max\{1, \|Z\|_2^2\} \max\{1, \|\widetilde{Z}\|_2^2\}}{c(A, B)c(\widetilde{A}, \widetilde{B})} \left|\!\left|\!\left| P_{Z^H} - P_{\widetilde{Z}^H} \right|\!\right|\!\right|,$$

where Σ is as defined in (3.4).

Lemmas 3.2, 3.3, and Theorem 3.2 yield

Theorem 3.5. *Under the conditions of Theorem* 3.4, *there exists a permutation σ of $\{1, \ldots, n\}$ such that*

$$|\!|\!|\Sigma|\!|\!| \le \frac{\pi}{2} \frac{\min\{\|Z\|_2, \|\widetilde{Z}\|_2\}}{c(A, B)c(\widetilde{A}, \widetilde{B})} \left|\!\left|\!\left| Z - \widetilde{Z} \right|\!\right|\!\right|, \tag{3.12}$$

where Σ is defined as in (3.4).

4. Proofs of theorems

One of our key tricks is the perturbation equation listed in the following lemma (see [9, pp. 244, 253]).

Lemma 4.1. *Let $A-\lambda B$ and $\widetilde{A}-\lambda\widetilde{B}$ be as described by* (3.1) *and* (3.2). *We have*

$$\begin{aligned}\widetilde{\Lambda}\widetilde{X}^{-1}X\Omega-\widetilde{\Omega}\widetilde{X}^{-1}X\Lambda &= -(\widetilde{\Lambda}\widetilde{X}^{-1},\widetilde{\Omega}\widetilde{X}^{-1})(P_{Z^H}-P_{\widetilde{Z}^H})\begin{pmatrix} X\Omega \\ -X\Lambda\end{pmatrix}\\ &\stackrel{\mathrm{def}}{=} E,\end{aligned}\tag{4.1}$$

$$\begin{aligned}\Lambda X^{-1}\widetilde{X}\widetilde{\Omega}-\Omega X^{-1}\widetilde{X}\widetilde{\Lambda} &= -(\Lambda X^{-1},\Omega X^{-1})(P_{\widetilde{Z}^H}-P_{Z^H})\begin{pmatrix} \widetilde{X}\widetilde{\Omega} \\ -\widetilde{X}\widetilde{\Lambda}\end{pmatrix}\\ &\stackrel{\mathrm{def}}{=} \widetilde{E}.\end{aligned}\tag{4.2}$$

Lemma 4.2 ([10]). *Suppose α_i, β_i, $\widetilde{\alpha}_j$, $\widetilde{\beta}_j \in \mathbb{R}$ satisfy $\alpha_i^2+\beta_i^2=\widetilde{\alpha}_j^2+\widetilde{\beta}_j^2=1$, $i,j=1,\dots,n$. Let Λ, Ω, $\widetilde{\Lambda}$ and $\widetilde{\Omega}$ be as defined in* (3.2), *and let $U\in\mathbb{U}_n$. Then there exists a permutation σ of $\{1,\dots,n\}$ such that for any unitarily invariant norm $|||\cdot|||$*

$$\begin{aligned}&\left|\left|\left|\operatorname{diag}\left(\rho\big((\alpha_1,\beta_1),(\widetilde{\alpha}_{\sigma(1)},\widetilde{\beta}_{\sigma(1)})\big),\dots,\rho\big((\alpha_n,\beta_n),(\widetilde{\alpha}_{\sigma(n)},\widetilde{\beta}_{\sigma(n)})\big)\right)\right|\right|\right|\\ &\qquad\le \tfrac{\pi}{2}\left|\left|\left|\widetilde{\Lambda}U\Omega-\widetilde{\Omega}U\Lambda\right|\right|\right|,\end{aligned}\tag{4.3}$$

where the constant $\pi/2$ is best possible.

Lemma 4.3. *Let U and V be two $n\times n$ unitary matrices, and let Γ be a positive diagonal matrix. Then for every unitarily invariant norm $|||\cdot|||$*

$$\|\Gamma^{-1}\|_2\,|||U\Gamma-\Gamma V|||\ge |||U-V|||.\tag{4.4}$$

This lemma was proved in [3]. Here, for the sake of completeness, we present a proof which is a little more direct.

Proof. Define $F\stackrel{\mathrm{def}}{=}U\Gamma-\Gamma V$. Then $VF^HU=V\Gamma-\Gamma U$, and

$$(U-V)\Gamma+\Gamma(U-V)=F-VF^HU.\tag{4.5}$$

Since U and V are unitary,

$$\left|\left|\left|F-VF^HU\right|\right|\right|\le |||F|||+\left|\left|\left|VF^HU\right|\right|\right|=2\,|||F|||.$$

By [6, Theorem 5.2],

$$\begin{aligned}2\,|||F||| &\ge |||(U-V)\Gamma+\Gamma(U-V)|||\\ &\ge 2\|\Gamma^{-1}\|_2^{-1}\,|||U-V|||,\end{aligned}$$

which proves (4.4). □

Lemma 4.4. *Let α_i, β_i, $\widetilde{\alpha}_j$, $\widetilde{\beta}_j$ and Λ, Ω, $\widetilde{\Lambda}$, $\widetilde{\Omega}$ be as described in Lemma* 4.2, *and let T be a nonsingular matrix. Then there exists a permutation σ of $\{1,\dots,n\}$ such that for any unitarily invariant norm $|||\cdot|||$*

$$\begin{aligned}&\left|\left|\left|\operatorname{diag}\left(\rho\big((\alpha_1,\beta_1),(\widetilde{\alpha}_{\sigma(1)},\widetilde{\beta}_{\sigma(1)})\big),\dots,\rho\big((\alpha_n,\beta_n),(\widetilde{\alpha}_{\sigma(n)},\widetilde{\beta}_{\sigma(n)})\big)\right)\right|\right|\right|\\ &\qquad\le \tfrac{\pi}{2}\|T^{-1}\|_2\left|\left|\left|\widetilde{\Lambda}T\Omega-\widetilde{\Omega}T\Lambda\right|\right|\right|,\end{aligned}\tag{4.6}$$

where the constant $\pi/2$ is best possible.

Proof. Let $R = \widetilde{\Lambda} T \Omega - \widetilde{\Omega} T \Lambda$, and let

$$\begin{pmatrix} L \\ M \end{pmatrix} = \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \Lambda \\ \Omega \end{pmatrix}, \quad \begin{pmatrix} \widetilde{L} \\ \widetilde{M} \end{pmatrix} = \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \widetilde{\Lambda} \\ \widetilde{\Omega} \end{pmatrix}.$$

Then the diagonal matrices L, M, $\widetilde{L}$ and $\widetilde{M}$ are all in $\mathbb{U}_n$. Hence, we have $\widetilde{L}TM - \widetilde{M}TL = 2R$, and this gives

$$\widetilde{M}^H \widetilde{L} T - T L M^H = 2 \widetilde{M}^H R M^H. \tag{4.7}$$

Let $T = U\Gamma V^H$ be the singular value decomposition of T. From (4.7) it follows that

$$U^H \widetilde{M}^H \widetilde{L} U \Gamma - \Gamma V^H L M^H V = 2 U^H \widetilde{M}^H R M^H V. \tag{4.8}$$

If $\widehat{U} \stackrel{\text{def}}{=} U^H \widetilde{M}^H \widetilde{L} U$ and $\widehat{V} \stackrel{\text{def}}{=} V^H L M^H V$, then $\widehat{U}$ and $\widehat{V}$ are unitary matrices. Therefore, from (4.8) and Lemma 4.3, we get

$$\begin{aligned} 2\|T^{-1}\|_2 \, |||R||| &= \|\Gamma^{-1}\|_2 \left|\left|\left| 2U^H \widetilde{M}^H R M^H V \right|\right|\right| \\ &\geq \left|\left|\left| U^H \widetilde{M}^H \widetilde{L} U - V^H L M^H V \right|\right|\right| \qquad \text{(Lemma 4.3)} \\ &= \left|\left|\left| \widetilde{M}^H \widetilde{L} U V^H - U V^H L M^H \right|\right|\right| \\ &= \left|\left|\left| \widetilde{L} U V^H M - \widetilde{M} U V^H L \right|\right|\right| \\ &= 2 \left|\left|\left| \widetilde{\Lambda} U V^H \Omega - \widetilde{\Omega} U V^H \Lambda \right|\right|\right| . \end{aligned}$$

The conclusion of this lemma now follows from (4.3). □

Lemma 4.5. *Let* $F = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix}$ *be a partitioned matrix. Then*

$$\left|\left|\left| \begin{pmatrix} F_{11} & 0 \\ 0 & F_{22} \end{pmatrix} \right|\right|\right| \leq \left|\left|\left| \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \right|\right|\right|$$

for every unitarily invariant norm $|||\cdot|||$.

The reader is referred to [1, p. 31] for this lemma.

Proof of Theorem 3.1. It follows from Lemmas 4.1 and 4.4 that there exists a permutation σ of $\{1, \ldots, n\}$ such that

$$|||\Sigma||| \leq \frac{\pi}{2} \|\widetilde{X}^{-1} X\|_2 \, |||E||| , \quad |||\Sigma||| \leq \frac{\pi}{2} \|X^{-1} \widetilde{X}\|_2 \left|\left|\left| \widetilde{E} \right|\right|\right| ,$$

where Σ is as defined in (3.4). Hence by Lemma 4.5,

(4.9)

$$\left|\left|\left|\begin{pmatrix} \Sigma & \\ & \Sigma \end{pmatrix}\right|\right|\right| \le \max\{\|\widetilde{X}^{-1}X\|_2, \|X^{-1}\widetilde{X}\|_2\} \left|\left|\left|\begin{pmatrix} E & \\ & \widetilde{E}^H \end{pmatrix}\right|\right|\right|$$

$$\le \max\{\|\widetilde{X}^{-1}X\|_2, \|X^{-1}\widetilde{X}\|_2\}$$

$$\times \left|\left|\left|\begin{pmatrix} \widetilde{\Lambda}\widetilde{X}^{-1} & \widetilde{\Omega}\widetilde{X}^{-1} \\ \widetilde{\Omega}\widetilde{X}^H & -\widetilde{\Lambda}\widetilde{X}^H \end{pmatrix}(P_{Z^H} - P_{\widetilde{Z}^H})\begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix}\right|\right|\right|$$

$$\le \max\{\|\widetilde{X}^{-1}X\|_2, \|X^{-1}\widetilde{X}\|_2\}$$

$$\times \left\|\begin{pmatrix} \widetilde{\Lambda}\widetilde{X}^{-1} & \widetilde{\Omega}\widetilde{X}^{-1} \\ \widetilde{\Omega}\widetilde{X}^H & -\widetilde{\Lambda}\widetilde{X}^H \end{pmatrix}\right\|_2 \left|\left|\left|P_{Z^H} - P_{\widetilde{Z}^H}\right|\right|\right| \left\|\begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix}\right\|_2.$$

The proof is completed by noting

$$\left\|\begin{pmatrix} \widetilde{\Lambda}\widetilde{X}^{-1} & \widetilde{\Omega}\widetilde{X}^{-1} \\ \widetilde{\Omega}\widetilde{X}^H & -\widetilde{\Lambda}\widetilde{X}^H \end{pmatrix}\right\|_2 \le \max\{\|\widetilde{X}^{-1}\|_2, \|\widetilde{X}\|_2\}, \tag{4.10}$$

$$\left\|\begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix}\right\|_2 \le \max\{\|X^{-1}\|_2, \|X\|_2\}, \tag{4.11}$$

proved in Li [9, p. 246]. □

Proofs of Theorems 3.2 and 3.3 are quite similar and are based on the following two identities proved also in Li [9, p. 247]:

$$\widetilde{\Lambda}\widetilde{X}^{-1}X\Omega - \widetilde{\Omega}\widetilde{X}^{-1}X\Lambda = -\widetilde{Y}^H(Z - \widetilde{Z})\begin{pmatrix} X & \\ & X \end{pmatrix}\begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}, \tag{4.12}$$

$$\Lambda X^{-1}\widetilde{X}\widetilde{\Omega} - \Omega X^{-1}\widetilde{X}\widetilde{\Lambda} = -Y^H(\widetilde{Z} - Z)\begin{pmatrix} \widetilde{X} & \\ & \widetilde{X} \end{pmatrix}\begin{pmatrix} \widetilde{\Omega} \\ -\widetilde{\Lambda} \end{pmatrix}. \tag{4.13}$$

Proof of Theorem 3.4. By Lemma 3.2, we know that $A - \lambda B$ and $\widetilde{A} - \lambda\widetilde{B}$ admit decompositions

$$\begin{cases} X^H AX = \Lambda, \\ X^H BX = \Omega, \end{cases} \quad \text{and} \quad \begin{cases} \widetilde{X}^H \widetilde{A}\widetilde{X} = \widetilde{\Lambda}, \\ \widetilde{X}^H \widetilde{B}\widetilde{X} = \widetilde{\Omega}, \end{cases} \tag{4.14}$$

where X, $\widetilde{X} \in \mathbb{C}^{n\times n}$ are nonsingular matrices, and Λ, Ω, $\widetilde{\Lambda}$ and $\widetilde{\Omega}$ are of the form (3.2) with α_i, β_i, $\widetilde{\alpha}_j$, $\widetilde{\beta}_j \in \mathbb{R}$ and $\alpha_i^2 + \beta_i^2 = \widetilde{\alpha}_j^2 + \widetilde{\beta}_j^2 = 1$, $i, j = 1, \ldots, n$. So from Theorem 3.1 it follows that there exists a permutation σ of $\{1, \ldots, n\}$ such that (3.3) holds. (3.11) follows from (3.3) and (3.10). □

5. Concluding remarks

1. P_{Z^H} is invariant under premultiplications of A and B by nonsingular matrices. Thus using $P_{Z^H} - P_{\widetilde{Z}^H}$ has advantages in the case when $\widetilde{A} - \lambda\widetilde{B} \approx Q(A - \lambda B)$ for some nonsingular matrix Q. In fact, if $\widetilde{A} - \lambda\widetilde{B} = Q(A - \lambda B)$, then $P_{Z^H} - P_{\widetilde{Z}^H} = 0$ and (3.3) is actually an equality.
2. Perturbation theory for matrices similar to unitary matrices, and that for diagonalizable matrix pencils with real spectra, are really the same once (4.1), (4.2), (4.12), and (4.13) are established. This can be seen from the proof of Lemma 4.4.

3. The constant $\frac{\pi}{2}$ is best possible in the sense that it cannot be replaced by any smaller number for all dimensionalities and unitarily invariant norms. On the other hand, theorems in [9] show that it can be replaced by 1 for the spectral norm and the Frobenius norm.
4. The constant $\frac{\pi}{2}$ comes from Lemma 4.2, which is proved in [10], based on a perturbation theorem of Bhatia, Davis and McIntosh [4] for unitary matrices:

$$\left|\left|\left|\operatorname{diag}\left(\alpha_1-\widetilde{\alpha}_{\sigma(1)},\ldots,\alpha_n-\widetilde{\alpha}_{\sigma(n)}\right)\right|\right|\right| \le \frac{\pi}{2}\left|\left|\left|A-\widetilde{A}\right|\right|\right|, \tag{5.1}$$

where A and $\widetilde{A}$ are unitary matrices having eigenvalues $\{\alpha_i\}_{i=1}^n$ and $\{\widetilde{\alpha}_i\}_{i=1}^n$, respectively, and σ is an appropriate permutation of $\{1,2,\ldots,n\}$. Again, $\frac{\pi}{2}$ here is best, but can be replaced by 1 for the spectral norm and the Frobenius norm. It is interesting to notice that the spectral norm is the Schatten ∞-norm, while the Frobenius norm is the Schatten 2-norm. So we conjecture that for the Schatten p-norms $(2 \le p \le \infty)$, the constant $\frac{\pi}{2}$ in (5.1) could be improved to 1. On the other hand, as we can see from Remark 2 above, any possible improvement of (5.1) in the future will lead to a corresponding improvement of theorems proved in this paper.

Acknowledgment

We would like to thank Professor G. W. Stewart for his constructive comments on our historical account for the inequality (1.1).

References

1. R. Bhatia, *Perturbation bounds for matrix eigenvalues*, Longman, Essex and Wiley, New York, 1987. MR **88k:**15020
2. R. Bhatia and Ch. Davis, A bound for the spectral variation of a unitary operator, *Linear Multilinear Algebra* **15**(1984), 71–76. MR **85b:**15020
3. R. Bhatia, Ch. Davis and F. Kittaneh, Some inequalities for commutators and an application to spectral variation, *Aequationes Mathematicae*, **41**(1991), 70–78. MR **92j:**47023
4. R. Bhatia, Ch. Davis and A. McIntosh, Perturbation of spectral subspaces and solution of linear operator equations, *Linear Algebra Appl.*, **52/53**(1983), 45–67. MR **85a:**47020
5. C. R. Crawford, A stable generalized eigenvalue problem, *SIAM J. Numer. Anal.*, **8**(1976), 854–860. MR **55:**4628
6. Ch. Davis, and W. M. Kahan, The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.*, **7**(1970), 1–46. MR **41:**9044
7. L. Elsner and J.-G. Sun, Perturbation theorems for the generalized eigenvalue problem, *Linear Algebra Appl.,* **48**(1982), 341–357. MR **84f:**15012
8. A. J. Hoffman and H. W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.*, **20**(1953), 37–39. MR **14:**611
9. R.-C. Li, On perturbations of matrix pencils with real spectra, *Math. Comp.*, **62**(1994), 231–265. MR **94f:**15011
10. ———, Norms of certain matrices with applications to variations of the spectra of matrices and matrix pencils, *Linear Algebra Appl.,* **182**(1993), 199–234. MR **94c:**15040
11. ———, A perturbation bound for definite pencils, *Linear Algebra Appl.,* **179**(1993), 191–202. MR **94h:**15007
12. V. B. Lidskii, The proper values of the sum and product of symmetric matrices, *Doklady Akademii Nauk SSSR*, **75**(1950), 769–772. In Russian. Translation by C. Benster available from the National Translation Center of the Library of Congress. MR **14:**528
13. K. Loewner, Über monotone Matrixfunktionen, *Math. Z.*, **38**(1934), 177–216.
14. L. Mirsky, Symmetric gauge functions and unitarily invariant norms, *Quart. J. Math. Oxford*, **11**(1960), 50–59. MR **22:**5639
15. G. W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, *Linear Algebra Appl.*, **23**(1979), 69–85. MR **80c:**15007

16. G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990. MR **92a:**65017
17. J.-G. Sun, A note on Stewart's theorem for definite matrix pairs, *Linear Algebra Appl.*, **48**(1982), 331–339. MR **84f:**15013
18. ———, The perturbation bounds for eigenspaces of a definite matrix-pair, *Numer. Math.*, **41**(1983), 321–343. MR **85c:**65045
19. ———, The perturbation bounds of generalized eigenvalues of a class of matrix-pairs, *Math. Numer. Sinica*, **4**(1982), 23–29 (Chinese). MR **85h:**15021
20. ———, On the perturbation of the eigenvalues of a normal matrix, *Math. Numer. Sinica*, **6**(1984), 334-336 (Chinese). MR **86d:**15010
21. H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung), *Mathematische Annalen*, **71**(1912), 441–479.
22. H. Wielandt, An extremum property of sums of eigenvalues, *Proc. Amer. Math. Soc.*, **6**(1955), 106–110. MR **16:**785a

INDIAN STATISTICAL INSTITUTE, 7, S. J. S. SANSANWAL MARG, NEW DELHI – 110016, INDIA
E-mail address: rbh@isid.ernet.in

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT BERKELEY, BERKELEY, CALIFORNIA 94720
E-mail address: li@math.berkeley.edu