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 $\tilde{A}$ be two $n \times n$ Hermitian matrices, and let $\lambda_1, \ldots, \lambda_n$ and $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ be their eigenvalues arranged in ascending order. Then
\[
\left\| \text{diag}(\lambda_1 - \tilde{\lambda}_1, \ldots, \lambda_n - \tilde{\lambda}_n) \right\| \leq \left\| A - \tilde{A} \right\|
\]
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 $\tilde{A}$ be two $n \times n$ Hermitian matrices, and let $\lambda_1, \ldots, \lambda_n$ and $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n$ respectively be their eigenvalues arranged in ascending order. Then for any unitarily invariant norm $\| \cdot \|
(1.1)
\[
\left\| \text{diag}(\lambda_1 - \tilde{\lambda}_1, \ldots, \lambda_n - \tilde{\lambda}_n) \right\| \leq \left\| A - \tilde{A} \right\|
\]
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
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, $\mathbb{C}^m = \mathbb{C}^{m \times 1}$, $\mathbb{C}^1 = \mathbb{C}$; 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 Minsky [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 $A \in \mathbb{C}^{m \times n}$, 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 $\det(A - \lambda B) \neq 0$. Let

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

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 $(\alpha, \beta)$ of $G_{1,2}$ is called a generalized eigenvalue of a regular pencil $A - \lambda B$ if $\det(\beta A - \alpha B) = 0$. An element $(\alpha, \beta)$ of $G_{1,2}$ 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

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

Let $(\alpha, \beta), (\gamma, \delta) \in G_{1,2}$. The chordal metric

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

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

$$\|P_{Z^H} - P_{\tilde{Z}^H}\| \quad \text{and} \quad \|Z - \tilde{Z}\|,$$
where

\[ Z = (A, B) \quad \text{and} \quad \tilde{Z} = (\tilde{A}, \tilde{B}). \]

3. Main results

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

\[
\begin{align*}
&\begin{cases}
Y^HAX = \Lambda, \\
Y^HBX = \Omega,
\end{cases} \quad \text{and} \quad \begin{cases}
\tilde{Y}^H\tilde{A}\tilde{X} = \tilde{\Lambda}, \\
\tilde{Y}^H\tilde{B}\tilde{X} = \tilde{\Omega},
\end{cases}
\end{align*}
\]

where \( X, Y, \tilde{X}, \tilde{Y} \in \mathbb{C}^{n \times n} \) are nonsingular matrices,

\[
\begin{align*}
\begin{cases}
\Lambda = \text{diag}(\alpha_1, \ldots, \alpha_n), \\
\Omega = \text{diag}(\beta_1, \ldots, \beta_n),
\end{cases} \quad \text{and} \quad \begin{cases}
\tilde{\Lambda} = \text{diag}(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n), \\
\tilde{\Omega} = \text{diag}(\tilde{\beta}_1, \ldots, \tilde{\beta}_n),
\end{cases}
\end{align*}
\]

and \( \alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}, i, j = 1, \ldots, n \). Clearly,

\[
\begin{align*}
\lambda(A, B) &= \{(\alpha_i, \beta_i), i = 1, \ldots, n\}, \\
\lambda(\tilde{A}, \tilde{B}) &= \{(	ilde{\alpha}_j, \tilde{\beta}_j), j = 1, \ldots, n\}.
\end{align*}
\]

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

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

\[
\| \left( \begin{array}{cc}
\Sigma & \\
\Sigma & \end{array} \right) \| \leq \frac{\pi}{2} \tilde{\kappa}(X) \tilde{\kappa}(\tilde{X}) \| P_{Z\sigma} - P_{\tilde{Z}\sigma} \|,
\]

where

\[
\begin{align*}
\tilde{\kappa}(X) &= \max \{ \| X^{-1} \|_2, \| X \|_2 \}^2, \\
\tilde{\kappa}(\tilde{X}) &= \max \{ \| \tilde{X}^{-1} \|_2, \| \tilde{X} \|_2 \}^2, \\
\Sigma &= \text{diag}(\rho(\alpha_1, \beta_1), \rho(\alpha_2, \beta_2), \ldots, \rho(\alpha_n, \beta_n)).
\end{align*}
\]

In Theorem 3.1, the distance between the two pencils \( A - \lambda B \) and \( \tilde{A} - \lambda \tilde{B} \) is measured via \( \| P_{Z\sigma} - P_{\tilde{Z}\sigma} \| \). In the following theorem, it is measured more directly by \( \| Z - \tilde{Z} \| \).

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

\[
\alpha_i^2 + \beta_i^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1, \quad i, j = 1, \ldots, n,
\]

then there exists a permutation \( \sigma \) of \( \{1, \ldots, n\} \) such that

\[
\| \Sigma \| \leq \frac{\pi}{2} \min \{ \kappa(X) \| \tilde{X}^{-1} \|_2, \kappa(X) \| \tilde{X} \|_2, \kappa(X) \| \tilde{X}^{-1} \|_2, \kappa(X) \| \tilde{X} \|_2 \} \| Z - \tilde{Z} \|,
\]

where \( \Sigma \) is defined as in (3.4), and

\[
\kappa(X) = \| X \|_2 \| X^{-1} \|_2, \quad \kappa(\tilde{X}) = \| \tilde{X} \|_2 \| \tilde{X}^{-1} \|_2.
\]
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 $\|Y^H\|_2 \leq \|Z^+\|_2 \|X^{-1}\|_2$.

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

\[ \|\Sigma\| \leq \frac{\pi}{2} \kappa(X) \kappa(\bar{X}) \min\{\|Z^+\|_2, \|\bar{Z}^+\|_2\} \|Z - \bar{Z}\|, \]

where $\Sigma$ 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

\[ c(A, B) \overset{\text{def}}{=} \min\{|x^H(A + iB)x| : \|x\|_2 = 1\} > 0. \]

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

\[ X^HAX = \text{diag}\,(\alpha_1, \ldots, \alpha_n), \quad X^HBX = \text{diag}\,(\beta_1, \ldots, \beta_n). \]

In Lemma 3.2, it is easily verified that $\alpha_i, \beta_i \in \mathbb{R}$, 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

\[ \|X\|_2 \leq \frac{1}{\sqrt{c(A, B)}}, \quad \|X^{-1}\|_2 \leq \frac{\|Z\|_2}{\sqrt{c(A, B)}}. \]

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 $\bar{A} - \lambda \bar{B}$ be two definite pencils of order $n$. Then there exists a permutation $\sigma$ of \{1, \ldots, n\} such that for any unitarily invariant norm $\| \cdot \|$

\[ \left\| \left( \begin{array}{cc} \Sigma & \bar{X} \\ X & \Sigma \end{array} \right) \right\| \leq \frac{\pi}{2} \max\{1, \|Z\|_2\} \max\{1, \|\bar{Z}\|_2\} \frac{\|P_{Z^H} - P_{\bar{Z}^H}\|}{c(A, B)c(A, \bar{B})}, \]

where $\Sigma$ 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 $\sigma$ of \{1, \ldots, n\} such that

\[ \|\Sigma\| \leq \frac{\pi}{2} \min\{\|Z\|_2, \|\bar{Z}\|_2\} \|Z - \bar{Z}\|, \]

where $\Sigma$ 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 \( \tilde{A} - \lambda \tilde{B} \) be as described by (3.1) and (3.2). We have

\[
\tilde{\Lambda} X^{-1} \tilde{\Omega} - \tilde{\Omega} X^{-1} \Lambda = -(\tilde{\Lambda} X^{-1}, \tilde{\Omega} X^{-1})(P_{Z^T} - P_{Z^0}) \begin{pmatrix}
 X \Omega \\
 -X \Lambda
\end{pmatrix}
\]

(4.1)

\[
\Lambda X^{-1} \tilde{\Omega} - \Omega X^{-1} \tilde{\Lambda} = -(\Lambda X^{-1}, \Omega X^{-1})(P_{Z^T} - P_{Z^0}) \begin{pmatrix}
 \tilde{X} \Omega \\
 -\tilde{X} \Lambda
\end{pmatrix}
\]

(4.2)

where the constant \( \pi/2 \) is best possible.

**Proof.** Define \( F \equiv UT - \Gamma V \). Then \( VF^H U = VT - \Gamma U \), and

\[
(U - V)\Gamma + \Gamma(U - V) = F - VF^H U.
\]

(4.5)

Since \( U \) and \( V \) are unitary,

\[
\| F - VF^H U \| \leq \| F \| + \| VF^H U \| = 2 \| F \|.
\]

By [6, Theorem 5.2],

\[ 2 \| F \| \geq \|(U - V)\Gamma + \Gamma(U - V)\| \geq 2\|\Gamma^{-1}\|_2^{-1} \| U - V \|, \]

which proves (4.4).

**Lemma 4.2** ([10]). Suppose \( \alpha_i, \beta_j, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R} \) satisfy \( \alpha_i^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1 \), \( i, j = 1, \ldots, n \). Let \( \Lambda, \Omega, \tilde{\Lambda} \) and \( \tilde{\Omega} \) be as defined in (3.2), and let \( U \in U_n \). Then there exists a permutation \( \sigma \) of \( \{1, \ldots, n\} \) such that for any unitarily invariant norm \( \| \cdot \| \)

\[
\left\| \text{diag} \left( \rho(\alpha_1, \beta_1), (\tilde{\alpha}_1(1), \tilde{\beta}_1(1)), \ldots, \rho(\alpha_n, \beta_n), (\tilde{\alpha}_n(\sigma(n)), \tilde{\beta}_n(\sigma(n))) \right) \right\|
\]

\[
\leq \frac{\pi}{2} \left\| \tilde{\Lambda}U\Omega - \tilde{\Omega}U\Lambda \right\|
\]

where the constant \( \pi/2 \) is best possible.

**Lemma 4.4.** Let \( \alpha_i, \beta_j, \tilde{\alpha}_j, \tilde{\beta}_j \) and \( \Lambda, \Omega, \tilde{\Lambda}, \tilde{\Omega} \) be as described in Lemma 4.2, and let \( T \) be a nonsingular matrix. Then there exists a permutation \( \sigma \) of \( \{1, \ldots, n\} \) such that for any unitarily invariant norm \( \| \cdot \| \)

\[
\left\| \text{diag} \left( \rho(\alpha_1, \beta_1), (\tilde{\alpha}_1(1), \tilde{\beta}_1(1)), \ldots, \rho(\alpha_n, \beta_n), (\tilde{\alpha}_n(\sigma(n)), \tilde{\beta}_n(\sigma(n))) \right) \right\|
\]

\[
\leq \frac{\pi}{2} \|T^{-1}\|_2 \left\| \tilde{\Lambda}T\Omega - \tilde{\Omega}T\Lambda \right\|
\]

where the constant \( \pi/2 \) is best possible.
Proof. Let \( R = \tilde{\Lambda}\tilde{T} - \tilde{\Omega}\tilde{T} \), and let
\[
\begin{pmatrix}
L \\
M
\end{pmatrix} = \begin{pmatrix}
I & iI \\
iI & I
\end{pmatrix} \begin{pmatrix}
\Lambda \\
\Omega
\end{pmatrix},
\begin{pmatrix}
L \\
M
\end{pmatrix} = \begin{pmatrix}
I & iI \\
iI & I
\end{pmatrix} \begin{pmatrix}
\tilde{\Lambda} \\
\tilde{\Omega}
\end{pmatrix}.
\]
Then the diagonal matrices \( L, M, \tilde{L} \) and \( \tilde{M} \) are all in \( U_n \). Hence, we have \( \tilde{L}TM - \tilde{M}TL = 2R \), and this gives
\[
\tilde{M}^H\tilde{LT} - TL^M = 2\tilde{M}^HRM^H.
\]
(4.7)
Let \( T = UT^HV^H \) be the singular value decomposition of \( T \). From (4.7) it follows that
\[
U^H\tilde{M}^H\tilde{L}U - \Gamma V^HLM^HV = 2U^H\tilde{M}^HRM^HV. \tag{4.8}
\]
If \( \tilde{U} \overset{\text{def}}{=} U^H\tilde{M}^H\tilde{L}U \) and \( \tilde{V} \overset{\text{def}}{=} V^HLM^HV \), then \( \tilde{U} \) and \( \tilde{V} \) are unitary matrices. Therefore, from (4.8) and Lemma 4.3, we get
\[
2\left\| T^{-1}\right\|_2 \left\| R \right\| = \left\| \Gamma^{-1} \right\|_2 \left\| 2U^H\tilde{M}^HRM^HV \right\|
\geq \left\| U^H\tilde{M}^H\tilde{L}U - V^HLM^HV \right\| \quad \text{(Lemma 4.3)}
\geq \left\| \tilde{M}^H\tilde{L}U\tilde{V}^HV - UV^HLM^HV \right\|
\geq \left\| \tilde{L}U\tilde{V}^HV - \tilde{M}UV^HLM^HV \right\|
\geq 2\left\| \tilde{M}UV^H\tilde{L} - \tilde{\Omega}U\tilde{V}^H\Lambda \right\|.
\]
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
\[
\left\| \begin{pmatrix}
F_{11} & 0 \\
0 & F_{22}
\end{pmatrix} \right\| \leq \left\| \begin{pmatrix}
F_{11} & F_{12} \\
F_{21} & F_{22}
\end{pmatrix} \right\|
\]
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 \( \sigma \) of \( \{1, \ldots, n\} \) such that
\[
\| \Sigma \| \leq \frac{\pi}{2} \| X^{-1}X \|_2 \| E \|, \quad \| \Sigma \| \leq \frac{\pi}{2} \| X^{-1}X \|_2 \| E \|
\]
where Σ is as defined in (3.4). Hence by Lemma 4.5,

\[(4.9)\]
\[
\left\| \begin{pmatrix} \Sigma \\ \Sigma \end{pmatrix} \right\| \leq \max\{\|\bar{X}^{-1}X\|_2, \|X^{-1}\bar{X}\|_2\} \left\| \begin{pmatrix} E \\ \bar{E}^H \end{pmatrix} \right\|
\]
\[
\leq \max\{\|\bar{X}^{-1}X\|_2, \|X^{-1}\bar{X}\|_2\} \times \left\| \begin{pmatrix} \bar{X}^{-1} \bar{\Omega} \bar{X}^{-1} \\ \bar{\Omega}^H \bar{X}^{-1} \\ \bar{X}^{-1} \bar{\Omega} \\ \bar{\Omega}^H \end{pmatrix} \right\| (P_{Zu} - P_{\bar{Z}u}) \left( \begin{pmatrix} \bar{X} \Omega \\ X^{-H} \bar{\Lambda} \\ -X \bar{\Lambda} \end{pmatrix} \right)
\]
\[
\leq \max\{\|\bar{X}^{-1}X\|_2, \|X^{-1}\bar{X}\|_2\} \times \left\| \begin{pmatrix} \bar{X}^{-1} \bar{\Omega} \bar{X}^{-1} \\ \bar{\Omega}^H \bar{X}^{-1} \\ \bar{X}^{-1} \bar{\Omega} \\ \bar{\Omega}^H \end{pmatrix} \right\| \left\| P_{Zu} - P_{\bar{Z}u} \right\| \left\| \begin{pmatrix} \bar{X} \Omega \\ X^{-H} \bar{\Lambda} \\ -X \bar{\Lambda} \end{pmatrix} \right\|_2.
\]

The proof is completed by noting

\[(4.10)\]
\[
\left\| \begin{pmatrix} \bar{X}^{-1} \bar{\Omega} \bar{X}^{-1} \\ \bar{\Omega}^H \bar{X}^{-1} \\ \bar{X}^{-1} \bar{\Omega} \\ \bar{\Omega}^H \end{pmatrix} \right\|_2 \leq \max\{\|\bar{X}^{-1}\|_2, \|\bar{X}\|_2\},
\]
\[(4.11)\]
\[
\left\| \begin{pmatrix} \bar{X} \Omega \\ X^{-H} \bar{\Lambda} \\ -X \bar{\Lambda} \end{pmatrix} \right\|_2 \leq \max\{\|X^{-1}\|_2, \|X\|_2\},
\]

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]:

\[(4.12)\]
\[
\bar{X}^{-1}X\Omega - \bar{\Omega} \bar{X}^{-1}X\Lambda = -\bar{Y}^H(Z - \bar{Z}) \begin{pmatrix} \bar{X} \\ \Omega \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix},
\]
\[(4.13)\]
\[
\bar{\Lambda} \bar{X}^{-1}\bar{\Omega} - \Omega X^{-1}\bar{\Lambda} = -Y^H(\bar{Z} - Z) \begin{pmatrix} \bar{X} \\ \bar{\Omega} \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}.
\]

**Proof of Theorem 3.4.** By Lemma 3.2, we know that \(\Lambda - \lambda \bar{B}\) and \(\bar{\Lambda} - \bar{\lambda} \bar{B}\) admit decompositions

\[(4.14)\]
\[
\begin{cases} X^HAX = \Lambda, \\ X^HBX = \Omega, \end{cases} \quad \begin{cases} \bar{X}^H\bar{A}\bar{X} = \bar{\Lambda}, \\ \bar{X}^H\bar{B}\bar{X} = \bar{\Omega}. \end{cases}
\]

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

5. **Concluding remarks**

1. \(P_{Zu}\) is invariant under premultiplications of \(A\) and \(B\) by nonsingular matrices. Thus using \(P_{Zu} - P_{\bar{Z}u}\) has advantages in the case when \(\bar{\Lambda} - \bar{\lambda} \bar{B} \approx Q(\Lambda - \lambda B)\) for some nonsingular matrix \(Q\). In fact, if \(\bar{A} - \bar{\lambda} \bar{B} = Q(\Lambda - \lambda B)\), then \(P_{Zu} - P_{\bar{Z}u} = 0\) 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:

$$\left\|\text{diag}(\alpha_1 - \tilde{\alpha}_{\sigma(1)}, \ldots, \alpha_n - \tilde{\alpha}_{\sigma(n)})\right\| \leq \frac{\pi}{2} \left\| A - \tilde{A} \right\|,$$

(5.1)

where $A$ and $\tilde{A}$ are unitary matrices having eigenvalues $\{\alpha_i\}_{i=1}^{n}$ and $\{\tilde{\alpha}_i\}_{i=1}^{n}$, respectively, and $\sigma$ 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 $\infty$-norm, while the Frobenius norm is the Schatten 2-norm. So we conjecture that for the Schatten $p$-norms ($2 \leq p \leq \infty$), 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


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