

**SOLUTION OF LINEAR MATRIX EQUATION  $AXD - BXC = S$   
AND PERTURBATION OF EIGENSPACES OF A MATRIX PENCIL. I†**

REN-CANG LI

INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS  
P. O. Box 8009  
BEIJING 100088, P. R. CHINA

ABSTRACT

Let  $A - \lambda B$  and  $C - \lambda D$  be two normal matrix pencils or Hermitian matrix pencils or definite matrix pencils, and let  $AXD - BXC = S$ . Our main results in this part are estimates of a unitarily invariant norm of the solution  $X$  under some conditions on the spectra of two pencils.

§1. INTRODUCTION

The first problem treated in this paper is to bound unitarily invariant norms of the solution  $X$  of linear matrix equation

$$AXD - BXC = S \tag{1.1}$$

under some conditions on the spectra of matrix pencils  $A - \lambda B$  and  $C - \lambda D$  (for definitions see §2.2 below). As a special case, when  $B$  and  $D$  are both unit matrices, (1.1) becomes popular and has been attracting a lot of attention in literature, see e.g., Lancaster[12].

We devote three sections to discuss the problem. The case of strong hypothesis on the separation of the spectra of the pencils is treated in §3, while §4 deals with more general case under the assumption that the spectra of the pencils lie in a band region on Riemann sphere, which is indeed satisfied in some cases for example both spectra are real. As a special unitarily invariant norm, Frobenius norm appeals often to researchers owing to its own properties. Here it is also the case. In §5, we show that the situation is rather special for the case of Frobenius norm.

The other problem we treat is the eigenspace perturbation problem. In searching for bounds, we rely on the construction of perturbation equations in Li[15] and on the application of results obtained in §§3 and 4. Doing in this way, we can get several  $\sin \theta$  theorems.

---

†Published in *Journal of Computational Mathematics*, **10**: Supplementary Issue (1992), 93–109.

(Not all these theorems are written out, instead, we will explain in detail how to develop them.) It is worth mentioning these  $\sin \theta$  theorems cover the complex generalized eigenvalue case whereas theorems of Li[15] do not, and some theorems here are valid under weaker hypotheses.

A real generalized eigenvalue has a natural geometric representation as described in Li[14, 15]. Such a representation enables us to draw sophisticated bounds for the perturbation of a diagonalizable matrix pencil with real spectrum and of a Hermitian matrix pencil[14], e.g., we can prove that there is a one-one pairing of the generalized eigenvalues with the perturbed eigenvalues and uniform bounds for their differences (in the sense of the chordal metric). For the complex spectrum case, the method used in [14] does not work owing to the difficulty of pairing complex generalized eigenvalues. We, relying on our foregoing results, will give in §7 a bound for the perturbation of a diagonalizable matrix pencil which, generally, may have complex generalized eigenvalues.

§2. PRELIMINARIES

**2.1 Notation and Essential Definitions.** Throughout the paper, capital letters are for matrices, lowercase Latin letters for column vectors or scalars and lowercase Greek letters for scalars;  $\mathbf{C}^{m \times n}$  for the set of  $m \times n$  complex matrices,  $\mathcal{U}_n \subset \mathbf{C}^{n \times n}$  for the set of  $n \times n$  unitary matrices,  $\mathbf{C}^m = \mathbf{C}^{m \times 1}$ ,  $\mathbf{C} = \mathbf{C}^1$ ,  $\mathbb{R}$  is the real number set. The symbol  $I^{(n)}$  stands for the  $n \times n$  unit matrix (also we just write  $I$  for convenience when no confusion arises).  $A > 0$  ( $A \geq 0$ ) denotes that  $A$  is a positive definite (positive semidefinite) Hermitian matrix, and  $A > B$  ( $A \geq B$ ) means  $A - B > 0$  ( $A - B \geq 0$ ).  $A^T$ ,  $A^H$ , and  $A^+$  denote the transpose, conjugate transpose, and Moore-Penrose inverse of  $A$ , respectively.  $\mathcal{R}(X)$  is the column space, the subspace spanned by the column vectors of  $X$ , and  $P_X$  is the orthogonal projection onto the column space  $\mathcal{R}(X)$ . It is easy to verify that (Sun[23, pp.38-41])

$$P_X = XX^+, \quad P_{X^H} = X^+X. \tag{2.1}$$

We denote by  $\|A\|_2$  and  $\|A\|_F$  the spectral norm and Frobenius norm of  $A \triangleq (a_{ij}) \in \mathbf{C}^{m \times n}$  respectively, i.e.,

$$\begin{aligned} \|A\|_2 &\triangleq \left( \text{the maximum eigenvalue of } A^H A \right)^{\frac{1}{2}} \\ &= \max_{\substack{x \in \mathbf{C}^n \\ \|x\|_2=1}} \|Ax\|_2, \\ \|A\|_F &\triangleq \left( \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Here and later,  $\triangleq$  denotes a definition and  $\|x\|_2$  is the Euclidean length of vector  $x = (\xi_1, \dots, \xi_n)^T \in \mathbf{C}^n$

$$\|x\|_2 \triangleq \left( \sum_{i=1}^n |\xi_i|^2 \right)^{\frac{1}{2}}.$$

More generally, we will consider unitarily invariant norms  $\|\cdot\|$  of matrices. In this we follow Sun[23]. To say that the norm is unitarily invariant on  $\mathbf{C}^{m \times n}$  means it satisfies besides the usual properties of any norm, also

$$(1) \|UAV\| = \|A\|, \text{ for any } U \in \mathcal{U}_m, \text{ and } V \in \mathcal{U}_n.$$

(2)  $\|A\| = \|A\|_2$ , for any  $A \in \mathbf{C}^{m \times n}$ ,  $\text{rank } A = 1$ .

Two unitarily invariant norms used frequently are  $\|\cdot\|_2$  and  $\|\cdot\|_F$ .

It is well-known that any unitarily invariant norm  $\|\cdot\|$  on  $\mathbf{C}^{m \times n}$  corresponds to an symmetric gauge function  $\Phi(\xi_1, \dots, \xi_N)$ , where  $N = \min\{m, n\}$ , and vice versa. By extension according to this property, we can define a unitarily invariant norm  $\|\cdot\|$  on  $\mathbf{C}^{m_1 \times n_1}$  ( $m_1 \leq n$ ,  $n_1 \leq n$ ) consistent with the original one as  $\|A\| = \Phi(\sigma_1, \dots, \sigma_{N_1}, 0, \dots, 0)$  if  $A \in \mathbf{C}^{m_1 \times n_1}$  with singular values  $\sigma_1, \dots, \sigma_{N_1}$  ( $N_1 = \min\{m_1, n_1\}$ ). In this paper, very often matrices with different dimensions enter our arguments together, so we make an agreement: assume we first have a matrix space with sufficiently large dimension  $M \times N$  and with a unitarily invariant norm  $\|\cdot\|$  on it, then by the extension mentioned, on every matrix space with dimension smaller exists the extended unitarily invariant norm denoted also by  $\|\cdot\|$ . In this way, we have (see Sun[23, pp.319-321])

$$\|CD\| \leq \begin{cases} \|C\|_2 \|D\| \\ \|C\| \|D\|_2 \end{cases}, \text{ for any } C \in \mathbf{C}^{m \times n}, D \in \mathbf{C}^{n \times \ell}. \quad (2.2)$$

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

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

$(\alpha, \beta) \in \mathbf{G}_{1,2}$  is called a *generalized eigenvalue* of regular pencil  $A - \lambda B$  if  $\det(\beta A - \alpha B) = 0$ . Nonzero vectors  $x, y \in \mathbf{C}^n$  are termed *the generalized eigenvector (the right generalized eigenvector at whites)* and *the left generalized eigenvector* corresponding to  $(\alpha, \beta)$ , respectively, if

$$\beta Ax = \alpha Bx, \beta y^H A = \alpha y^H B.$$

It is easy to see that if  $(\alpha, \beta) \in \mathbf{G}_{1,2}$  is a generalized eigenvalue of  $A - \lambda B$  so is  $(\xi\alpha, \xi\beta)$  for any complex number  $\xi \neq 0$ .  $(\alpha, \beta) \in \mathbf{G}_{1,2}$  is said to be *real* if there exists  $0 \neq \xi \in \mathbf{C}$  such that  $\xi\alpha, \xi\beta \in \mathbb{R}$ , for an instance  $(i, i)$  is real. The spectrum of regular pencil  $A - \lambda B$  consists of its all generalized eigenvalues (counted according to their algebraic multiplicities), and is denoted by  $\lambda(A, B)$ . Let  $(\alpha, \beta), (\gamma, \delta) \in \mathbf{G}_{1,2}$ , we use *the chordal metric* on Riemann sphere

$$\rho((\alpha, \beta), (\gamma, \delta)) \triangleq \frac{|\delta\alpha - \gamma\beta|}{\sqrt{|\alpha|^2 + |\beta|^2} \sqrt{|\gamma|^2 + |\delta|^2}}$$

to measure the difference between the two points.

**2.2 Several Classes of Pencils.** The following listed classes of regular matrix pencils were introduced in the past few years by several numerical analysts in order to develop a perturbation theory for the eigenvalue problem of matrix pencils.

Elsner & Sun[18] introduced the classes of diagonalizable (normalizable) matrix pencils so that they could establish a Bauer-Fike type theorem and they indeed succeeded.

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

$$\begin{cases} Y^H AX = \Lambda \triangleq \text{diag}(\alpha_1, \dots, \alpha_n) \\ Y^H BX = \Omega \triangleq \text{diag}(\beta_1, \dots, \beta_n) \end{cases}. \quad (2.3)$$

We denote by  $\mathbf{D}_g(n)$  the set of  $n \times n$  diagonalizable pencils of order  $n$ .

**THEOREM 2.1.**  $A - \lambda B$  is as described in Definition 2.1, and moreover  $|\alpha_i|^2 + |\beta_i|^2 = 1$ ,  $i = 1, 2, \dots, n$ . Then

$$\begin{cases} \|Y\|_2 \leq \|X^{-1}\|_2 \|Z^+\|_2 \\ \|Y^{-1}\|_2 \leq \|X\|_2 \|Z\|_2 \end{cases} \quad \text{and} \quad \begin{cases} \|X\|_2 \leq \|Y^{-1}\|_2 \|Z_d^+\|_2 \\ \|X^{-1}\|_2 \leq \|Y\|_2 \|Z_d\|_2 \end{cases}, \quad (2.4)$$

where  $Z = (A, B)$  and  $Z_d = \begin{pmatrix} A \\ B \end{pmatrix}$ .

**PROOF:** As an example, we show  $\|Y\|_2 \leq \|X^{-1}\|_2 \|Z^+\|_2$ . To this end, we note  $ZZ^+ = I$ , and thus

$$\begin{aligned} Y^H Z \begin{pmatrix} X & \\ & X \end{pmatrix} &= (\Lambda, \Omega) \\ \Rightarrow Y^H &= (\Lambda, \Omega) \begin{pmatrix} X^{-1} & \\ & X^{-1} \end{pmatrix} Z^+ \\ \Rightarrow \|Y\|_2 &\leq \|(\Lambda, \Omega)\|_2 \left\| \begin{pmatrix} X^{-1} & \\ & X^{-1} \end{pmatrix} \right\|_2 \|Z^+\|_2 \\ &= \|X^{-1}\|_2 \|Z^+\|_2, \end{aligned}$$

since  $\|(\Lambda, \Omega)\|_2 = \max_{1 \leq i \leq n} \sqrt{|\alpha_i|^2 + |\beta_i|^2} = 1$ . ■

The generalization of the noted Hoffman-Wielandt theorem[11] leads Sun[20] to define the class of normal matrix pencils. Later, Li[13, 14, 15] went further to introduce left normal matrix pencil, Hermitian matrix pencil and left Hermitian matrix pencil. The definition of Hermitian matrix pencil gives another choice of counterparts of Hermitian matrix in the standard eigenvalue problem. (Stewart[19] and Sun[22, 23] had already given a choice—Definite pencil.)

**DEFINITION 2.2.** A regular matrix pencil  $A - \lambda B$  of order  $n$  is called a normal matrix pencil of order  $n$ , if there exist an invertible matrix  $Y \in \mathbb{C}^{n \times n}$  and  $V \in \mathcal{U}_n$  such that

$$\begin{cases} Y^H A V = \Lambda \triangleq \text{diag}(\alpha_1, \dots, \alpha_n) \\ Y^H B V = \Omega \triangleq \text{diag}(\beta_1, \dots, \beta_n) \end{cases}. \quad (2.5)$$

If, in addition,  $Y, \Lambda, \Omega, V$  can be chosen such that  $\alpha_i, \beta_i \in \mathbb{R}$  (or in other words  $(\alpha_i, \beta_i)$  is real),  $i = 1, 2, \dots, n$ , then  $A - \lambda B$  is termed a Hermitian matrix pencil of order  $n$ .

**DEFINITION 2.3.** A regular matrix pencil  $A - \lambda B$  of order  $n$  is called a left normal matrix pencil of order  $n$ , if there exist  $U \in \mathcal{U}_n$  and an invertible matrix  $X \in \mathbb{C}^{n \times n}$  such that

$$\begin{cases} U^H A X = \Lambda \triangleq \text{diag}(\alpha_1, \dots, \alpha_n) \\ U^H B X = \Omega \triangleq \text{diag}(\beta_1, \dots, \beta_n) \end{cases}. \quad (2.6)$$

If, in addition,  $U, \Lambda, \Omega, X$  can be chosen such that  $\alpha_i, \beta_i \in \mathbb{R}$  (or in other words  $(\alpha_i, \beta_i)$  is real),  $i = 1, 2, \dots, n$ , then  $A - \lambda B$  is termed a left Hermitian matrix pencil of order  $n$ .

Sometimes, normal pencil is also called *right normal pencil* in contrast with left normal pencil, and Hermitian pencil is also called *right Hermitian pencil* in contrast with left Hermitian pencil. We denote by  $\mathbf{H}_R(n)$  and  $\mathbf{H}_L(n)$  the set of Hermitian pencils of order  $n$  and

the set of left Hermitian pencils of order  $n$ , respectively, denote by  $\mathbf{N}_R(n)$  and  $\mathbf{N}_L(n)$  the set of normal pencils of order  $n$  and the set of left normal pencils of order  $n$ , respectively. Besides, we write

$$\mathbf{H}(n) = \mathbf{H}_R(n) \cap \mathbf{H}_L(n), \quad \mathbf{N}(n) = \mathbf{N}_R(n) \cap \mathbf{N}_L(n). \quad (2.7)$$

**THEOREM 2.2.** Let  $Z = (A, B) \in \mathbf{C}^{n \times 2n}$  and  $Z_d = \begin{pmatrix} A \\ B \end{pmatrix} \in \mathbf{C}^{2n \times n}$ .

- (i)  $A - \lambda B \in \mathbf{H}_R(n) \iff A^H - \lambda B^H \in \mathbf{H}_L(n) \iff A^T - \lambda B^T \in \mathbf{H}_L(n)$ ;
- (ii)  $A - \lambda B \in \mathbf{N}_R(n) \iff A^H - \lambda B^H \in \mathbf{N}_L(n) \iff A^T - \lambda B^T \in \mathbf{N}_L(n)$ ;
- (iii)  $A - \lambda B \in \mathbf{H}_R(n) \iff (ZZ^H)^{-\frac{1}{2}}(A - \lambda B) \in \mathbf{H}(n)$ ,  
 $A - \lambda B \in \mathbf{H}_L(n) \iff (A - \lambda B)(Z_d^H Z_d)^{-\frac{1}{2}} \in \mathbf{H}(n)$ ;
- (iv)  $A - \lambda B \in \mathbf{N}_R(n) \iff (ZZ^H)^{-\frac{1}{2}}(A - \lambda B) \in \mathbf{N}(n)$ ,  
 $A - \lambda B \in \mathbf{N}_L(n) \iff (A - \lambda B)(Z_d^H Z_d)^{-\frac{1}{2}} \in \mathbf{N}(n)$ .

**PROOF:** As an example, we show the first one of (iii) to be true. Suppose  $A - \lambda B \in \mathbf{H}_R(n)$  admits (2.5), and without loss of generality, we assume that  $\alpha_i, \beta_i \in \mathbb{R}$ ,  $|\alpha_i|^2 + |\beta_i|^2 = 1$ ,  $i = 1, 2, \dots, n$ . It is easy to see  $ZZ^H = Y^{-H}Y^{-1}$ , and thus

$$Y^H(YY^H)^{-\frac{1}{2}}(ZZ^H)^{-\frac{1}{2}}AV = \Lambda, \quad Y^H(YY^H)^{-\frac{1}{2}}(ZZ^H)^{-\frac{1}{2}}BV = \Omega.$$

Because of  $(YY^H)^{-\frac{1}{2}}Y \in \mathcal{U}_n$ , we have proved  $(ZZ^H)^{-\frac{1}{2}}(A - \lambda B) \in \mathbf{H}(n)$ . On the other hand,  $\Leftarrow$  is evident.  $\blacksquare$

Definitions 2.2–2.3 have their own equivalent statements in geometry (see Sun[23, Chapter 4]). And there are easy ways to verify whether a matrix pencil is normal (Hermitian) or not (see Li[14] and Sun[20, 23]).

Another important class of matrix pencils is that of definite pencils, the perturbation of which eigenvalues was first studied by Crawford[7], but a thorough study of the perturbation theory for the generalized eigenvalue problem of definite pencils was not done until the appearance of [19], [21, 22], [14, 15]. We may say now that a little well-developed theory has been established for the perturbation of the generalized eigenvalue problem of definite pencils.

**DEFINITION 2.4.** Let Hermitian matrices  $A, B \in \mathbf{C}^{n \times n}$ .  $A - \lambda B$  is said to be a definite pencil of order  $n$ , if

$$c(A, B) \triangleq \min_{\substack{x \in \mathbf{C}^n \\ \|x\|_2=1}} |x^H(A + \sqrt{-1}B)x| > 0. \quad (2.8)$$

$c(A, B)$  is called Crawford number of the definite pencil  $A - \lambda B$ .  $\mathbf{D}(n)$  denotes the set of all definite pencils of order  $n$ .

Definitions 2.1–2.3 imply that (left) normal pencils and (left) Hermitian pencils are all diagonalizable. The following theorem shows that it is also the case for definite pencils.

**THEOREM 2.3.** Let  $A - \lambda B \in \mathbf{D}(n)$ . Then there is a nonsingular matrix  $X \in \mathbf{C}^{n \times n}$  such that

$$\begin{cases} X^H A X = \text{diag}(\alpha_1, \dots, \alpha_n), \\ X^H B X = \text{diag}(\beta_1, \dots, \beta_n). \end{cases} \quad (2.9)$$

In Theorem 2.3, it is easily verified that  $\alpha_i, \beta_i \in \mathbb{R}$ , and by appropriate choice of  $X$ , we can make  $|\alpha_i|^2 + |\beta_i|^2 = 1$ . Readers are referred to Stewart[19] or Sun[23, pp.244-246] for a proof of Theorem 2.3.

THEOREM 2.4. In Theorem 2.3, if  $|\alpha_i|^2 + |\beta_i|^2 = 1, i = 1, 2, \dots, n$ , then

$$\|X\|_2 \leq \frac{1}{\sqrt{c(A, B)}}, \|X^{-1}\|_2 \leq \frac{\|Z\|_2}{\sqrt{c(A, B)}}, \tag{2.10}$$

where  $Z = (A, B)$ .

**2.3. Existence and Uniqueness of Solution of the Equation  $AXD - BXC = S$ .**

THEOREM 2.5. Let  $A - \lambda B$  and  $C - \lambda D$  be two regular matrix pencils of orders  $m, n$ , respectively. If  $\lambda(A, B) \cap \lambda(C, D) = \emptyset$ , then the equation  $AXD - BXC = S$  has a unique solution  $X \in \mathbb{C}^{m \times n}$ .

A proof of Theorem 2.5 can be found in Chu [5, Theorem 3].

§3. BOUNDS UNDER STRONG HYPOTHESES

For the sake of convenience, we henceforth write  $A - \lambda B$  and  $C - \lambda D$  for two regular matrix pencils, and denote by

$$Z = (A, B), \quad Z_d = \begin{pmatrix} A \\ B \end{pmatrix}, \tag{3.1}$$

$$W = (C, D), \quad W_d = \begin{pmatrix} C \\ D \end{pmatrix}. \tag{3.2}$$

For the use of this section only, we assign numbers  $\xi, \delta, \eta$  and a point  $(\alpha_0, \beta_0)$  satisfying

$$\begin{cases} (\alpha_0, \beta_0) \in \mathbf{G}_{1,2}, & |\alpha_0|^2 + |\beta_0|^2 = 1, \\ \xi > 0, \delta > 0, \xi + \delta \leq 1, \\ \eta \triangleq (\xi + \delta)\sqrt{1 - \xi^2} - \xi\sqrt{1 - (\xi + \delta)^2} > 0. \end{cases} \tag{3.3}$$

The following lemma is the most essential results of the section.

LEMMA 3.1. Suppose  $A - \lambda B \in \mathbf{N}(m), C - \lambda D \in \mathbf{N}(n)$  and  $ZZ^H = I$  and  $WW^H = I$ . If

$$\begin{cases} \|\beta_0 A - \alpha_0 B\|_2 \leq \xi, \\ \|(\beta_0 C - \alpha_0 D)^{-1}\|_2 \leq (\xi + \delta)^{-1}, \end{cases} \tag{3.4}$$

then the equation  $AXD - BXC = S$  has a unique solution  $X \in \mathbb{C}^{m \times n}$ , and

$$\eta \|X\| \leq \|S\|. \tag{3.5}$$

PROOF: Without loss of generality, we assume that  $(\alpha_0, \beta_0) = (1, 0)$ , otherwise let

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} \beta_0 I & -\alpha_0 I \\ \bar{\alpha}_0 I & \bar{\beta}_0 I \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix}, \quad \begin{pmatrix} C_1 \\ D_1 \end{pmatrix} = \begin{pmatrix} \beta_0 I & -\alpha_0 I \\ \bar{\alpha}_0 I & \bar{\beta}_0 I \end{pmatrix} \begin{pmatrix} C \\ D \end{pmatrix},$$

then it is easy to verify that  $A_1 - \lambda B_1 \in \mathbf{N}(m), C_1 - \lambda D_1 \in \mathbf{N}(n), A_1 A_1^H + B_1 B_1^H = I$  and  $C_1 C_1^H + D_1 D_1^H = I$ , and moreover

$$\|A_1\|_2 \leq \xi, \quad \|C_1^{-1}\|_1 \leq (\xi + \delta)^{-1}$$

and  $A_1XD_1 - B_1XC_1 = S$ .

Let  $(\alpha_0, \beta_0) = (1, 0)$ . We claim that

$$\|B^{-1}\|_2 \leq \frac{1}{\sqrt{1 - \xi^2}}. \quad (3.6)$$

In fact, from Definition 2.2 follows that there exist  $U, V \in \mathcal{U}_m$  and

$$\begin{cases} \Lambda_1 = \text{diag}(\alpha_1, \dots, \alpha_n), \\ \Omega_1 = \text{diag}(\beta_1, \dots, \beta_n), \end{cases} \quad (3.7)$$

such that

$$A = U\Lambda_1V^H, \quad B = U\Omega_1V^H. \quad (3.8)$$

$ZZ^H = I$  yields  $|\alpha_i|^2 + |\beta_i|^2 = 1$ ,  $i = 1, 2, \dots, n$ . Thus from (3.4) and (3.7) and (3.8) we have

$$\begin{aligned} \|A\|_2 &= \max_{1 \leq i \leq m} |\alpha_i| \leq \xi \\ \Rightarrow |\beta_i|^2 &= 1 - |\alpha_i|^2 \geq 1 - \xi^2 \\ \Rightarrow \|B^{-1}\|_2 &= \max_{1 \leq i \leq m} \frac{1}{|\beta_i|} \leq \frac{1}{\sqrt{1 - \xi^2}}. \end{aligned}$$

Similarly, we also have

$$\|D\|_2 \leq \sqrt{1 - (\xi + \delta)^2}. \quad (3.9)$$

On the other hand,  $\lambda(A, B) \cap \lambda(C, D) = \emptyset$  can be proved easily, so the equation has a unique solution  $X$ . Now by (3.4), (3.6) and (3.9), we see

$$\begin{aligned} X &= B^{-1}(AXD - S)C^{-1} \\ \Rightarrow \|X\| &\leq \|B^{-1}\|_2 (\|A\|_2 \|X\| \|D\|_2 - \|S\|) \|C^{-1}\|_2 \\ &\leq \frac{1}{(\xi + \delta)\sqrt{1 - \xi^2}} (\xi\sqrt{1 - (\xi + \delta)^2} \|X\| + \|S\|) \\ \Rightarrow \eta \|X\| &\leq \|S\|. \quad \blacksquare \end{aligned}$$

Inequality (3.4), as a matter of fact, is nothing but a description upon the separation of  $\lambda(A, B)$  and  $\lambda(C, D)$  (see Proposition 3.1 below). For the case of only real generalized eigenvalues appearing, this description was first proposed and used by Sun[22] to derive perturbation bounds for eigenspaces of a definite pencil, and its geometric interpretation was discussed in Li[14, 15] in some detail. It is worth mentioning description (3.4) covers the complex generalized eigenvalue case.

**PROPOSITION 3.1.** *The conditions and notation are as described in Lemma 3.1, then (3.4) holds if and only if*

$$\max_{(\mu, \nu) \in \lambda(A, B)} \rho((\alpha_0, \beta_0), (\mu, \nu)) \leq \xi, \quad (3.10a)$$

$$\min_{(\omega, \tau) \in \lambda(C, D)} \rho((\alpha_0, \beta_0), (\omega, \tau)) \geq \xi + \delta. \quad (3.10b)$$

*Remark 3.1.* The normalization condition  $ZZ^H = I$  and  $WW^H = I$  in Lemma 3.1 can be replaced by  $Z_d^H Z_d = I$  and  $W_d^H W_d = I$ . In fact, it is not difficult to show that if  $A - \lambda B \in \mathbf{N}(m)$  then  $ZZ^H = I \iff Z_d^H Z_d = I$ .

Relying on Lemma 3.1, we now can give several results. For the sake of convenience, we list the following four conditions under which our bounds are derived.

$$\|(ZZ^H)^{-\frac{1}{2}}(\beta_0 A - \alpha_0 B)\|_2 \leq \xi, \tag{3.11}$$

$$\|(\beta_0 A - \alpha_0 B)(Z_d^H Z_d)^{-\frac{1}{2}}\|_2 \leq \xi, \tag{3.12}$$

$$\left\| \left[ (WW^H)^{-\frac{1}{2}}(\beta_0 C - \alpha_0 D) \right]^{-1} \right\|_2 \leq (\xi + \delta)^{-1}, \tag{3.13}$$

$$\left\| \left[ (\beta_0 C - \alpha_0 D)(W_d^H W_d)^{-\frac{1}{2}} \right]^{-1} \right\|_2 \leq (\xi + \delta)^{-1}. \tag{3.14}$$

These inequalities are, essentially, separation conditions imposed upon the spectra of the pencils  $A - \lambda B$  and  $C - \lambda D$ . In fact, similarly to Proposition 3.1, we have

PROPOSITION 3.2. (3.11)  $\iff$  (3.10a) if  $A - \lambda B \in \mathbf{N}_R(m)$ ; (3.12)  $\iff$  (3.10a) if  $A - \lambda B \in \mathbf{N}_L(m)$ ; (3.13)  $\iff$  (3.10b) if  $C - \lambda D \in \mathbf{N}_R(n)$ ; (3.14)  $\iff$  (3.10b) if  $C - \lambda D \in \mathbf{N}_L(n)$ .

THEOREM 3.1. Let  $AXD - BXC = S$ .

(1) If  $A - \lambda B \in \mathbf{N}_R(m)$ ,  $C - \lambda D \in \mathbf{N}_L(n)$  and (3.11) and (3.14) hold, the the equation has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and

$$\eta \|X\| \leq \left\| (ZZ^H)^{-\frac{1}{2}} S (W_d^H W_d)^{-\frac{1}{2}} \right\|. \tag{3.15}$$

(2) If  $A - \lambda B \in \mathbf{N}_R(m)$ ,  $C - \lambda D \in \mathbf{N}_R(n)$  and (3.11) and (3.13) hold, the the equation has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and

$$\eta \|X\| \leq \left\| (ZZ^H)^{-\frac{1}{2}} S \right\| \| (W^H W)^{-\frac{1}{2}} \|_2. \tag{3.16}$$

(3) If  $A - \lambda B \in \mathbf{N}_L(m)$ ,  $C - \lambda D \in \mathbf{N}_L(n)$  and (3.12) and (3.14) hold, the the equation has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and

$$\eta \|X\| \leq \| (Z_d^H Z_d)^{-\frac{1}{2}} \|_2 \left\| S (W_d^H W_d)^{-\frac{1}{2}} \right\|. \tag{3.17}$$

(4) If  $A - \lambda B \in \mathbf{N}_L(m)$ ,  $C - \lambda D \in \mathbf{N}_R(n)$  and (3.12) and (3.13) hold, the the equation has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and

$$\eta \|X\| \leq \| (Z_d^H Z_d)^{-\frac{1}{2}} \|_2 \|S\| \| (W W^H)^{-\frac{1}{2}} \|_2. \tag{3.18}$$

PROOF: As an example, we show (3.16), and other inequalities may be proved in a similar way. It follows from Theorem 2.2 and  $A - \lambda B \in \mathbf{N}_R(m)$  and  $C - \lambda D \in \mathbf{N}_R(n)$  that

$$(ZZ^H)^{-\frac{1}{2}}(A - \lambda B) \in \mathbf{N}(m), \quad (WW^H)^{-\frac{1}{2}}(C - \lambda D) \in \mathbf{N}(n).$$

Thus bearing (3.11) and (3.13) and Lemma 3.1 in mind, we have

$$\begin{aligned} & (ZZ^H)^{-\frac{1}{2}} A \left[ X (W W^H)^{\frac{1}{2}} \right] (W W^H)^{-\frac{1}{2}} D - (ZZ^H)^{-\frac{1}{2}} B \left[ X (W W^H)^{\frac{1}{2}} \right] (W W^H)^{-\frac{1}{2}} C \\ & = (ZZ^H)^{-\frac{1}{2}} S \\ \Rightarrow & \eta \left\| X (W W^H)^{\frac{1}{2}} \right\| \leq \left\| (ZZ^H)^{-\frac{1}{2}} S \right\| \\ \Rightarrow & \eta \|X\| \leq \eta \left\| X (W W^H)^{\frac{1}{2}} \right\| \| (W W^H)^{-\frac{1}{2}} \|_2 \\ & \leq \left\| (ZZ^H)^{-\frac{1}{2}} S \right\| \| (W W^H)^{-\frac{1}{2}} \|_2. \quad \blacksquare \end{aligned}$$

Generally, for two diagonalizable pencils, we have

THEOREM 3.2. Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $C - \lambda D \in \mathbf{D}_g(n)$  admit decompositions

$$\begin{cases} Y_1^H A X_1 = \Lambda_1 \\ Y_1^H B X_1 = \Omega_1 \end{cases} \quad \text{and} \quad \begin{cases} Y_2^H C X_2 = \Lambda_2 \\ Y_2^H D X_2 = \Omega_2 \end{cases}, \quad (3.19a)$$

where nonsingular matrices  $X_1, Y_1 \in \mathbf{C}^{m \times m}$ ,  $X_2, Y_2 \in \mathbf{C}^{n \times n}$ ,

$$\begin{cases} \Lambda_1 = \text{diag}(\alpha_1, \dots, \alpha_m) \\ \Omega_1 = \text{diag}(\beta_1, \dots, \beta_m) \end{cases} \quad \text{and} \quad \begin{cases} \Lambda_2 = \text{diag}(\gamma_1, \dots, \gamma_n) \\ \Omega_2 = \text{diag}(\delta_1, \dots, \delta_n) \end{cases}. \quad (3.19b)$$

If  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1$ ,  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, n$  and (3.10) holds, then the unique solution  $X \in \mathbf{C}^{m \times n}$  to the equation  $AXD - BXC = S$  satisfies

$$\eta \|X\| \leq \|X_1\|_2 \|Y_1^H S X_2\| \|Y_2\|_2. \quad (3.20)$$

PROOF: It follows from the hypotheses of the theorem that  $\Lambda_1 - \lambda \Omega_1 \in \mathbf{N}(m)$ ,  $\Lambda_2 - \lambda \Omega_2 \in \mathbf{N}(n)$  and  $\lambda(A, B) \cap \lambda(C, D) = \emptyset$ , so the equation has a unique solution  $X$ . On the other hand  $AXD - BXC = S$  produces

$$\begin{aligned} \Lambda_1(X_1^{-1} X Y_2^{-H}) \Omega_2 - \Omega_1(X_1^{-1} X Y_2^{-H}) \Lambda_2 &= Y_1^H S X_2 \\ \Rightarrow \eta \|X_1^{-1} X Y_2^{-H}\| &\leq \|Y_1^H S X_2\| \quad (\text{by (3.10) and Lemma 3.1}) \\ \Rightarrow \eta \|X\| &\leq \eta \|X_1\|_2 \|X_1^{-1} X Y_2^{-H}\| \|Y_2^H\|_2 \\ &\leq \|X_1\|_2 \|Y_1^H S X_2\| \|Y_2\|_2. \quad \blacksquare \end{aligned}$$

Combining Theorem 2.1 and Theorem 3.2 leads to

COROLLARY 3.1. The conditions and notation are as described in Theorem 3.2, then

$$\begin{aligned} \eta \|X\| &\leq \kappa(X_1) \kappa(X_2) \|Z^+\|_2 \|W^+\|_2 \|S\|, \\ \eta \|X\| &\leq \kappa(Y_1) \kappa(Y_2) \|Z_d^+\|_2 \|W_d^+\|_2 \|S\|. \end{aligned} \quad (3.21)$$

Here  $\kappa(X) = \|X\|_2 \|X^{-1}\|_2$  is the spectral condition number of nonsingular matrix  $X$ . Also, combining Theorem 2.4 and Theorem 3.2 yields

THEOREM 3.3. Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $C - \lambda D \in \mathbf{D}(n)$ , and suppose (3.10) holds, then the equation  $AXD - BXC = S$  has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and

$$\eta \|X\| \leq \frac{\|S\|}{c(A, B) c(C, D)}. \quad (3.22)$$

PROOF: By Theorem 2.4, we can suppose that  $A - \lambda B$  and  $C - \lambda D$  admit decompositions of form (3.19) with  $Y_i = X_i$ ,  $i = 1, 2$  and  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1$ ,  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, n$ . The uniqueness of the solution is evident. From Theorem 3.2 and Theorem 2.4, we see

$$\begin{aligned} \eta \|X\| &\leq \|X_1\|_2 \|X_1 S X_2\| \|X_2\|_2 \\ &\leq \|X_1\|_2^2 \|S\| \|X_2\|_2^2 \\ &\leq \frac{\|S\|}{c(A, B) c(C, D)}. \quad \blacksquare \end{aligned}$$

*Remark 3.2.* The foregoing results are all derived under the condition (3.10). Obviously, they remain valid when we interchange the positions of  $A - \lambda B$  and  $C - \lambda D$  in (3.10).

§4. BOUNDS IN MORE GENERAL SITUATIONS

Notations (3.1) and (3.2) are conserved throughout this section. And assume a weaker hypothesis than (3.10) as

$$\min_{\substack{(\mu, \nu) \in \lambda(A, B) \\ (\omega, \tau) \in \lambda(C, D)}} \rho((\mu, \nu), (\omega, \tau)) \geq \eta > 0. \tag{4.1}$$

We claim that (refer to Propositions 3.1, 3.2)

$$(3.10) \Rightarrow (4.1). \tag{4.2}$$

To see this, we assume, without loss of generality, that  $(\alpha_0, \beta_0) = (1, 0)$ . (Otherwise, we carry out a transformation just like we do in the proof of Lemma 3.1.) If (3.10) holds, then for any  $(\mu, \nu) \in \lambda(A, B)$  with  $|\mu|^2 + |\nu|^2 = 1$  and  $(\omega, \tau) \in \lambda(C, D)$  with  $|\omega|^2 + |\tau|^2 = 1$ ,

$$\begin{aligned} \xi &\geq \rho((1, 0), (\mu, \nu)) = |\nu| \Rightarrow |\mu| \geq \sqrt{1 - \xi^2}, \\ \xi + \delta &\leq \rho((1, 0), (\omega, \tau)) = |\tau| \Rightarrow |\omega| \leq \sqrt{1 - (\xi + \delta)^2}. \end{aligned}$$

These produce

$$\begin{aligned} \rho((\mu, \nu), (\omega, \tau)) &= |\mu\tau - \omega\nu| \\ &\geq |\mu\tau| - |\omega\nu| \\ &\geq (\xi + \delta)\sqrt{1 - \xi^2} - \xi\sqrt{1 - (\xi + \delta)^2}, \end{aligned}$$

by (3.3) this establishes (4.2). It is not difficult for one to give an counterexample to illustrate that (3.10) is not implied by (4.1).

Our following discussion also involves a hypothesis on the distribution of the spectra of  $A - \lambda B$  and of  $C - \lambda D$  on Riemann sphere.

**HYPOTHESIS I.** *There are  $(\gamma_0, \delta_0) \in \mathbf{G}_{1,2}$  and  $0 < \gamma \leq \Gamma \leq 1$  such that*

$$\begin{cases} \gamma \leq \rho((\gamma_0, \delta_0), (\mu, \nu)), \rho((\gamma_0, \delta_0), (\omega, \tau)) \leq \Gamma, \\ \text{for any } (\mu, \nu) \in \lambda(A, B), (\omega, \tau) \in \lambda(C, D). \end{cases} \tag{4.3}$$

Let  $(\alpha, \beta) \in \mathbf{G}_{1,2}$  be real, thus we may assume that  $\alpha, \beta \in \mathbb{R}$ . Therefore  $(i = \sqrt{-1})$

$$\rho((1, i), (\alpha, \beta)) = \rho((i, 1), (\alpha, \beta)) = \frac{1}{\sqrt{2}}.$$

So if  $A - \lambda B$  and  $C - \lambda D$  both have only real generalized eigenvalues, then (4.3) holds with  $(\gamma_0, \delta_0) = (1, i)$  or  $= (i, 1)$  and  $\gamma = \Gamma = \frac{1}{\sqrt{2}}$ . This is an important fact. Generally, it is very useful to choose a  $(\gamma_0, \delta_0)$  with  $1 \leq \frac{\Gamma}{\gamma}$  of moderate size. (This observation will impress us deeply with the following results.)

**LEMMA 4.1.** *Let  $(\alpha_i, \beta_i), (\gamma_j, \delta_j) \in \mathbf{G}_{1,2}$  satisfy  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1$ ,  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, n$ . And  $\Lambda_j$  and  $\Omega_j$  are defined by (3.19b) ( $j = 1, 2$ ). Assume also (4.1) and (4.3) hold for  $\Lambda_j - \lambda\Omega_j$ ,  $j = 1, 2$ . Then the equation  $\Lambda_1 X \Omega_2 - \Omega_1 X \Lambda_2 = S$*

has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and there is a positive constant  $c$  independent of  $m$  and  $n$  such that

$$\eta \|X\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \|S\|. \quad (4.4)$$

Especially, if  $\alpha_i, \beta_i, \gamma_j, \delta_j$  are all real, then with the same constant  $c$  as (4.4)

$$\eta \|X\| \leq c \|S\|. \quad (4.5)$$

PROOF: We noting our foregoing observation, (4.5) is implied by (4.4).

We rely on the main result of [1]. Let  $M \in \mathbf{C}^{m \times m}$  and  $N \in \mathbf{C}^{n \times n}$  be two normal matrices,  $\lambda(M)$  and  $\lambda(N)$  be their spectra respectively, if

$$\delta \triangleq \min_{\substack{\lambda \in \lambda(M) \\ \hat{\lambda} \in \lambda(N)}} |\lambda - \hat{\lambda}| > 0, \quad (4.6)$$

then the equation  $MX - XN = S$  has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and there is a positive constant  $c$  independent of  $m$  and  $n$  such that

$$\delta \|X\| \leq c \|S\|. \quad (4.7)$$

We do not know what the best possible value of the constant  $c$  for general unitarily invariant norms is, but it is proved by Bhatia *et al* (see [2]) that

$$\frac{\pi}{2} < c \leq \frac{\pi}{2} \int_0^\pi \frac{\sin t}{t} dt < 2.91. \quad (4.8)$$

Without loss of generality, we assume  $|\gamma_0|^2 + |\delta_0|^2 = 1$ . Set

$$\begin{pmatrix} \tilde{\Lambda}_j \\ \tilde{\Omega}_j \end{pmatrix} = \begin{pmatrix} \bar{\gamma}_0 I & \bar{\delta}_0 I \\ \delta_0 & -\gamma_0 \end{pmatrix} \begin{pmatrix} \Lambda_j \\ \Omega_j \end{pmatrix}, \quad j = 1, 2.$$

Hence it follows from (4.3) that  $\tilde{\Omega}_j$  are nonsingular, thus

$$\begin{aligned} \Lambda_1 X \Omega_2 - \Omega_1 X \Lambda_2 &= S \\ \iff \tilde{\Lambda}_1 X \tilde{\Omega}_2 - \tilde{\Omega}_1 X \tilde{\Lambda}_2 &= S \\ \iff \tilde{\Omega}_1^{-1} \tilde{\Lambda}_1 X - X \tilde{\Lambda}_2 \tilde{\Omega}_2^{-1} &= \tilde{\Omega}_1^{-1} S \tilde{\Omega}_2^{-1}. \end{aligned} \quad (4.9)$$

It is easy to verify that

$$\begin{aligned} \tilde{\Omega}_1^{-1} \tilde{\Lambda}_1 &= \text{diag} \left( \frac{\bar{\gamma}_0 \alpha_1 + \bar{\delta}_0 \beta_1}{\gamma_0 \alpha_1 - \delta_0 \beta_1}, \dots, \frac{\bar{\gamma}_0 \alpha_m + \bar{\delta}_0 \beta_m}{\gamma_0 \alpha_m - \delta_0 \beta_m} \right), \\ \tilde{\Lambda}_2 \tilde{\Omega}_2^{-1} &= \text{diag} \left( \frac{\bar{\gamma}_0 \gamma_1 + \bar{\delta}_0 \delta_1}{\gamma_0 \gamma_1 - \delta_0 \delta_1}, \dots, \frac{\bar{\gamma}_0 \gamma_n + \bar{\delta}_0 \delta_n}{\gamma_0 \gamma_n - \delta_0 \delta_n} \right) \end{aligned} \quad (4.10)$$

and (by (4.3))

$$\gamma \leq |\delta_0 \alpha_i - \gamma_0 \beta_i|, |\delta_0 \gamma_j - \gamma_0 \delta_j| \leq \Gamma,$$

$$\left| \frac{\bar{\gamma}_0 \alpha_i + \bar{\delta}_0 \beta_i}{\gamma_0 \alpha_i - \delta_0 \beta_i} - \frac{\bar{\gamma}_0 \gamma_j + \bar{\delta}_0 \delta_j}{\gamma_0 \gamma_j - \delta_0 \delta_j} \right| = \frac{|\alpha_i \delta_j - \beta_i \gamma_j|}{|\delta_0 \alpha_i - \gamma_0 \beta_i| \cdot |\delta_0 \gamma_j - \gamma_0 \delta_j|} \geq \frac{\eta}{\Gamma^2}.$$

So

$$\min_{\substack{\lambda \in \lambda(\tilde{\Omega}_1^{-1} \tilde{\Lambda}_1) \\ \hat{\lambda} \in \lambda(\tilde{\Lambda}_2 \tilde{\Omega}_2^{-1})}} |\lambda - \hat{\lambda}| \geq \frac{\eta}{\Gamma^2} > 0, \tag{4.11}$$

$$\left\| \tilde{\Omega}_1^{-1} S \tilde{\Omega}_2^{-1} \right\| \leq \frac{\|S\|}{\gamma^2}. \tag{4.12}$$

(4.9)–(4.12) and (4.7)–(4.8) yield

$$\frac{\eta}{\Gamma^2} \|X\| \leq c \frac{\|S\|}{\gamma^2},$$

this proves (4.4), and the constant  $c$  can be chosen so that (4.8) holds. The proof of Li[14, §4, Lemma 3] shows that the best possible value of  $c$  in (4.4) and (4.5) is not less than  $\frac{\pi}{2}$ . ■

**THEOREM 4.1.** *Let  $AXD - BXC = S$ , and let (4.1) holds, thus the equation has a unique solution  $X \in \mathbf{C}^{m \times n}$ . Assume also that Hypothesis I is true and  $c$  as described in Lemma 4.1.*

(1) *If  $A - \lambda B \in \mathbf{N}_R(m)$  and  $C - \lambda D \in \mathbf{N}_L(n)$  then*

$$\eta \|X\| \leq c \left(\frac{\Gamma}{\gamma}\right)^2 \left\| (ZZ^H)^{-\frac{1}{2}} S (W_d^H W_d)^{-\frac{1}{2}} \right\|. \tag{4.13}$$

(2) *If  $A - \lambda B \in \mathbf{N}_R(m)$  and  $C - \lambda D \in \mathbf{N}_R(n)$  then*

$$\eta \|X\| \leq c \left(\frac{\Gamma}{\gamma}\right)^2 \left\| (ZZ^H)^{-\frac{1}{2}} S \right\| \|(W^H W)^{-\frac{1}{2}}\|_2. \tag{4.14}$$

(3) *If  $A - \lambda B \in \mathbf{N}_L(m)$  and  $C - \lambda D \in \mathbf{N}_L(n)$  then*

$$\eta \|X\| \leq c \left(\frac{\Gamma}{\gamma}\right)^2 \|(Z_d^H Z_d)^{-\frac{1}{2}}\|_2 \left\| S (W_d^H W_d)^{-\frac{1}{2}} \right\|. \tag{4.15}$$

(4) *If  $A - \lambda B \in \mathbf{N}_L(m)$  and  $C - \lambda D \in \mathbf{N}_R(n)$  then*

$$\eta \|X\| \leq c \left(\frac{\Gamma}{\gamma}\right)^2 \|(Z_d^H Z_d)^{-\frac{1}{2}}\|_2 \|S\| \|(W W^H)^{-\frac{1}{2}}\|_2. \tag{4.16}$$

**PROOF:** To explain the idea of the proof, we take (4.14) as an example. By Theorem 2.2, we deduce from  $A - \lambda B \in \mathbf{N}_R(m)$  and  $C - \lambda D \in \mathbf{N}_R(n)$  that

$$(ZZ^H)^{-\frac{1}{2}}(A - \lambda B) \in \mathbf{N}(m), \quad (W W^H)^{-\frac{1}{2}}(C - \lambda D) \in \mathbf{N}(n),$$

therefore there exist  $U_1, V_1 \in \mathcal{U}_m, U_2, V_2 \in \mathcal{U}_n$  such that

$$\begin{cases} (ZZ^H)^{-\frac{1}{2}}A = U_1\Lambda_1V_1^H \\ (ZZ^H)^{-\frac{1}{2}}B = U_1\Omega_1V_1^H \end{cases} \quad \text{and} \quad \begin{cases} (WW^H)^{-\frac{1}{2}}C = U_2\Lambda_2V_2^H \\ (WW^H)^{-\frac{1}{2}}D = U_2\Omega_2V_2^H \end{cases},$$

where  $\Lambda_j, \Omega_j$  are of form (3.19b) with  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1$ . Hence from  $AXD - BXC = S$  follow

$$\begin{aligned} \Lambda_1V_1^HX(WW^H)^{\frac{1}{2}}U_2\Omega_2 - \Omega_1V_1^HX(WW^H)^{\frac{1}{2}}U_2\Lambda_2 &= U_1^H(ZZ^H)^{-\frac{1}{2}}SV_2 \\ \Rightarrow \eta \left\| V_1^HX(WW^H)^{\frac{1}{2}}U_2 \right\| &= \eta \left\| X(WW^H)^{\frac{1}{2}} \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \left\| U_1^H(ZZ^H)^{-\frac{1}{2}}SV_2 \right\| \\ &= c \left( \frac{\Gamma}{\gamma} \right)^2 \left\| (ZZ^H)^{-\frac{1}{2}}S \right\| \\ \Rightarrow \eta \|X\| &= \eta \left\| X(WW^H)^{\frac{1}{2}} \right\| \|(WW^H)^{-\frac{1}{2}}\|_2 \\ &\leq c \left( \frac{\Gamma}{\gamma} \right)^2 \left\| (ZZ^H)^{-\frac{1}{2}}S \right\| \|(W^HW)^{-\frac{1}{2}}\|_2. \end{aligned}$$

This establishes (4.14).  $\blacksquare$

Generally, for two diagonalizable pencils, we have

**THEOREM 4.2.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $C - \lambda D \in \mathbf{D}_g(n)$  admit decompositions (3.19) with  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1, i = 1, 2, \dots, m, j = 1, 2, \dots, n$  and let (4.1) holds, thus the equation  $AXD - BXC = S$  has a unique solution  $X \in \mathbf{C}^{m \times n}$ . If also Hypothesis I is true, then the solution satisfies*

$$\eta \|X\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \|X_1\|_2 \|Y_1^H S X_2\| \|Y_2\|_2. \quad (4.17)$$

*Epecially, if  $A - \lambda B$  and  $C - \lambda D$  both have only real generalized eigenvalues, then  $\left(\frac{\Gamma}{\gamma}\right)^2$  can be removed, i.e.,*

$$\eta \|X\| \leq c \|X_1\|_2 \|Y_1^H S X_2\| \|Y_2\|_2. \quad (4.18)$$

The constant  $c$  in (4.17) and (4.18) is as described in Lemma 4.1.

We have noted that for Hermitian pencils or definite pencils  $(\gamma_0, \delta_0)$  can be chosen so that  $\Gamma = \gamma = \frac{1}{\sqrt{2}}$ . Therefore, from Theorems 4.1 and 4.2 follow

**THEOREM 4.3.** *Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $C - \lambda D \in \mathbf{D}(n)$ , and suppose (4.1) holds, then the equation  $AXD - BXC = S$  has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and*

$$\eta \|X\| \leq c \frac{\|S\|}{c(A, B) c(C, D)}, \quad (4.19)$$

where the constant  $c$  is as described in Lemma 4.1.

**THEOREM 4.4.** *Let  $AXD - BXC = S$ , and (4.1) holds. (Thus the equation has a unique solution.) The (1)–(4) of Theorem 4.1 with  $\frac{\Gamma}{\gamma} = 1$  and all the  $\mathbf{N}$  replaced by  $\mathbf{H}$  remain valid.*

## §5. THE FROBENIUS NORM

One of unitarily invariant norms to which the theorems of §§3 and 4 apply is Frobenius norm. We will show in this section that it is a special case, not only (as usual) in its simplicity, but in the strength of the estimates and weakness of hypothesis. (Note that Hypotheses I in §4 is, now, not required.)

**THEOREM 5.1.** *Let  $AXD - BXC = S$ , and (4.1) holds. (Thus the equation has a unique solution.) The (1)–(4) of Theorem 4.1 with  $c = 1$  and  $\frac{\Gamma}{\gamma}$  removed from the inequalities and with  $\|\cdot\| = \|\cdot\|_F$  remain valid.*

Similar results hold generally for two diagonalizable pencils, of course for definite pencils (cf. Theorems 4.2 and 4.3). To prove this theorem, it is enough to illustrate

**LEMMA 5.1.** *Let  $(\alpha_i, \beta_i), (\gamma_j, \delta_j) \in \mathbf{G}_{1,2}$  satisfy  $|\alpha_i|^2 + |\beta_i|^2 = |\gamma_j|^2 + |\delta_j|^2 = 1$ ,  $i = 1, 2, \dots, m$ ,  $j = 1, 2, \dots, n$ . And  $\Lambda_j$  and  $\Omega_j$  are defined by (3.19b) ( $j = 1, 2$ ). If (4.1) holds for  $\Lambda_j - \lambda\Omega_j$ ,  $j = 1, 2$ , then the equation  $\Lambda_1 X \Omega_2 - \Omega_1 X \Lambda_2 = S$  has a unique solution  $X \in \mathbf{C}^{m \times n}$ , and*

$$\eta \|X\|_F \leq \|S\|_F. \quad (5.1)$$

**PROOF:** Denote by  $X = (q_{ij})$ , then

$$\begin{aligned} \|S\|_F^2 &= \|\Lambda_1 X \Omega_2 - \Omega_1 X \Lambda_2\|_F^2 \\ &= \sum_{i,j} |\alpha_i \delta_j - \beta_i \gamma_j|^2 |q_{ij}|^2 \\ &\geq \eta^2 \sum_{i,j} |q_{ij}|^2 \\ &= \eta^2 \|X\|_F^2. \end{aligned}$$

This establishes (5.1). ■

Theorem 5.1 says the constants obtained in theorems of §4 for arbitrary norms can be improved for special norms.

## REFERENCES

- [1] Bhatia, R., Davis, Ch. and McIntosh, A., Perturbation of spectral subspaces and the solution of linear equations, *Lin. Alg. Appl.*, **52/53**(1983), 45-67.
- [2] Bhatia, R. and Kittaneh, F., On some perturbation inequities for operators, *Lin. Alg. Appl.*, **106**(1988), 271-279.
- [3] Bhatia, R., Davis, Ch. and Koosis, P., An extremal problem in Fourier analysis with applications to operator theory, *J. Funct. Anal.*, **82**(1989), 138-150.
- [4] Bondy, J. A. and Murty, U. S. R., *Graph Theory with Applications*, MacMillan Press, London, 1976.
- [5] Chu, K. -W. E., Exclusion theorems and the perturbation analysis of the generalized eigenvalue problem, *SIAM J. Numer. Anal.*, **24**(1987), 1114-1125.
- [6] Crawford, C. R., A stable generalized eigenvalue problem, *SIAM J. Numer. Anal.*, **8**(1976), 854-860.
- [7] Davis, Ch. and Kahan, W. M., The rotation of eigenvectors by a perturbation. III, *SIAM J. Numer. Anal.*, **7**(1970), 1-46.
- [8] Davis, Ch. and Kahan, W. M., Some new bounds on perturbation of subspaces, *Bull. Amer. Math. Soc.*, **75**(1969), 863-868.
- [9] Elsner, L. and Sun, Ji-guang, Perturbation theorems for the generalized eigenvalue problem, *Lin. Alg. Appl.*, **48**(1982), 341-357.
- [10] Gohberg, I. C. and Krein, M. G., *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, R. I., 1969.
- [11] Hoffman, A. J. and Wielandt, H. W., The variation of the spectrum of a normal matrix, *Duke Math. J.*, **20**(1953), 37-39.
- [12] Lancaster, P., Explicit solutions of linear matrix equations, *SIAM Rev.*, **12**(1970), 544-566.
- [13] Li, Ren-cang, *M.S. dissertation*, Computing Center, Academia Sinica, 1987. (Chinese)
- [14] Li, Ren-cang, On perturbations of matrix pencils with real spectra, to appear.
- [15] Li, Ren-cang, On perturbations of eigenspaces of matrix pencils with real spectra, to appear.
- [16] Li, Ren-cang, Bounds on perturbations of generalized singular values and associated subspaces, to appear.
- [17] Mirsky, L., Symmetric gauge functions and unitarily invariant norms, *Quart. J. Math. Oxford*, **11**(1960), 50-59.
- [18] Stewart, G. W., On the sensitivity of the eigenvalue problem  $Ax = \lambda Bx$ , *SIAM J. Numer. Anal.*, **9**(1972), 669-86.
- [19] Stewart, G. W., Perturbation bounds for the definite generalized eigenvalue problem, *Lin. Alg. Appl.*, **23**(1979), 69-83.
- [20] Sun, Ji-guang, The perturbation bounds of generalized eigenvalues of a class of matrix-pairs, *Math. Numer. Sinica*, **4**(1982), 23-29.
- [21] Sun, Ji-guang, A note on Stewart's theorem for definite matrix pairs, *Lin. Alg. Appl.*, **48**(1982), 331-339.
- [22] Sun, Ji-guang, The perturbation bounds for eigenspaces of a definite matrix-pair, *Numer. Math.*, **41**(1983), 321-343. (Chinese)
- [23] Sun, Ji-guang, *Matrix Perturbation Analysis*, Science Press, Beijing, 1987. (Chinese)

**SOLUTION OF LINEAR MATRIX EQUATION  $AXD - BXC = S$   
AND PERTURBATION OF EIGENSPACES OF A MATRIX PENCIL. II\***

REN-CANG LI

INSTITUTE OF APPLIED PHYSICS AND COMPUTATIONAL MATHEMATICS  
P. O. Box 8009  
BEIJING 100088, P. R. CHINA

ABSTRACT

We illustrate how to apply estimates of norms of the solution  $X$  to the linear matrix equation  $AXD - BXC = S$  to estimate the eigenspace and eigenvalue variations of matrix pencils of certain classes. As examples, we will derive here several perturbation bounds

§6. PERTURBATION OF EIGENSPACES

For the sake of convenience, in this section we use  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  for two regular matrix pencils in attempting to search bounds upon the unitarily invariant norms of the sine of the angle between eigenspaces belong to certain spectra of two pencils. The key to our method is how to relate the sine of the angle to solution of a linear matrix equation of form  $AXD - BXC = S$ . In this way, we rely on results of Li[15] to constructe perturbation equation. It does not appeal us to write out all results as possible as we could. Nevertheless, we will try to explain our main idea in detail, so as to readers may establish, without difficulty, other bounds necessary to their research or applications (refer to [15]).

We take some results of §4 as examples.

Suppose that  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  admit decompositions

$$\begin{cases} Y^H AX = \Lambda \\ Y^H BX = \Omega \end{cases} \quad \text{and} \quad \begin{cases} \tilde{Y}^H \tilde{A} \tilde{X} = \tilde{\Lambda} \\ \tilde{Y}^H \tilde{B} \tilde{X} = \tilde{\Omega} \end{cases}, \tag{6.1a}$$

where

$$\begin{cases} X = (X_1, X_2), \tilde{X} = (\tilde{X}_1, \tilde{X}_2), Y = (Y_1, Y_2), \tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2), \\ X^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix}, \tilde{X}^{-1} = \begin{pmatrix} \tilde{W}_1^H \\ \tilde{W}_2^H \end{pmatrix}, Y^{-1} = \begin{pmatrix} S_1^H \\ S_2^H \end{pmatrix}, \tilde{Y}^{-1} = \begin{pmatrix} \tilde{S}_1^H \\ \tilde{S}_2^H \end{pmatrix}, \\ X_1, \tilde{X}_1, Y_1, \tilde{Y}_1, W_1, \tilde{W}_1, S_1, \tilde{S}_1 \in \mathbf{C}^{n \times \ell}, \end{cases} \tag{6.1b}$$

---

\*This is the second part of the article. We adhere to the notation introduced in Li[24], the first part of it.

$$\begin{cases} \Lambda = \text{diag}(\alpha_1, \dots, \alpha_n) \triangleq \text{diag}(\Lambda_1, \Lambda_2), \\ \Omega = \text{diag}(\beta_1, \dots, \beta_n) \triangleq \text{diag}(\Omega_1, \Omega_2), \\ \Lambda_1, \Omega_1 \in \mathbf{C}^{\ell \times \ell}, \Lambda_2, \Omega_2 \in \mathbf{C}^{(n-\ell) \times (n-\ell)}, \end{cases} \quad (6.2a)$$

$$\begin{cases} \tilde{\Lambda} = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n) \triangleq \text{diag}(\tilde{\Lambda}_1, \tilde{\Lambda}_2), \\ \tilde{\Omega} = \text{diag}(\tilde{\beta}_1, \dots, \tilde{\beta}_n) \triangleq \text{diag}(\tilde{\Omega}_1, \tilde{\Omega}_2), \\ \tilde{\Lambda}_1, \tilde{\Omega}_1 \in \mathbf{C}^{\ell \times \ell}, \tilde{\Lambda}_2, \tilde{\Omega}_2 \in \mathbf{C}^{(n-\ell) \times (n-\ell)}. \end{cases} \quad (6.2b)$$

Set

$$\mathcal{X}_1 = \mathcal{R}(X_1), \quad \tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{X}_1), \quad (6.3)$$

$$Z = (A, B), \quad \tilde{Z} = (\tilde{A}, \tilde{B}). \quad (6.4)$$

LEMMA 6.1 (Li[15]). *Suppose that  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  admit decompositions (6.1)–(6.2), then*

$$\tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda = -(\tilde{\Lambda}, \tilde{\Omega}) \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}, \quad (6.5a)$$

$$\tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1 = -(\tilde{\Lambda}_2, \tilde{\Omega}_2) \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ -\Lambda_1 \end{pmatrix} \quad (6.5b)$$

and

$$\tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda = \tilde{Y}^H (\tilde{Z} - Z) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}, \quad (6.6a)$$

$$\tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1 = \tilde{Y}_2^H (\tilde{Z} - Z) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ -\Lambda_1 \end{pmatrix}. \quad (6.6b)$$

Identities (6.5) and (6.6) are implied by the proofs of Theorem 1 and Theorem 1a in Li[15]. From the proofs, we know that (6.5) and (6.6) are valid for  $\alpha_i, \beta_i, \gamma_j, \delta_j \in \mathbf{C}$ , though the original proofs of (6.5) and (6.6) in Li[15] are done under the assumption  $\alpha_i, \beta_i, \gamma_j, \delta_j \in \mathbb{R}$ . In order to make the results of §4 applicable, we assume

HYPOTHESIS II. *There are  $(\gamma_0, \delta_0) \in \mathbf{G}_{1,2}$ ,  $0 < \gamma \leq \Gamma \leq 1$ , so that*

$$\gamma \leq \rho((\gamma_0, \delta_0), (\alpha_i, \beta_i)), \rho((\gamma_0, \delta_0), (\tilde{\alpha}_j, \tilde{\beta}_j)) \leq \Gamma, \quad (6.7)$$

for  $i = 1, 2, \dots, \ell, j = \ell + 1, \dots, n$ .

THEOREM 6.1. *Suppose that  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  admit decompositions (6.1)–(6.2), and suppose that Hypothesis II is satisfied. Let*

$$\eta = \min_{\substack{1 \leq i \leq \ell \\ \ell + 1 \leq j \leq n}} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)). \quad (6.8)$$

If  $\eta > 0$ , then for any unitarily invariant norm  $\|\cdot\|$

$$\left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{\|X_1^+\|_2 \|\tilde{W}_2^+\|_2}{\eta} \left\| \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\|. \quad (6.9)$$

If, in addition,  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$ ,  $i, j = 1, 2, \dots, n$ , then

$$\left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{\|X_1^+\|_2 \|\tilde{W}_2^+\|_2}{\eta} \left\| \tilde{Y}_2^H (\tilde{Z} - Z) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\|. \quad (6.10)$$

Here

$$\Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \triangleq \arccos((X_1^H X_1)^{-1/2} X_1^H \tilde{X}_1 (\tilde{X}_1^H \tilde{X}_1)^{-1} \tilde{X}_1^H X_1 (X_1^H X_1)^{-1/2})^{1/2} \geq 0.$$

It can be shown that (see, e.g., Li[15, §2, Lemma 2])

$$\left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\| = \left\| (\tilde{W}_2^H \tilde{W}_2)^{-\frac{1}{2}} \tilde{W}_2^H X_1 (X_1^H X_1)^{-\frac{1}{2}} \right\|.$$

In other words, the set of singular values of the matrix  $\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)$  is the same as that of the matrix  $(\tilde{W}_2^H \tilde{W}_2)^{-\frac{1}{2}} \tilde{W}_2^H X_1 (X_1^H X_1)^{-\frac{1}{2}}$ .

PROOF OF THEOREM 6.1: By the results of §4, if  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$  then there exists a positive absolute constant  $c$  so that

$$\begin{aligned} & c \left( \frac{\Gamma}{\gamma} \right)^2 \left\| \tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1 \right\| \\ & \geq \eta \left\| \tilde{W}_2^H X_1 \right\| \\ & \geq \eta \left\| (\tilde{W}_2^H \tilde{W}_2)^{-\frac{1}{2}} \right\|_2^{-1} \left\| (\tilde{W}_2^H \tilde{W}_2)^{-\frac{1}{2}} \tilde{W}_2^H X_1 (X_1^H X_1)^{-\frac{1}{2}} \right\| \left\| (X_1^H X_1)^{-\frac{1}{2}} \right\|_2^{-1} \\ & \geq \eta \|\tilde{W}_2^+\|_2^{-1} \|X_1^+\|_2^{-1} \left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\|. \end{aligned} \quad (6.11)$$

Now, we notice (6.5) doesn't contain  $Y$  and  $\tilde{Y}$ , thus by adjusting  $Y, \tilde{Y}$  suitably, we have  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$ . Therefore, (6.5) and (6.11) produce (6.9) while (6.10) is the consequence of (6.6) and (6.11). ■

Consider normal pencil case. Suppose  $A - \lambda B \in \mathbf{N}_R(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{N}_R(n)$  admit the decompositions

$$\begin{cases} Y^H A V = \Lambda \\ Y^H B V = \Omega \end{cases} \quad \text{and} \quad \begin{cases} \tilde{Y}^H \tilde{A} \tilde{V} = \tilde{\Lambda} \\ \tilde{Y}^H \tilde{B} \tilde{V} = \tilde{\Omega} \end{cases}, \quad (6.12a)$$

where  $V, \tilde{V} \in \mathcal{U}_n$ ,  $\Lambda, \Omega, \dots, \tilde{\alpha}_1, \tilde{\beta}_1, \dots$  are of form (6.2),

$$\begin{cases} V = (V_1, V_2), \tilde{V} = (\tilde{V}_1, \tilde{V}_2), \\ V_1, \tilde{V}_1 \in \mathbf{C}^{n \times \ell}. \end{cases} \quad (6.12b)$$

Now, (6.5) becomes

$$\tilde{\Lambda} \tilde{V}^H V \Omega - \tilde{\Omega} \tilde{V}^H V \Lambda = -(\tilde{\Lambda}, \tilde{\Omega}) \begin{pmatrix} \tilde{V}^H & \\ & \tilde{V}_H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V & \\ & V \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}, \quad (6.13)$$

$$\begin{aligned} & \tilde{\Lambda}_2 \tilde{V}_2^H V_1 \Omega_1 - \tilde{\Omega}_2 \tilde{V}_2^H V_1 \Lambda_1 \\ & = -(\tilde{\Lambda}_2, \tilde{\Omega}_2) \begin{pmatrix} \tilde{V}_2^H & \\ & \tilde{V}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V_1 & \\ & V_1 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ -\Lambda_1 \end{pmatrix} \triangleq E. \end{aligned} \quad (6.14)$$

By treating  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  symmetrically, from (6.13) follows

$$\begin{aligned}
& \Lambda V^H \tilde{V} \tilde{\Omega} - \Omega V^H \tilde{V} \tilde{\Lambda} = -(\Lambda, \Omega) \begin{pmatrix} V^H & \\ & V^H \end{pmatrix} (P_{\tilde{Z}^H} - P_{Z^H}) \begin{pmatrix} \tilde{V} & \\ & \tilde{V} \end{pmatrix} \begin{pmatrix} \tilde{\Omega} \\ -\tilde{\Lambda} \end{pmatrix} \\
\Rightarrow & \tilde{\Lambda}^H \tilde{V}^H V \Omega^H - \tilde{\Omega}^H \tilde{V}^H V \Lambda^H = -(\tilde{\Omega}^H, -\tilde{\Lambda}^H) \begin{pmatrix} \tilde{V}^H & \\ & \tilde{V}^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V & \\ & V \end{pmatrix} \begin{pmatrix} \Lambda^H \\ \Omega^H \end{pmatrix} \\
\Rightarrow & \tilde{\Lambda}_2^H \tilde{V}_2^H V_1 \Omega_1^H - \tilde{\Omega}_2^H \tilde{V}_2^H V_1 \Lambda_1^H \\
& = -(\tilde{\Omega}_2^H, -\tilde{\Lambda}_2^H) \begin{pmatrix} \tilde{V}_2^H & \\ & \tilde{V}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V_1 & \\ & V_1 \end{pmatrix} \begin{pmatrix} \Lambda_1^H \\ \Omega_1^H \end{pmatrix} \triangleq F. \quad (6.15)
\end{aligned}$$

LEMMA 6.2. Let  $M_i, N_i, i = 1, 2$  be four matrices with suitable dimensions. If for any unitarily invariant norm  $\| \cdot \|$

$$\|M_i\| \leq \|N_i\|, \quad i = 1, 2$$

hold, then so does the following inequality

$$\left\| \begin{pmatrix} M_1 & \\ & M_2 \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} N_1 & \\ & N_2 \end{pmatrix} \right\|. \quad (6.16)$$

PROOF: By a theorem of Ky Fan (see also [17]) that  $\|M\| \leq \|N\|$  holds for all unitarily invariant norm if and only if the sum of the  $k$  largest singular values of  $M$  is less than that of  $N$  ( $k = 1, 2, \dots$ ), we see easily from  $\|M_i\| \leq \|N_i\|$  ( $i = 1, 2$ ) that the sum of the  $k$  largest singular values of  $\text{diag}(M_1, M_2)$  is less than that of  $\text{diag}(N_1, N_2)$  ( $k = 1, 2, \dots$ ). This proves (6.16). ■

Armed with the identities (6.13) and (6.15) and Lemma 6.2, we now prove

THEOREM 6.2. Besides the hypotheses of Theorem 6.1, we assume more strongly that  $A - \lambda B \in \mathbf{N}_R(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{N}_R(n)$  and that they admit the decompositions (6.12).  $\eta > 0$  is defined by (6.8). Then for any unitarily invariant norm  $\| \cdot \|$

$$\begin{aligned}
& \left\| \begin{pmatrix} \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) & \\ & \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \end{pmatrix} \right\| \\
& \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{1}{\eta} \left\| \begin{pmatrix} \tilde{V}_2^H & \\ & \tilde{V}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V_1 & \\ & V_1 \end{pmatrix} \right\|, \quad (6.17)
\end{aligned}$$

where  $\mathcal{X}_1 = \mathcal{R}(V_1)$ ,  $\tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{V}_1)$ .

PROOF: Without loss of generality, we may assume that  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$ . It follows from (6.13) and (6.15) and the results of §4 and Lemma 6.2 that

$$\begin{aligned}
& \left\| \tilde{V}_2^H V_1 \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{1}{\eta} \|E\| \\
& \left\| \tilde{V}_2^H V_1 \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{1}{\eta} \|F\| \\
\Rightarrow & \left\| \begin{pmatrix} \tilde{V}_2^H V_1 & \\ & \tilde{V}_2^H V_1 \end{pmatrix} \right\| \leq c \left( \frac{\Gamma}{\gamma} \right)^2 \frac{1}{\eta} \left\| \begin{pmatrix} E & \\ & F \end{pmatrix} \right\|. \quad (6.18)
\end{aligned}$$

Since the nonzero singular values of  $\tilde{V}_2^H V_1$  are the same as those of  $\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)$ , therefore

$$\left\| \begin{pmatrix} \tilde{V}_2^H V_1 & \\ & \tilde{V}_2^H V_1 \end{pmatrix} \right\| \equiv \left\| \begin{pmatrix} \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) & \\ & \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \end{pmatrix} \right\|. \quad (6.19)$$

On the other hand, we have

$$\begin{pmatrix} E & * \\ * & F \end{pmatrix} = - \begin{pmatrix} \tilde{\Lambda}_2 & \tilde{\Omega}_2 \\ \tilde{\Omega}_2^H & -\tilde{\Lambda}_2^H \end{pmatrix} \begin{pmatrix} \tilde{V}_2^H & \\ & \tilde{V}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V_1 & \\ & V_1 \end{pmatrix} \begin{pmatrix} \Omega_1 & \Lambda_1^H \\ -\Lambda_1 & \Omega_1^H \end{pmatrix},$$

hence it follows from [10, Theorem 5.1, Chapter II] and

$$\begin{pmatrix} \tilde{\Lambda}_2 & \tilde{\Omega}_2 \\ \tilde{\Omega}_2^H & -\tilde{\Lambda}_2^H \end{pmatrix} \in \mathcal{U}_{2(n-\ell)}, \quad \begin{pmatrix} \Omega_1 & \Lambda_1^H \\ -\Lambda_1 & \Omega_1^H \end{pmatrix} \in \mathcal{U}_{2\ell}.$$

that

$$\left\| \begin{pmatrix} E & \\ & F \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} \tilde{V}_2^H & \\ & \tilde{V}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} V_1 & \\ & V_1 \end{pmatrix} \right\|. \quad (6.20)$$

(6.17) is the consequence of (6.18)–(6.20).  $\blacksquare$

*Remark 6.1.* Let  $\mathcal{X}_2 = \mathcal{R}(V_2)$ ,  $\tilde{\mathcal{X}}_2 = \mathcal{R}(\tilde{V}_2)$ . It is easy to see that

$$\left\| \sin \Theta(\mathcal{X}_2, \tilde{\mathcal{X}}_2) \right\| = \left\| V_1^H \tilde{V}_2 \right\| = \left\| \tilde{V}_2^H V_1 \right\|,$$

i.e., the nonzero singular values of  $\sin \Theta(\mathcal{X}_2, \tilde{\mathcal{X}}_2)$  are the same as those of  $\tilde{V}_2^H V_1$ . Thus by (6.18) and (6.19) the right hand side of (6.17) can be replaced by

$$\left\| \begin{pmatrix} \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) & \\ & \sin \Theta(\mathcal{X}_2, \tilde{\mathcal{X}}_2) \end{pmatrix} \right\|.$$

The case when the considered pencils have only real generalized eigenvalues is rather special as we have mentioned in §4. For the very case, we remark that  $\left(\frac{\Gamma}{\gamma}\right)^2$  can be removed from the inequalities (6.9) and (6.10). To illustrate the idea, we consider the definite pencil case. Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}(n)$  admit the decompositions

$$\begin{cases} X^H A X = \Lambda \\ X^H B X = \Omega \end{cases} \quad \text{and} \quad \begin{cases} \tilde{X}^H \tilde{A} \tilde{X} = \tilde{\Lambda} \\ \tilde{X}^H \tilde{B} \tilde{X} = \tilde{\Omega} \end{cases}, \quad (6.21a)$$

where  $\Lambda, \Omega, \dots, \tilde{\alpha}_1, \tilde{\beta}_1, \dots$  are of form (6.2),

$$\begin{cases} X = (X_1, X_2), \tilde{X} = (\tilde{X}_1, \tilde{X}_2), \\ X^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix}, \tilde{X}^{-1} = \begin{pmatrix} \tilde{W}_1^H \\ \tilde{W}_2^H \end{pmatrix} \\ X_1, \tilde{X}_1, W_1, \tilde{W}_1 \in \mathbf{C}^{n \times \ell}. \end{cases} \quad (6.21b)$$

**THEOREM 6.3.** Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}(n)$  admit the decompositions (6.21), and  $\eta > 0$  is defined by (6.8). Then for any unitarily invariant norm  $\|\cdot\|$

$$\left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\| \leq c \cdot \frac{\|Z\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \cdot \frac{1}{\eta} \cdot \left\| \tilde{X}_{20}^H (\tilde{Z} - Z) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\|, \quad (6.22)$$

where  $X_{10} = X_1(X_1^H X_1)^{-\frac{1}{2}}$ ,  $\tilde{X}_{20} = \tilde{X}_2(\tilde{X}_2^H \tilde{X}_2)^{-\frac{1}{2}}$ .

PROOF: Without loss of generality, we assume that  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$  in (6.21). Since Hypothesis II holds for  $(\gamma_0, \delta_0) = (1, \sqrt{-1})$  and  $\gamma = \Gamma = \frac{1}{\sqrt{2}}$ , therefore it follows from Theorem 6.1 that

$$\begin{aligned} \left\| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \right\| &\leq c \frac{\|X_1^+\|_2 \|\tilde{W}_2^+\|_2}{\eta} \left\| \tilde{X}_2^H (\tilde{Z} - Z) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\| \\ &\leq c \frac{\|X^{-1}\|_2 \|\tilde{X}\|_2}{\eta} \|(\tilde{X}_2^H \tilde{X}_2)^{\frac{1}{2}}\|_2 \left\| \tilde{X}_{20}^H (\tilde{Z} - Z) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\| \|(X_1^H X_1)^{\frac{1}{2}}\|_2 \\ &\leq c \frac{\|X^{-1}\|_2 \|X\|_2 \|\tilde{X}\|_2^2}{\eta} \left\| \tilde{X}_{20}^H (\tilde{Z} - Z) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\|. \end{aligned}$$

This together with Theorem 2.4 establish (6.22).  $\blacksquare$

If we ignore the weaker assumption on the separation between spectra, the appearance of the constant  $c$  in Theorem 6.3 is the only difference in form between (6.22) and Theorem 9b of [15, §5].

*Remark 6.2.* The applications of results in §6 will lead to the  $\sin \theta$  theorems obtained by Li[15]. Perturbation bounds for right eigenspaces of  $A - \lambda B$  can be obtained by applying the above results to the pencil  $A^H - \lambda B^H$ .

#### §7. APPLICATION TO GENERALIZED EIGENVALUE ESTIMATION

As another application, we prove in this section

**THEOREM 7.1.** *Suppose that  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  admit decompositions (6.1)–(6.2), and there exist  $(\gamma_0, \delta_0) \in \mathbf{G}_{1,2}$ ,  $0 < \gamma \leq \Gamma \leq 1$  such that*

$$\gamma \leq \rho((\gamma_0, \delta_0), (\alpha_i, \beta_i)), \rho((\gamma_0, \delta_0), (\tilde{\alpha}_j, \tilde{\beta}_j)) \leq \Gamma, \quad (7.1)$$

for  $i, j = 1, 2, \dots, n$ , then there exist a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  and an absolute positive number  $d$  independent of  $n$  and the pencils so that

$$\max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq d \left( \frac{\Gamma}{\gamma} \right)^2 \kappa(X) \kappa(\tilde{X}) \|P_{Z^H} - P_{\tilde{Z}^H}\|_2 \quad (7.2)$$

PROOF: By a well-known combinatorial theorem—Marriage Theorem ([4, p.73]), it is enough to prove the following assertion (cf. [1, §5]): if  $\mathbf{K}$  be a set consisting of  $k$  of the  $(\alpha_i, \beta_i)$ , then the set

$$\mathbf{K}_1 \triangleq \left\{ (\tilde{\alpha}_j, \tilde{\beta}_j) \left| \min_{(\alpha_i, \beta_i) \in \mathbf{K}} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)) \leq \epsilon \right. \right\}$$

containing at least  $k$  of the  $(\tilde{\alpha}_j, \tilde{\beta}_j)$ . Here  $\epsilon \triangleq d \left( \frac{\Gamma}{\gamma} \right)^2 \kappa(X) \kappa(\tilde{X}) \|P_{Z^H} - P_{\tilde{Z}^H}\|_2$ . Assume to the contrary that this assertion does not hold, then without loss of generality, we would have  $\mathbf{K} = \{(\alpha_i, \beta_i), i = 1, 2, \dots, k\}$  and  $\mathbf{K}_1 = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = k, \dots, n\}$  such that

$$\min_{\substack{1 \leq i \leq k \\ k \leq j \leq n}} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)) \geq \eta > \epsilon \quad (7.3)$$

for some  $\eta$ . Partition  $X, \tilde{X}, \Lambda, \Omega, \tilde{\Lambda}, \tilde{\Omega}$  as in the forms of (6.1) and (6.2) but with

$$X_1, W_1 \in \mathbb{C}^{n \times k}, \quad \tilde{X}_1, \tilde{W}_1 \in \mathbb{C}^{n \times (k-1)}, \tag{7.4a}$$

$$\Lambda_1, \Omega_1 \in \mathbb{C}^{k \times k}, \quad \tilde{\Lambda}_1, \tilde{\Omega}_1 \in \mathbb{C}^{(k-1) \times (k-1)}. \tag{7.4b}$$

At present time, (6.5a) is, of course, true, and yields

$$\tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1 = -(\tilde{\Lambda}_2, \tilde{\Omega}_2) \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ -\Lambda_1 \end{pmatrix}.$$

(We noting the current partitions, this identity is different from (6.5b), although formally, the two is the same.) Now from (7.1), (7.3) and the results of §4 follows

$$\begin{aligned} \eta \|\tilde{W}_2^H X_1\|_2 &\leq c \left(\frac{\Gamma}{\gamma}\right)^2 \left\| (\tilde{\Lambda}_2, \tilde{\Omega}_2) \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \begin{pmatrix} \Omega_1 \\ -\Lambda_1 \end{pmatrix} \right\|_2 \\ &\leq c \left(\frac{\Gamma}{\gamma}\right)^2 \|\tilde{W}_2\|_2 \|P_{Z^H} - P_{\tilde{Z}^H}\|_2 \|X_1\|_2 \\ &\leq c \left(\frac{\Gamma}{\gamma}\right)^2 \|\tilde{X}^{-1}\|_2 \|X\|_2 \|P_{Z^H} - P_{\tilde{Z}^H}\|_2. \end{aligned} \tag{7.5}$$

On the other hand (cf. (6.11))

$$\|\tilde{W}_2^H X_1\|_2 \geq \|\tilde{W}_2^+\|_2^{-1} \|X_1^+\|_2^{-1} \|\tilde{W}_{20}^H X_{10}\|_2, \tag{7.6}$$

where  $\tilde{W}_{20} = \tilde{W}_2 (\tilde{W}_2^H \tilde{W}_2)^{-\frac{1}{2}}$ ,  $X_{10} = X_1 (X_1^H X_1)^{-\frac{1}{2}}$ . (7.4) produces

$$X_{10} \in \mathbb{C}^{n \times k}, \quad \tilde{W}_{20} \in \mathbb{C}^{n \times (n-k+1)}$$

which together with  $k + (n - k + 1) = n + 1 > n$  yield  $\mathcal{R}(X_{10}) \cap \mathcal{R}(\tilde{W}_{20}) \neq \emptyset$ . Therefore there exist  $g \in \mathbb{C}^k, h \in \mathbb{C}^{n-k+1}$  with  $\|g\|_2 = \|h\|_2 = 1$  so that

$$X_{10}g = \tilde{W}_{20}h \triangleq x.$$

Obviously,  $\|x\|_2 = 1$ , hence

$$\|\tilde{W}_{20}^H X_{10}\|_2 \geq h^H \tilde{W}_{20}^H X_{10}g = x^H x = 1.$$

So, (7.6) yields

$$\|\tilde{W}_2^H X_1\|_2 \geq \|\tilde{W}_2^+\|_2^{-1} \|X_1^+\|_2^{-1} \geq \|\tilde{X}\|_2^{-1} \|X^{-1}\|_2^{-1}.$$

Combining this with (7.5), we obtain

$$\eta \leq c \left(\frac{\Gamma}{\gamma}\right)^2 \kappa(X) \kappa(\tilde{X}) \|P_{Z^H} - P_{\tilde{Z}^H}\|_2.$$

Our taking  $d = c$  produces  $\eta \leq \epsilon$ , a contradiction to (7.3). This completes the proof. ■

It can be seen that if  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  both have only real generalized eigenvalues, then  $\left(\frac{\Gamma}{\gamma}\right)^2 = 1$ . Theorem 1 of Li[14] tells us that for this very case the constant  $d$  can also be removed. It remains to study whether the bothering factor  $\left(\frac{\Gamma}{\gamma}\right)^2$  entering our bounds might be crossed out or not.

REFERENCES\*

[24] Li, Ren-cang, Solution of linear matrix equation  $AXD - BXC = S$  and perturbation of eigenspaces of a matrix pencil. I.

\*References [1]–[23] are as listed in Li[24], the first part of the paper.