

## ON PERTURBATIONS OF MATRIX PENCILS WITH REAL SPECTRA

REN-CANG LI

**ABSTRACT.** Perturbation bounds for the generalized eigenvalue problem of a diagonalizable matrix pencil  $A - \lambda B$  with real spectrum are developed. It is shown how the chordal distances between the generalized eigenvalues and the angular distances between the generalized eigenspaces can be bounded in terms of the angular distances between the matrices. The applications of these bounds to the spectral variations of definite pencils are conducted in such a way that extra attention is paid to their peculiarities so as to derive more sophisticated perturbation bounds. Our results for generalized eigenvalues are counterparts of some celebrated theorems for the spectral variations of Hermitian matrices such as the Weyl-Lidskii theorem and the Hoffman-Wielandt theorem; and those for generalized eigenspaces are counterparts of the celebrated Davis-Kahan  $\sin \theta$ ,  $\sin 2\theta$  theorems for the eigenspace variations of Hermitian matrices.

The paper consists of two parts. Part I is for generalized eigenvalue perturbations, while Part II deals with generalized eigenspace perturbations.

### 1. INTRODUCTION

The study of perturbations of eigenvalues and eigenspaces of a matrix always demands a great deal of attention not only by operator theorists, but also by numerical analysts. During the past few years significant advances have been made in the perturbation theory for the generalized eigenvalue problem  $Ax = \lambda Bx$ . By now, almost all the celebrated perturbation theorems for the standard eigenvalue problem  $Ax = \lambda x$  have been generalized to the generalized eigenvalue problem by several authors, e.g., [3, 5, 6, 11–18, 21–22, 24, and 25–29]. Stewart's and Sun's book [25] is a very well-written comprehensive review for both the development of perturbation theory for the standard eigenvalue problem and that for the generalized eigenvalue problem.

As to the perturbation of eigenvalues of the standard eigenvalue problem  $Ax = \lambda x$ , we have the following well-known result due to several mathematicians. Let  $A$  and  $\tilde{A}$  be two  $n \times n$  Hermitian matrices, and let  $\lambda_1, \dots, \lambda_n$  and  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$  be their eigenvalues arranged in ascending order, respectively. Then for any unitarily invariant norm  $\|\cdot\|$  (for definition see §2 below),

$$(1.1) \quad \|\text{diag}(\lambda_1 - \tilde{\lambda}_1, \dots, \lambda_n - \tilde{\lambda}_n)\| \leq \|A - \tilde{A}\|.$$

---

Received by the editor November 16, 1992 and, in revised form, December 10, 1992.

1991 *Mathematics Subject Classification.* Primary 65F15, 15A22, 15A42.

*Key words and phrases.* Diagonalizable matrix pencil, definite pencils, real spectrum, perturbation bounds.

It is worth mentioning that if the norm used in (1.1) is the Frobenius norm, then, and only then, (1.1) even holds for normal matrices  $A$  and  $\tilde{A}$  with their eigenvalues arranged properly [8].

As to the perturbation of eigenspaces of the standard eigenvalue problem  $Ax = \lambda x$ , Davis and Kahan [4] studied the case for a Hermitian matrix (operator); their results are now known as  $\sin \theta$ ,  $\sin 2\theta$ ,  $\tan \theta$ ,  $\tan 2\theta$  theorems, and have influenced much later work on the same topic, e.g., Sun [29] and Stewart [24]. The case for a general matrix was studied by Stewart [22], and his method has a simple extension which can be used to deal with the generalized eigenvalue problem  $Ax = \lambda Bx$ .

The above-mentioned results are playing important roles in the perturbation theory for the standard eigenvalue problem; therefore it is of great importance to generalize them to the generalized eigenvalue problem.

A few generalizations of (1.1) to definite pencils were made by Stewart [24] and Sun [28]. Sun [29] extended the Davis and Kahan theorems [4] to definite pencils as well. Our generalizations in this paper are applicable not only to definite pencils, but also to a wider class of matrix pencils.

The paper is organized as follows. Preliminaries necessary to our presentation are outlined in §2. The main results for generalized eigenvalue perturbations are presented and proved in §3. Applications to definite pencils are given in §4. Section 5 contains the main results for generalized eigenspace perturbations, whose applications are given in §6. We conclude our paper with miscellaneous remarks in §7.

## 2. PRELIMINARIES

Throughout the paper, capital letters are for matrices, lowercase Latin letters for column vectors or scalars, and lowercase Greek letters for scalars;  $\mathbb{C}^{m \times n}$  denotes the set of  $m \times n$  complex matrices,  $\mathcal{U}_n \subset \mathbb{C}^{n \times n}$  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 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).  $A > 0$  ( $A \geq 0$ ) means that  $A$  is a positive definite (positive semidefinite) Hermitian matrix, and  $A > B$  ( $A \geq B$ ) means  $A, B$  Hermitian and  $A - B > 0$  ( $A - B \geq 0$ ). The matrix  $A^{1/2}$  is the unique positive definite (semidirect) square root of  $A \geq 0$ , and  $A^{-1/2} = (A^{1/2})^{-1}$  for  $A > 0$ . The matrices  $A^T$ ,  $A^H$ , 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

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

We will consider unitarily invariant norms  $|||\cdot|||$  of matrices. In this we follow Mirsky [20] and [25]. To say that the norm is unitarily invariant on  $\mathbb{C}^{m \times n}$  means it satisfies besides the usual properties of any norm, also

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

Two unitarily invariant norms used frequently are the spectral norm  $\|\cdot\|_2$  and the Frobenius norm  $\|\cdot\|_F$ . It is well known that any unitarily invariant norm  $|||\cdot|||$  on  $\mathbb{C}^{m \times n}$  corresponds to a *symmetric gauge function*  $\Phi(\xi_1, \dots, \xi_N)$ , where  $N = \min\{m, n\}$ , and vice versa. By extension according to this property,

we define a unitarily invariant norm  $|||\cdot|||$  on  $\mathbb{C}^{m_1 \times n_1}$  ( $m_1 \leq n, n_1 \leq n$ ) consistent with the original one as  $|||A||| = \Phi(\sigma_1, \dots, \sigma_{N_1}, 0, \dots, 0)$  if  $A \in \mathbb{C}^{m_1 \times n_1}$  with singular values  $\sigma_1, \dots, \sigma_{N_1}$  ( $N_1 = \min\{m_1, n_1\}$ ). In the second half of this paper, very often matrices with different dimensions enter our arguments together, so we make the following agreements: 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 smaller dimension there exists the extended unitarily invariant norm denoted also by  $|||\cdot|||$ . In this way, we have (see [25, Chapter 2, §3])

$$(2.1) \quad |||CD||| \leq \begin{cases} \|C\|_2 |||D||| \\ |||C||| \|D\|_2 \end{cases} \quad \text{for any } C \in \mathbb{C}^{m \times n}, D \in \mathbb{C}^{n \times l}.$$

Consider the pencil  $A - \lambda B$  with  $A, B \in \mathbb{C}^{n \times n}$  arbitrary constant matrices. The pencil is said to be *regular* if  $\det(A - \lambda B) \neq 0$ . Denote by

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

The pair  $(\alpha, \beta) \in \mathbf{G}_{1,2}$  is called a *generalized eigenvalue* of a regular pencil  $A - \lambda B$  if  $\det(\beta A - \alpha B) = 0$ . Nonzero vectors  $x, y \in \mathbb{C}^n$  are termed the *generalized eigenvector* (the *right generalized eigenvector*, sometimes) 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  $\xi \neq 0$ . The pair  $(\alpha, \beta) \in \mathbf{G}_{1,2}$  is said to be *real* if there exists  $0 \neq \xi \in \mathbb{C}$  such that  $\xi\alpha, \xi\beta \in \mathbb{R}$ ; for an instance,  $(i, i)$  is real. The spectrum of a regular pencil  $A - \lambda B$  consists of its all generalized eigenvalues (counted according to their algebraic multiplicities), and is denoted by  $\lambda(A, B)$ . (And similarly,  $\lambda(A)$  denotes the spectrum of the square matrix  $A$ .)

Roughly speaking, the class of matrix pencils treated here is that of diagonalizable pencils with real spectra.

**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

$$(2.2) \quad \begin{cases} Y^H A X = \Lambda \equiv \text{diag}(\alpha_1, \dots, \alpha_n), \\ Y^H B X = \Omega \equiv \text{diag}(\beta_1, \dots, \beta_n). \end{cases}$$

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

To study the perturbation of generalized eigenvalues, we need metrics on  $\mathbf{G}_{1,2}$  and on the space of matrix pencils. Let  $(\alpha, \beta), (\gamma, \delta) \in \mathbf{G}_{1,2}$ . G. W. Stewart [21] was the first one who used the *chordal metric* on the Riemannian sphere,

$$(2.3) \quad \rho((\alpha, \beta), (\gamma, \delta)) \stackrel{\text{def}}{=} \frac{|\delta\alpha - \gamma\beta|}{\sqrt{|\alpha|^2 + |\beta|^2} \sqrt{|\gamma|^2 + |\delta|^2}},$$

to measure the difference between the two points. We shall adopt the chordal metric, too. To measure the difference between two regular pencils  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  of order  $n$ , Sun [27] was the first one who realized that metrics on the *Grassmann manifold* of all  $n \times 2n$  matrices having full row rank are more

suitable than any other natural metrics such as  $\|(A - \tilde{A}, B - \tilde{B})\|_2$ . In this paper, we will employ

$$(2.4) \quad d_2(Z, \tilde{Z}) \stackrel{\text{def}}{=} \|\sin \Theta_r(Z, \tilde{Z})\|_2 = \|I - Z_1 \tilde{Z}_1^H \tilde{Z}_1 Z_1^H\|_2^{1/2} \\ = \|P_{Z^H} - P_{\tilde{Z}^H}\|_2,$$

$$(2.5) \quad d_F(Z, \tilde{Z}) \stackrel{\text{def}}{=} \|\sin \Theta_r(Z, \tilde{Z})\|_F = \|I - Z_1 \tilde{Z}_1^H \tilde{Z}_1 Z_1^H\|_F^{1/2} \\ = \frac{1}{\sqrt{2}} \|P_{Z^H} - P_{\tilde{Z}^H}\|_F,$$

where

$$(2.6) \quad Z = (A, B), \quad \tilde{Z} = (\tilde{A}, \tilde{B}),$$

and the angular matrix  $\Theta_r(Y, \tilde{Y})$  between two matrices  $Y$  and  $\tilde{Y} \in \mathbb{C}^{m \times q}$  ( $1 \leq m \leq q$ ), both having full row rank, is defined by

$$(2.7) \quad \Theta_r(Y, \tilde{Y}) \stackrel{\text{def}}{=} \arccos((Y Y^H)^{-1/2} Y \tilde{Y}^H (\tilde{Y} \tilde{Y}^H)^{-1} \tilde{Y} Y^H (Y Y^H)^{-1/2})^{-1/2} \geq 0.$$

One must note that  $\text{rank } Z = \text{rank } \tilde{Z} = n$ , which can be easily verified by the regularity of  $A - \lambda B$  and that of  $\tilde{A} - \lambda \tilde{B}$ . For a detailed discussion of the above definitions and results, the reader is referred to [25, Chapter 2, §4]. In the literature, several pseudometrics on  $\mathbf{G}_{1,2}$  also entered the study for perturbations of generalized eigenvalues at whiles (see Elsner and Lancaster [5], Li [11, 13]). Although these pseudometrics are not metrics, they are all equivalent to the chordal metric.

The perturbation of generalized eigenspaces in Part II requires the definition of the angles between two subspaces with the same dimension. Let  $X_1, \tilde{X}_1 \in \mathbb{C}^{n \times l}$  ( $1 \leq l \leq n-1$ ) have full column rank. The angle between  $\mathcal{X}_1 = \mathcal{R}(X_1)$  and  $\tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{X}_1)$  is defined by

$$(2.8) \quad \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) \stackrel{\text{def}}{=} \text{diag}(\theta_1, \dots, \theta_l),$$

where  $\theta_1, \dots, \theta_l$  are the eigenvalues of  $\Theta_r(X_1^H, \tilde{X}_1^H)$ . Two different choices of the bases of  $\mathcal{X}_1$  and  $\tilde{\mathcal{X}}_1$  may result in two different angular matrices, which differ one from the other only by a unitary similarity transformation. Hence,  $\Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)$  is well defined.

**Lemma 2.1.** *Let  $X_1, \tilde{X}_1 \in \mathbb{C}^{n \times l}$  ( $1 \leq l \leq n-1$ ) with  $X_1^H X_1 = \tilde{X}_1^H \tilde{X}_1 = I$ , and let  $\mathcal{X}_1 = \mathcal{R}(X_1)$  and  $\tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{X}_1)$ . If  $\tilde{X} = (\tilde{X}_1, \tilde{X}_2) \in \mathcal{U}_n$ , then for any unitarily invariant norm  $\|\cdot\|$  we have  $\|\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\| = \|\tilde{X}_2^H X_1\|$ .*

For a proof of this lemma, the reader is referred to, e.g., [25, Chapter 1].

**Lemma 2.2.** *Let  $X_1, \tilde{X}_1 \in \mathbb{C}^{n \times l}$  ( $1 \leq l \leq n-1$ ) have full column rank. Suppose that  $\tilde{X} = (\tilde{X}_1, \tilde{X}_2) \in \mathbb{C}^{n \times n}$  is a nonsingular matrix with*

$$\tilde{X}^{-1} = \begin{pmatrix} \tilde{W}_1^H \\ \tilde{W}_2^H \end{pmatrix}, \quad \tilde{W}_1 \in \mathbb{C}^{n \times l}.$$

Set  $\mathcal{X}_1 = \mathcal{R}(X_1)$  and  $\tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{X}_1)$ . Then for any unitarily invariant norm  $\|\cdot\|$  we have

$$(2.9) \quad \|\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\| = \|(\tilde{W}_2^H \tilde{W}_2)^{-1/2} \tilde{W}_2^H X_1 (X_1^H X_1)^{-1/2}\|.$$

*Proof.* We have

$$I^{(n)} = \tilde{X}^{-1} \tilde{X} = \begin{pmatrix} \tilde{W}_1^H \\ \tilde{W}_2^H \end{pmatrix} (\tilde{X}_1, \tilde{X}_2) = \begin{pmatrix} \tilde{W}_1^H \tilde{X}_1 & \tilde{W}_1^H \tilde{X}_2 \\ \tilde{W}_2^H \tilde{X}_1 & \tilde{W}_2^H \tilde{X}_2 \end{pmatrix} \Rightarrow \tilde{W}_2^H \tilde{X}_1 = 0.$$

Let  $Y = (\tilde{X}_1 (\tilde{X}_1^H \tilde{X}_1)^{-1/2}, \tilde{W}_2 (\tilde{W}_2^H \tilde{W}_2)^{-1/2})$ . One can verify easily with the help of  $\tilde{W}_2^H \tilde{X}_1 = 0$  that  $Y^H Y = I$  implies  $Y \in \mathcal{U}_n$ . Now (2.9) is a consequence of Lemma 2.1.  $\square$

**Lemma 2.3** (Sun [29]). *Suppose  $A_1 \in \mathbb{C}^{m \times n_1}$ ,  $A_2 \in \mathbb{C}^{m \times n_2}$ ,  $m \leq n_1 + n_2$  and  $A_1 A_1^H + A_2 A_2^H = I$ ; then there exist  $U \in \mathcal{U}_m$ ,  $V_1 \in \mathcal{U}_{n_1}$  and  $V_2 \in \mathcal{U}_{n_2}$  such that*

$$A_1 = U \Sigma_1 V_1^H, \quad A_2 = U \Sigma_2 V_2^H,$$

where  $\Sigma_1 = \text{diag}(\alpha_1, \alpha_2, \dots)$  and  $\Sigma_2 = \text{diag}(\beta_1, \beta_2, \dots)$  satisfying  $\alpha_i \geq 0$ ,  $\beta_i \geq 0$  and  $|\alpha_i|^2 + |\beta_i|^2 = 1$  for  $i = 1, 2, \dots$ .

### PART I. PERTURBATION BOUNDS FOR GENERALIZED EIGENVALUES

This part is composed of three sections. In §3, we prove several perturbation theorems in each of which  $\|\cdot\|_2$  or  $\|\cdot\|_F$  is used. To this end, we also give a detailed analysis of a geometric representation of real generalized eigenvalues, which serves as an essential tool in our proofs, as well as a basic one in Part II below. The applications of the theorems in §3 to definite pencils seem rather straightforward since definite pencils not only are diagonalizable but also have real spectra. However, extra attention has to be paid for exploiting the speciality of definite pencils. We consider such applications in §4. The theorems in this part are generalizations of some well-known perturbation theorems for the standard eigenvalue problem of Hermitian matrices, as is easily seen from Remarks 7.1–7.3 in the last section, §7, of this paper. It is shown that under perturbations, the generalized eigenvalues of a diagonalizable pencil with real spectrum behave just like the eigenvalues of a diagonalizable matrix with real spectrum in the sense 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).

Throughout the rest of this paper,  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  are always reserved for two regular matrix pencils, and  $Z, \tilde{Z} \in \mathbb{C}^{n \times 2n}$  are defined by (2.6).

#### 3. EIGENVALUE PERTURBATION OF A DIAGONALIZABLE PENCIL

The following theorems are two of our main results. They deal with the most general case treated in this paper.

**Theorem 3.1.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  both have only real generalized eigenvalues, and let*

$$\begin{aligned} \lambda(A, B) &= \{(\alpha_i, \beta_i), i = 1, 2, \dots, n\}, \\ \lambda(\tilde{A}, \tilde{B}) &= \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = 1, 2, \dots, n\} \end{aligned}$$

with  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$ ,  $i, j = 1, 2, \dots, n$ . Assume they admit decompositions

$$(3.1a) \quad \begin{cases} Y^H A X = \Lambda, \\ Y^H B X = \Omega, \end{cases} \quad \text{and} \quad \begin{cases} \tilde{Y}^H \tilde{A} \tilde{X} = \tilde{\Lambda}, \\ \tilde{Y}^H \tilde{B} \tilde{X} = \tilde{\Omega}, \end{cases}$$

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

$$(3.1b) \quad \begin{cases} \Lambda = \text{diag}(\alpha_1, \dots, \alpha_n), \\ \Omega = \text{diag}(\beta_1, \dots, \beta_n), \end{cases} \quad \text{and} \quad \begin{cases} \tilde{\Lambda} = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n), \\ \tilde{\Omega} = \text{diag}(\tilde{\beta}_1, \dots, \tilde{\beta}_n). \end{cases}$$

Then there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that

$$(3.2) \quad \begin{aligned} & \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \\ & \leq \|\tilde{X}_2\| \left\| \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X & \\ & X \end{pmatrix} \right\|_2 \|X^{-1}\|_2 \\ & \leq \kappa(\tilde{X})\kappa(X)d_2(Z, \tilde{Z}), \end{aligned}$$

where  $\kappa(X) = \|X\|_2 \|X^{-1}\|_2$  is the spectral condition number of  $X$ .

**Theorem 3.2.** Under the conditions of Theorem 3.1, there exists a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that

$$(3.3) \quad \begin{aligned} & \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \\ & \leq \max\{\|\tilde{X}^{-1}\|_2, \|\tilde{X}\|_2\}^2 \max\{\|X^{-1}\|_2, \|X\|_2\}^2 d_F(Z, \tilde{Z}). \end{aligned}$$

To prove the above theorems, we need the help of results developed in §§3.1–3.2 below.

**3.1. Geometric representation of a real generalized eigenvalue.** We have already noted that a generalized eigenvalue is invariant under a nonzero factor; hence for a given real  $(\alpha, \beta) \in \mathbf{G}_{1,2}$ , without loss of generality, we may assume

$$(3.4) \quad \alpha, \beta \in \mathbb{R}, \quad \beta \geq 0, \quad \alpha^2 + \beta^2 = 1, \quad \text{and} \quad \alpha = 1 \text{ if } \beta = 0.$$

Thus, there is a 1-1 correspondence between the set of real generalized eigenvalues and the upper half  $\Gamma$  of the unit circle, containing  $(1, 0)$  but not  $(-1, 0)$ , in the following way. The pair  $(\alpha, \beta)$  satisfying (3.4) corresponds to a point  $z \in \Gamma$  as shown in Figure 3.1. In the other direction, every  $z \in \Gamma$  determines an element of  $\mathbf{G}_{1,2}$  satisfying (3.4) by its coordinates. For convenience, in the following, we treat  $z$  and  $(\alpha, \beta)$  equally and write  $z = (\alpha, \beta)$ . The symbol  $\Gamma$  is always reserved for the upper half of the unit circle.

If  $z, w \in \Gamma$  correspond to  $(\alpha, \beta)$  and  $(\gamma, \delta)$ , respectively, they have only four different relative positions. By treating  $z$  and  $w$  symmetrically, we consider the two cases described by Figures 3.2 and 3.3.

We define

$$\begin{aligned} \widehat{zw} & \stackrel{\text{def}}{=} \text{the circular arc corresponding to angle } \theta, \\ \widehat{wz} & \stackrel{\text{def}}{=} \bigcup_{i=1}^2 \{\text{the circular arc corresponding to angle } \theta_i\} \end{aligned}$$

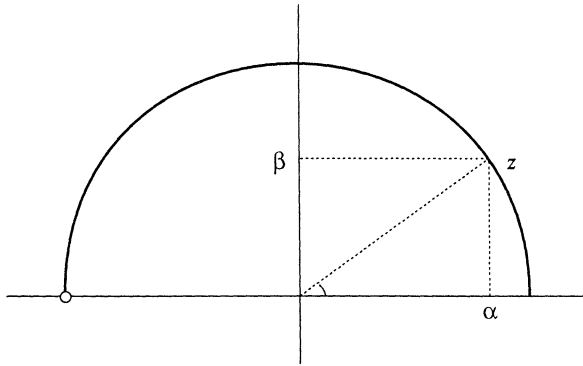


FIGURE 3.1. Geometric representation:  $z = (\alpha, \beta)$

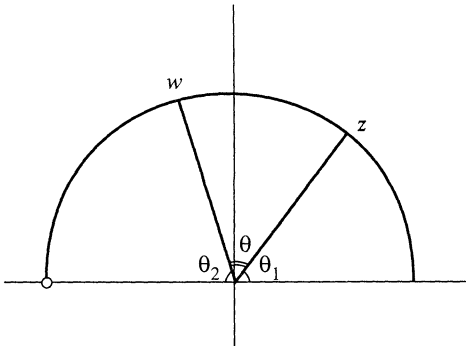


FIGURE 3.2. The case  $\theta < 90^\circ$

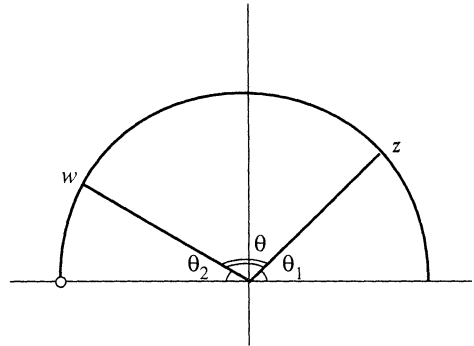


FIGURE 3.3. The case  $\theta > 90^\circ$

and define the anticlockwise angles between  $z$  and  $w$  by

$$\theta(z, w) \stackrel{\text{def}}{=} \theta = 180^\circ - (\theta_1 + \theta_2), \quad \theta(w, z) \stackrel{\text{def}}{=} \theta_1 + \theta_2 = 180^\circ - \theta.$$

The notation  $z \prec w$  means  $\theta(z, w) < 90^\circ$ ; therefore,  $z \prec w$  in Figure 3.2 and  $w \prec z$  in Figure 3.3. The distance on  $\Gamma$  is given by

$$d(z, w) \stackrel{\text{def}}{=} \rho((\alpha, \beta), (\gamma, \delta)).$$

It is easy to prove that  $(\Gamma, d)$  is a complete metric space (see Proposition 3.1 below).

We will use the notation  $(z w v)$  to mean that the points  $z, w, v \in \Gamma$  appear in counterclockwise cyclic order on  $\Gamma$ , as shown by Figure 3.4 (next page). Note that this order relation can also be described by one of  $(w v z)$  and  $(v z w)$ . We deal with more than three points in a similar way.

**Proposition 3.1.** *Let  $z, w$  be as described in Figure 3.2 or Figure 3.3; then*

$$\rho((\alpha, \beta), (\gamma, \delta)) = \sin \theta = \sin(\theta_1 + \theta_2).$$

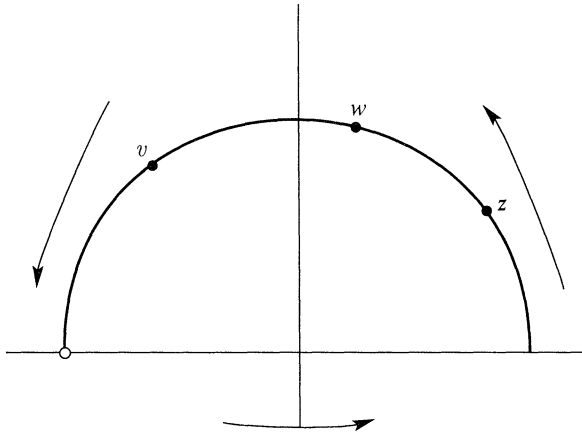


FIGURE 3.4. Ordering

*Proof.* Because of  $0 \leq \theta < 180^\circ$ , we have  $\sin \theta \geq 0$ . Thus,

$$\begin{aligned} \sin \theta &= \sqrt{\sin^2 \theta} = \sqrt{1 - \cos^2 \theta} = \sqrt{1 - (\alpha\gamma + \beta\delta)^2} \\ &= \sqrt{(\delta\alpha - \gamma\beta)^2} = |\delta\alpha - \gamma\beta| = \frac{|\delta\alpha - \gamma\beta|}{\sqrt{|\alpha|^2 + |\beta|^2} \sqrt{|\gamma|^2 + |\delta|^2}}, \end{aligned}$$

as claimed.  $\square$

Two connected circular arcs  $\Gamma_1$  and  $\Gamma_2$ , which are separated from each other, have also only four different relative positions. By treating them symmetrically, we consider the following two cases described by Figures 3.5 and 3.6. To understand the connectedness of  $\Gamma_2$  in Figure 3.6, one must regard the two points  $(-1, 0)$  and  $(1, 0)$  as a single one. This view is much the same as for quotient space in topology. The separation between  $\Gamma_1$  and  $\Gamma_2$  may be described by one of the following description. (The first “=” in (3.5a) below can be verified easily.)

(i)

$$(3.5a) \quad \min_{\substack{z \in \Gamma_1 \\ w \in \Gamma_2}} d(z, w) = \min\{\sin \theta_1, \sin \theta_2\} \stackrel{\text{def}}{=} \eta,$$

where  $\theta_2 = \theta_{21} + \theta_{22}$  in the case described by Figure 3.5.

(ii) There exist  $a \in \Gamma$ ,  $\alpha \geq 0$ ,  $\delta > 0$ ,  $\alpha + \delta \leq 1$ , such that

$$(3.5b) \quad \max_{z \in \Gamma_1} d(z, a) \leq \alpha \quad \text{and} \quad \min_{w \in \Gamma_2} d(w, a) \geq \alpha + \delta,$$

or

$$(3.5c) \quad \min_{z \in \Gamma_1} d(z, a) \geq \alpha + \delta \quad \text{and} \quad \max_{w \in \Gamma_2} d(w, a) \leq \alpha.$$

Description (ii) has already been used in Sun [29] to develop perturbation bounds for eigenspaces of a definite pencil. We claim that the two descriptions are equivalent. A geometric interpretation of (3.5a) is clearly seen from Figure 3.5 and/or Figure 3.6, whereas that of (3.5b) is clearly seen from Figure 3.7



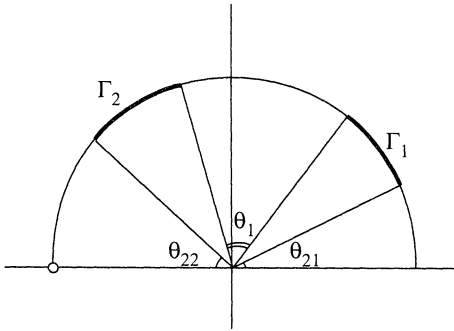


FIGURE 3.5.  $\Gamma_1$  is separated from  $\Gamma_2$

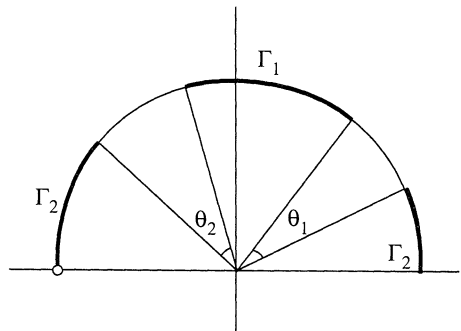


FIGURE 3.6.  $\Gamma_1$  is separated from  $\Gamma_2$

below, where  $\sin \phi = \alpha$ ,  $\sin(\theta + \phi) = \alpha + \delta$ , and  $\theta = \min\{\theta_1, \theta_2\}$ ; similarly for (3.5c).

It is easy to prove

**Proposition 3.2.** *If (3.5b) or (3.5c) holds, then (3.5a) holds with*

$$(3.6) \quad \eta \geq (\alpha + \delta)\sqrt{1 - \alpha^2} - \alpha\sqrt{1 - (\alpha + \delta)^2}.$$

On the other hand, if (3.5a) holds, it is not difficult to find  $a \in \Gamma$ ,  $\alpha \geq 0$ ,  $\delta > 0$ ,  $\alpha + \delta \leq 1$ , such that (3.5b) and/or (3.5c) holds.

Given a set  $\mathbf{K}$  consisting of several generalized eigenvalues, hereafter we prefer to use the notation  $\mathbf{K} \subset \Gamma_1$  (or  $\mathbf{K} \subset \Gamma_2$ ) which means that all points in  $\Gamma$  corresponding to elements of  $\mathbf{K}$  belong to  $\Gamma_1$  (or to  $\Gamma_2$ ).

**3.2. Auxiliary propositions.**

**Proposition 3.3.** *Let  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$  and*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1 \quad \text{for } i, j = 1, 2, \dots, n.$$

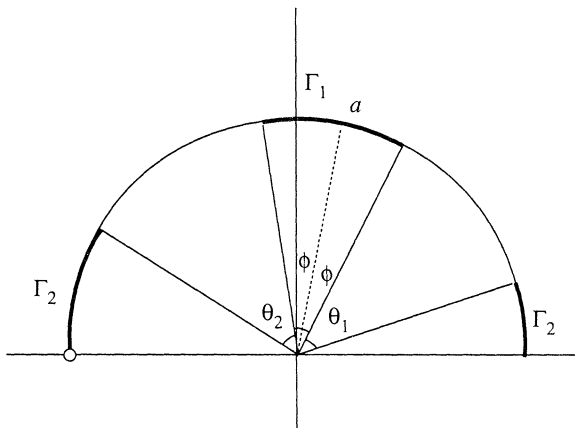


FIGURE 3.7. A geometric interpretation of (3.5b)

Define  $\Lambda$ ,  $\Omega$ ,  $\tilde{\Lambda}$ , and  $\tilde{\Omega}$  by (3.1b) in terms of these numbers. Let  $\mathbf{K}_1$  be a set containing  $k$  of the  $(\alpha_i, \beta_i)$  and  $\mathbf{K}_2$  a set containing  $l$  of the  $(\tilde{\alpha}_j, \tilde{\beta}_j)$  with  $k + l > n$ . Assume also that  $\Gamma_1$  and  $\Gamma_2$  are as described in one of the Figures 3.5 and 3.6. If  $\mathbf{K}_1 \subset \Gamma_1$  and  $\mathbf{K}_2 \subset \Gamma_2$ , or vice versa, then

$$(3.7) \quad \|Q^{-1}\|_2 \|\tilde{\Lambda}Q\Omega - \tilde{\Omega}Q\Lambda\|_2 \geq \eta,$$

where  $\eta$  is defined by (3.5a), and  $Q \in \mathbb{C}^{n \times n}$  is nonsingular.

*Proof.* Without loss of generality, assume that  $\mathbf{K}_1 = \{(\alpha_i, \beta_i), i = 1, 2, \dots, k\}$  and  $\mathbf{K}_2 = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = 1, 2, \dots, l\}$ , since for suitable  $n \times n$  permutation matrices  $P$  and  $\tilde{P}$ , the first  $k$  diagonal elements of the diagonal matrices  $P^T \Lambda P$  and  $P^T \Omega P$  and the first  $l$  diagonal elements of the diagonal matrices  $\tilde{P}^T \tilde{\Lambda} \tilde{P}$  and  $\tilde{P}^T \tilde{\Omega} \tilde{P}$  correspond to the elements of  $\mathbf{K}_1$  and  $\mathbf{K}_2$  in order, respectively. Now replace  $Q$ ,  $\Lambda$ ,  $\Omega$ ,  $\tilde{\Lambda}$ , and  $\tilde{\Omega}$  by  $\tilde{P}^T Q P$ ,  $P^T \Lambda P$ ,  $P^T \Omega P$ ,  $\tilde{P}^T \tilde{\Lambda} \tilde{P}$ , and  $\tilde{P}^T \tilde{\Omega} \tilde{P}$ , respectively. Having done so, we shall still have  $\|(\tilde{P}^T Q P)^{-1}\|_2 = \|Q^{-1}\|_2$  and

$$\|\tilde{P}^T \tilde{\Lambda} \tilde{P} \tilde{P}^T Q P P^T \Omega P - \tilde{P}^T \tilde{\Omega} \tilde{P} \tilde{P}^T Q P P^T \Lambda P\|_2 = \|\tilde{\Lambda}Q\Omega - \tilde{\Omega}Q\Lambda\|_2$$

by a simple verification. Set

$$\begin{cases} \Lambda_1 = \text{diag}(\alpha_1, \dots, \alpha_k), & \Lambda_2 = \text{diag}(\alpha_{k+1}, \dots, \alpha_n), \\ \Omega_1 = \text{diag}(\beta_1, \dots, \beta_k), & \Omega_2 = \text{diag}(\beta_{k+1}, \dots, \beta_n), \\ \tilde{\Lambda}_1 = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_l), & \tilde{\Lambda}_2 = \text{diag}(\tilde{\alpha}_{l+1}, \dots, \tilde{\alpha}_n), \\ \tilde{\Omega}_1 = \text{diag}(\tilde{\beta}_1, \dots, \tilde{\beta}_l), & \tilde{\Omega}_2 = \text{diag}(\tilde{\beta}_{l+1}, \dots, \tilde{\beta}_n); \end{cases}$$

then  $\Lambda = \text{diag}(\Lambda_1, \Lambda_2)$ ,  $\Omega = \text{diag}(\Omega_1, \Omega_2)$ , and  $\tilde{\Lambda} = \text{diag}(\tilde{\Lambda}_1, \tilde{\Lambda}_2)$ ,  $\tilde{\Omega} = \text{diag}(\tilde{\Omega}_1, \tilde{\Omega}_2)$ . Partition  $Q$  as  $Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$  with  $Q_{11} \in \mathbb{C}^{l \times k}$ . Then

$$E \stackrel{\text{def}}{=} \tilde{\Lambda}Q\Omega - \tilde{\Omega}Q\Lambda = \begin{pmatrix} \tilde{\Lambda}_1 Q_{11} \Omega_1 - \tilde{\Omega}_1 Q_{11} \Lambda_1 & \tilde{\Lambda}_1 Q_{12} \Omega_2 - \tilde{\Omega}_1 Q_{12} \Lambda_2 \\ \tilde{\Lambda}_2 Q_{21} \Omega_1 - \tilde{\Omega}_2 Q_{21} \Lambda_1 & \tilde{\Lambda}_2 Q_{22} \Omega_2 - \tilde{\Omega}_2 Q_{22} \Lambda_2 \end{pmatrix},$$

so

$$(3.8) \quad \|E\| \geq \|\tilde{\Lambda}_1 Q_{11} \Omega_1 - \tilde{\Omega}_1 Q_{11} \Lambda_1\|_2.$$

To estimate further a lower bound for  $\|E\|_2$ , we consider the case of  $\mathbf{K}_1 \subset \Gamma_1$ ,  $\mathbf{K}_2 \subset \Gamma_2$ , and moreover the midpoint of arc  $\Gamma_1 = \widehat{zw}$  coincides with  $(0, 1)$  (refer to Figures 3.5 or 3.6). Assume that the angle between the  $\beta$ -axis and the ray  $Oz$  or  $Ow$  is  $\phi$ , and let  $\theta = \min\{\theta_1, \theta_2\}$ ; then

$$(3.9) \quad \begin{cases} \|\tilde{\Lambda}_1^{-1}\|_2^{-1} = \min_{1 \leq j \leq l} |\tilde{\alpha}_j| \geq \sin(\theta + \phi), & \|\Omega_1^{-1}\|_2^{-1} = \min_{1 \leq i \leq k} |\beta_i| \geq \cos \phi, \\ \|\tilde{\Omega}_1\|_2 = \max_{1 \leq j \leq l} |\tilde{\beta}_j| \leq \cos(\theta + \phi), & \|\Lambda_1\|_2 = \max_{1 \leq i \leq k} |\alpha_i| \leq \sin \phi. \end{cases}$$

Figure 3.8 shows a geometric interpretation of the inequalities in (3.9). Hence, it follows from (3.8) and  $\|Q_{11}\|_2 \geq \|Q^{-1}\|_2^{-1}$ , by Corollary 2.2 of Li [17], that

$$(3.10) \quad \begin{aligned} \|E\|_2 &\geq \|\tilde{\Lambda}_1 Q_{11} \Omega_1\|_2 - \|\tilde{\Omega}_1 Q_{11} \Lambda_1\|_2 \\ &\geq \|\tilde{\Lambda}_1^{-1}\|_2^{-1} \|Q_{11}\|_2 \|\Omega_1^{-1}\|_2^{-1} - \|\tilde{\Omega}_1\|_2 \|Q_{11}\|_2 \|\Lambda_1\|_2 \\ &\geq \|Q_{11}\|_2 (\sin(\theta + \phi) \cos \phi - \cos(\theta + \phi) \sin \phi) \\ &= \|Q_{11}\|_2 \sin \theta \geq \eta \|Q^{-1}\|_2^{-1}, \end{aligned}$$

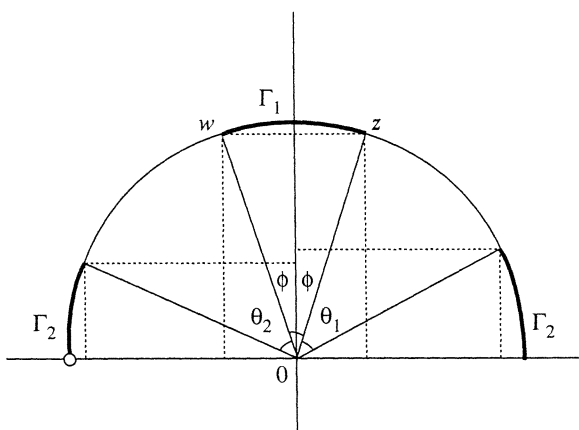


FIGURE 3.8. An interpretation of the inequalities (3.9)

which leads to the inequality (3.7) for the present case.

Now, consider the case when the midpoint  $m$  of  $\Gamma_1 = \widehat{zw}$  does not coincide with  $(0, 1)$ , i.e.,  $m = (\zeta, \xi) \neq (0, 1)$ . We apply a plane rotation in anticlockwise direction as

$$\begin{cases} \widehat{\Lambda}_1 = \zeta\Lambda - \zeta\Omega_1 \stackrel{\text{def}}{=} \text{diag}(\hat{\alpha}_1, \dots, \hat{\alpha}_k), \\ \widehat{\Omega}_1 = \zeta\Lambda_1 + \xi\Omega_1 \stackrel{\text{def}}{=} \text{diag}(\hat{\beta}_1, \dots, \hat{\beta}_k), \\ \widetilde{\Lambda}_1 = \xi\widetilde{\Lambda}_1 - \zeta\widetilde{\Omega}_1 \stackrel{\text{def}}{=} \text{diag}(\hat{\alpha}_1, \dots, \hat{\alpha}_l), \\ \widetilde{\Omega}_1 = \zeta\widetilde{\Lambda}_1 + \xi\widetilde{\Omega}_1 \stackrel{\text{def}}{=} \text{diag}(\hat{\beta}_1, \dots, \hat{\beta}_l), \end{cases}$$

and let

$$\begin{aligned} \widehat{\Gamma}_1 &= \{(\hat{\alpha}, \hat{\beta}) = (\alpha\xi - \beta\zeta, \alpha\zeta + \beta\xi) : (\alpha, \beta) \in \Gamma_1\} \subset \Gamma, \\ \widehat{\Gamma}_2 &= \left\{ (\hat{\alpha}, \hat{\beta}) = \begin{cases} (\tilde{\alpha}\xi - \tilde{\beta}\zeta, \tilde{\alpha}\zeta + \tilde{\beta}\xi) & \text{if } \tilde{\alpha}\zeta + \tilde{\beta}\xi > 0, \\ (-\tilde{\alpha}\xi + \tilde{\beta}\zeta, -\tilde{\alpha}\zeta - \tilde{\beta}\xi) & \text{otherwise,} \end{cases} \mid (\tilde{\alpha}, \tilde{\beta}) \in \Gamma_2 \right\} \subset \Gamma. \end{aligned}$$

It is easily verified that

- (i) the midpoint of  $\widehat{\Gamma}_1$  is  $\hat{m} = (0, 1)$ ;
- (ii) (3.5a) remains valid with  $\Gamma_1$  and  $\Gamma_2$  replaced by  $\widehat{\Gamma}_1$  and by  $\widehat{\Gamma}_2$ , respectively;
- (iii)  $(\hat{\alpha}_i, \hat{\beta}_i) \in \widehat{\Gamma}_1$ ,  $i = 1, 2, \dots, k$ , and for suitable  $d_j = \pm 1$  we have  $d_j\hat{\beta}_j \geq 0$  ( $d_j = -1$ , if  $\hat{\beta}_j = 0$ ), and thus  $(d_j\hat{\alpha}_j, d_j\hat{\beta}_j) \in \widehat{\Gamma}_2$ ,  $j = 1, 2, \dots, l$ ;
- (iv)  $\widehat{\Lambda}_1 Q_{11} \widehat{\Omega}_1 - \widetilde{\Omega}_1 Q_{11} \widehat{\Lambda}_1 = \widetilde{\Lambda}_1 Q_{11} \Omega_1 - \widetilde{\Omega}_1 Q_{11} \Lambda_1$ .

So from (3.9) we get

$$\|E\|_2 \geq \|\widehat{\Lambda}_1 Q_{11} \widehat{\Omega}_1 - \widetilde{\Omega}_1 Q_{11} \widehat{\Lambda}_1\|_2 = \|(D\widehat{\Lambda}_1)Q_{11}\widehat{\Omega}_1 - (D\widetilde{\Omega}_1)Q_{11}\widehat{\Lambda}_1\|_2,$$

where  $D = \text{diag}(d_1, \dots, d_l)$ . The rest of our proof is a repeat of the last part of the foregoing proof for the case when the midpoint of  $\Gamma_1$  is just  $(0, 1)$ .

The case for  $\mathbf{K}_1 \subset \Gamma_2$ ,  $\mathbf{K}_2 \subset \Gamma_1$  is clearly true by symmetry. The proof is completed.  $\square$

**Proposition 3.4.** *Let  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$  and*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1 \quad \text{for } i, j = 1, 2, \dots, n.$$

*Define  $\Lambda, \Omega, \tilde{\Lambda}$  and  $\tilde{\Omega}$  by (3.1b) in terms of these numbers. Then there is a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that*

$$(3.11) \quad \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq \|Q^{-1}\|_2 \|\tilde{\Lambda}Q\Omega - \tilde{\Omega}Q\Lambda\|_2,$$

where  $Q \in \mathbb{C}^{n \times n}$  is nonsingular.

*Proof.* Let the  $(\alpha_i, \beta_i)$  and  $(\tilde{\alpha}_j, \tilde{\beta}_j)$  correspond to  $z_i$  and  $\tilde{z}_j$ , respectively. We assume also, for convenience, that all the  $z_i$  and  $\tilde{z}_j$  are distinct, all the  $d(z_i, \tilde{z}_j)$  are distinct, and all  $d(z_i, \tilde{z}_j) < 1$ . By a density argument, it suffices to prove the theorem in this special case. In fact, for the moment suppose it is true in this special case. Let  $\{z_i^{(k)}\}_{k=1}^{+\infty}$  and  $\{\tilde{z}_j^{(k)}\}_{k=1}^{+\infty}$  be two sequences falling into this special case for each  $k$  such that  $z_i^{(k)} \rightarrow z_i, \tilde{z}_j^{(k)} \rightarrow \tilde{z}_j$  as  $k \rightarrow +\infty$ . Then there is a permutation  $\sigma_k$  corresponding to each  $\{z_i^{(k)}\}$  and  $\{\tilde{z}_j^{(k)}\}$ . Because the number of different permutations of  $\{1, 2, \dots, n\}$  is  $n!$ , a finite number,  $\{\sigma_k\}_{k=1}^{+\infty}$  contains a subsequence, say  $\{\sigma_{k_l}\}_{l=1}^{+\infty}$ , so that  $\sigma_{k_1} = \sigma_{k_2} = \dots$ . Taking the limit for such a subsequence, we will eventually see that the theorem is also true for  $z_i$  and  $\tilde{z}_j$ . Hereafter, indices will be understood to repeat cyclically, e.g.,  $z_{n+1} = z_1$ .

We do not deal explicitly with the case  $n = 2$  since it is very simple.

We number the  $z_i$  cyclically, i.e., we require  $(z_1 z_2 \dots z_n)$  (cf. §3.1 for notation). We define

$$(3.12) \quad \eta \stackrel{\text{def}}{=} \min_{\sigma} \max_{1 \leq i \leq n} d(z_i, \tilde{z}_{\sigma(i)}) = \min_{\sigma} \max_{1 \leq i \leq n} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})).$$

The numbering of the  $\tilde{z}_j$  is determined so that

$$(3.13) \quad \max_{i \in \mathcal{F}} d(z_i, \tilde{z}_i) \leq \max_{i \in \mathcal{F}} d(z_i, \tilde{z}_{\sigma(i)}) \text{ holds for any } \mathcal{F} \subset \{1, 2, \dots, n\} \text{ and for any permutation } \sigma \text{ of } \mathcal{F}.$$

Clearly, we now have  $\eta = \max_{1 \leq i \leq n} d(z_i, \tilde{z}_i)$ . Without loss of generality, assume this maximum is attained for  $i = 1$  and (in the notation introduced in §3.1)  $z_1 \prec \tilde{z}_1$ .

The following four claims are quite similar to Steps 1, 2, and 3 in [1, pp. 72–73], the influence of which the author wishes to acknowledge.

*Claim 1.* For any  $i$ , if  $\tilde{z}_i \prec z_i$ , then neither  $(z_i \tilde{z}_i \tilde{z}_1)$  nor  $(z_1 z_i \tilde{z}_1)$ .

This follows by applying (3.13) to the indices  $1, i$ .

*Claim 2.* If  $z_i \prec \tilde{z}_i$ , then for any  $j \neq i$ , neither  $(z_i z_j \tilde{z}_j \tilde{z}_i)$  nor  $(z_i \tilde{z}_j z_j \tilde{z}_i)$ ; if  $\tilde{z}_i \prec z_i$ , then for any  $j \neq i$ , neither  $(\tilde{z}_i z_j \tilde{z}_j z_i)$  nor  $(\tilde{z}_i \tilde{z}_j z_j z_i)$ .

These follow by applying (3.13) to the indices  $i, j$ .

*Claim 3.* There is a  $t$  such that  $d(z_{t+1}, \tilde{z}_t) > \eta$  and  $(z_1 \tilde{z}_1 \tilde{z}_t z_{t+1})$ .

As a matter of fact, for every  $j = 1, 2, \dots, n$ , had we  $d(z_{j+1}, \tilde{z}_j) < \eta$ , then we could pair each  $\tilde{z}_j$  with  $z_{j+1}$  and reduce the maximum distance in (3.12); but this contradicts the definition of  $\eta$ . For such  $t$ ,  $(z_1 \tilde{z}_1 \tilde{z}_t z_{t+1})$  follows from an application of Claim 1 to the index  $t$ , together with the inequalities imposed.

*Claim 4.* For  $1 < i < t$  we have  $(\tilde{z}_1 \tilde{z}_i \tilde{z}_t)$ .

This can be easily verified in case  $z_t \prec \tilde{z}_t$ , so in what follows we assume  $\tilde{z}_t \prec z_t$ . By Claim 3 and  $d(z_1, \tilde{z}_1) > d(z_t, \tilde{z}_t)$ , we see  $(z_1 \tilde{z}_1 \tilde{z}_t z_t z_{t+1})$ . Now if  $z_i \prec \tilde{z}_i$ , then unless  $(\tilde{z}_1 \tilde{z}_i \tilde{z}_t)$  we can get a contradiction to (3.13) for these two indices. So we need to consider the case  $\tilde{z}_i \prec z_i$ . Now  $(z_1 \tilde{z}_i \tilde{z}_1)$  is impossible by Claim 1;  $(\tilde{z}_t \tilde{z}_i z_t)$  is false by applying Claim 2 to the indices  $i, t$ ; and  $(z_t \tilde{z}_i z_1)$  can be ruled out, for otherwise, if  $(\tilde{z}_i z_i z_t)$ , then by  $\tilde{z}_i \prec z_i$ , we would have  $d(z_i, \tilde{z}_i) > d(z_1, \tilde{z}_1) = \eta$ , a contradiction, and if  $(z_1 z_i \tilde{z}_1)$ , our applying Claim 2 to the indices  $1, i$  would also lead to a contradiction. The remaining alternative is  $(\tilde{z}_1 \tilde{z}_i \tilde{z}_t)$ , so the verification of Claim 4 is completed.

Now, let  $\Gamma_1 = \widehat{z_{t+1} z_1}$ ,  $\mathbf{K}_1 = \{(\alpha_i, \beta_i); z_i \in \Gamma_1, i = t + 1, \dots, n, 1\}$ , and let  $\Gamma_2 = \widehat{\tilde{z}_1 \tilde{z}_t}$ ,  $\mathbf{K}_2 = \{(\tilde{\alpha}_j, \tilde{\beta}_j); \tilde{z}_j \in \Gamma_2, j = 1, 2, \dots, t\}$ . Obviously, we have

$$\min_{\substack{z \in \Gamma_1 \\ w \in \Gamma_2}} d(z, w) = d(z_1, \tilde{z}_1) = \eta$$

by (3.5a). By noting that  $\mathbf{K}_1$  contains  $n - t + 1$  of the  $(\alpha_i, \beta_i)$ , that  $\mathbf{K}_2$  contains  $t$  of the  $(\tilde{\alpha}_j, \tilde{\beta}_j)$ , and that  $(n - t + 1) + t = n + 1 > n$ , we see that  $\mathbf{K}_1, \mathbf{K}_2, \Gamma_1$ , and  $\Gamma_2$  satisfy the conditions of Proposition 3.3. Now applying this proposition leads to the inequality (3.11).  $\square$

**3.3. Proof of Theorem 3.1.** Without loss of generality, we may assume that all the  $(\alpha_i, \beta_i)$  and all the  $(\tilde{\alpha}_j, \tilde{\beta}_j)$  satisfy (3.4). Otherwise, set

$$D = \text{diag} \left( d_1 / \sqrt{\alpha_1^2 + \beta_1^2}, \dots, d_n / \sqrt{\alpha_n^2 + \beta_n^2} \right),$$

$$\tilde{D} = \text{diag} \left( \tilde{d}_1 / \sqrt{\tilde{\alpha}_1^2 + \tilde{\beta}_1^2}, \dots, \tilde{d}_n / \sqrt{\tilde{\alpha}_n^2 + \tilde{\beta}_n^2} \right).$$

By choosing  $d_i = \pm 1$  and  $\tilde{d}_j = \pm 1$  suitably, and replacing  $Y, \Lambda, \Omega$ , and  $\tilde{Y}, \tilde{\Lambda}, \tilde{\Omega}$  by  $YD, D\Lambda, D\Omega$ , and  $\tilde{Y}\tilde{D}, \tilde{D}\tilde{\Lambda}, \tilde{D}\tilde{\Omega}$ , respectively, we will get a decomposition with the needed properties.

It follows from (3.1) that

$$AXX^H B^H - BXX^H A^H = Y^{-H} \Lambda \Omega Y^{-1} - Y^{-H} \Omega \Lambda Y = 0$$

$$\Rightarrow Z^+(AXX^H B^H - BXX^H A^H) = 0;$$

hence

$$-\tilde{Z}^+(\tilde{A}X X^H B^H - \tilde{B}X X^H A^H)$$

$$= -\tilde{Z}^+(\tilde{A}X X^H B^H - \tilde{B}X X^H A^H) + Z^+(AXX^H B^H - BXX^H A^H)$$

$$= Z^+(A, B) \begin{pmatrix} X X^H & \\ & X X^H \end{pmatrix} \begin{pmatrix} B^H \\ -A^H \end{pmatrix}$$

$$- \tilde{Z}^+(\tilde{A}, \tilde{B}) \begin{pmatrix} X X^H & \\ & X X^H \end{pmatrix} \begin{pmatrix} B^H \\ -A^H \end{pmatrix}$$

$$= (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X X^H & \\ & X X^H \end{pmatrix} \begin{pmatrix} B^H \\ -A^H \end{pmatrix}.$$

Since  $\text{rank } \tilde{Z} = n$ ,  $\tilde{Z}\tilde{Z}^+ = I^{(n)}$ , premultiplying the two sides of the above equation by  $\tilde{Z}$  leads to

$$(3.14a) \quad \tilde{A}X X^H B^H - \tilde{B}X X^H A^H = -\tilde{Z}(P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X X^H & \\ & X X^H \end{pmatrix} \begin{pmatrix} B^H \\ -A^H \end{pmatrix}.$$

With the help of (3.1), we have

$$(3.14b) \quad \begin{aligned} & \tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda \\ &= -(\tilde{\Lambda}, \tilde{\Omega}) \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix} \stackrel{\text{def}}{=} E. \end{aligned}$$

So by Proposition 3.4, we see that there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that

$$\begin{aligned} & \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq \|(\tilde{X}^{-1} X)^{-1}\|_2 \|\tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda\|_2 \\ & \leq \|\tilde{X}\|_2 \|(\tilde{\Lambda}, \tilde{\Omega})\|_2 \left\| \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X & \\ & X \end{pmatrix} \right\|_2 \\ & \quad \times \left\| \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix} \right\|_2 \|X^{-1}\|_2 \\ & = \|\tilde{X}\|_2 \left\| \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X & \\ & X \end{pmatrix} \right\|_2 \|X^{-1}\|_2. \end{aligned}$$

The last “=” is true because  $\|(\tilde{\Lambda}, \tilde{\Omega})\|_2 = \max_{1 \leq j \leq n} \sqrt{\tilde{\alpha}_j^2 + \tilde{\beta}_j^2} = 1$  and  $\| \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix} \|_2 = \max_{1 \leq j \leq n} \sqrt{\alpha_j^2 + \beta_j^2} = 1$  by Li [11, Lemma 2.1].  $\square$

**3.4. Proof of Theorem 3.2.** Without loss of generality, we assume that (3.4) holds for all  $(\alpha_i, \beta_i)$  and  $(\tilde{\alpha}_j, \tilde{\beta}_j)$ . By treating  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  equally, one easily obtains the following identity from (3.14b):

$$(3.15) \quad \Lambda X^{-1} \tilde{X} \tilde{\Omega} - \Omega X^{-1} \tilde{X} \tilde{\Lambda} = -(\Lambda X^{-1}, \Omega X^{-1}) (P_{\tilde{Z}^H} - P_{Z^H}) \begin{pmatrix} \tilde{X} \tilde{\Omega} \\ -\tilde{X} \tilde{\Lambda} \end{pmatrix} \stackrel{\text{def}}{=} F.$$

To complete our proof, we need

**Lemma 3.1** (Sun [30]). *Let*

$$(3.16) \quad \Sigma = \text{diag}(\varepsilon_1, \dots, \varepsilon_n), \quad \varepsilon_j > 0 \text{ for } j = 1, 2, \dots, n,$$

and let  $M, N$  be two  $n \times n$  normal matrices with  $\lambda(M) = \{\lambda_1, \dots, \lambda_n\}$  and  $\lambda(N) = \{\mu_1, \dots, \mu_n\}$ . Then there exists a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that

$$\min_{1 \leq j \leq n} \varepsilon_j \sqrt{\sum_{j=1}^n |\lambda_j - \mu_{\tau(j)}|^2} \leq \|M \Sigma - \Sigma N\|_F.$$

**Lemma 3.2.** *Suppose that  $\Lambda, \Omega, \tilde{\Lambda}$ , and  $\tilde{\Omega}$  are defined by (3.1b) with  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$  and  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, i, j = 1, 2, \dots, n$ , and suppose  $X \in \mathbb{C}^{n \times n}$  is nonsingular; then there exists a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that*

$$(3.17) \quad \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \leq \|X^{-1}\|_2 \|\tilde{\Lambda} X \Omega - \Omega X \Lambda\|_F.$$

*Proof.* Let  $T = \tilde{\Lambda} X \Omega - \Omega X \Lambda$  and (here  $i = \sqrt{-1}$ )

$$\begin{pmatrix} \hat{\Lambda} \\ \hat{\Omega} \end{pmatrix} = \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \Lambda \\ \Omega \end{pmatrix}, \quad \begin{pmatrix} \hat{\tilde{\Lambda}} \\ \hat{\tilde{\Omega}} \end{pmatrix} = \begin{pmatrix} I & iI \\ iI & I \end{pmatrix} \begin{pmatrix} \tilde{\Lambda} \\ \tilde{\Omega} \end{pmatrix}.$$

We claim that the diagonal matrices  $\widehat{\Lambda}$ ,  $\widehat{\Omega}$ ,  $\widetilde{\Lambda}$ , and  $\widetilde{\Omega} \in \mathcal{U}_n$ . In fact,  $\widehat{\Lambda}\widehat{\Lambda}^H = (\Lambda + i\Omega)(\Lambda^H - i\Omega^H) = \Lambda^2 + \Omega^2 = I$  and  $\widehat{\Omega}\widehat{\Omega}^H = I$ . Similarly for  $\widetilde{\Lambda}$  and  $\widetilde{\Omega}$ . Hence, we have

$$(3.18) \quad \widehat{\Lambda}X\widehat{\Omega} - \widetilde{\Omega}X\widetilde{\Lambda} = 2T \Rightarrow \widehat{\Omega}^H\widehat{\Lambda}X - X\widetilde{\Lambda}\widehat{\Omega}^H = 2\widehat{\Omega}^H T\widehat{\Omega}^H.$$

Let  $X = U\Sigma V^H$  be the singular value decomposition of  $X$  with  $U, V \in \mathcal{U}_n$  and  $\Sigma$  defined by (3.16) with the singular values  $\varepsilon_j$  of  $X$ ; then from (3.18) it follows that

$$U^H\widehat{\Omega}^H\widehat{\Lambda}U\Sigma - \Sigma V^H\widetilde{\Lambda}\widehat{\Omega}^H V = 2U^H\widehat{\Omega}^H T\widehat{\Omega}^H V.$$

It is easy to verify that  $\widehat{U} \stackrel{\text{def}}{=} U^H\widehat{\Omega}^H\widehat{\Lambda}U$  and  $\widehat{V} \stackrel{\text{def}}{=} V^H\widetilde{\Lambda}\widehat{\Omega}^H V$  are normal matrices with

$$\begin{aligned} \lambda(\widehat{V}) &= \left\{ \lambda_k = \frac{\alpha_k + i\beta_k}{i\alpha_k + \beta_k}, k = 1, 2, \dots, n \right\}, \\ \lambda(\widehat{U}) &= \left\{ \mu_j = \frac{\tilde{\alpha}_j + i\tilde{\beta}_j}{i\tilde{\alpha}_j + \tilde{\beta}_j}, j = 1, 2, \dots, n \right\} \end{aligned}$$

and

$$\left| \frac{\alpha_k + i\beta_k}{i\alpha_k + \beta_k} - \frac{\tilde{\alpha}_j + i\tilde{\beta}_j}{i\tilde{\alpha}_j + \tilde{\beta}_j} \right| = \frac{2|\tilde{\beta}_j\alpha_k - \tilde{\alpha}_j\beta_k|}{\sqrt{|\alpha_k|^2 + |\beta_k|^2}\sqrt{|\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2}}.$$

So by Lemma 3.1, we have

$$\begin{aligned} 2\|T\|_F &= \|2U^H\widehat{\Omega}^H T\widehat{\Omega}^H V\|_F \geq \min_{1 \leq j \leq n} \varepsilon_j \sqrt{\sum_{j=1}^n |\lambda_j - \mu_{\tau(j)}|^2} \\ &= 2\|X^{-1}\|_2^{-1} \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2}. \end{aligned}$$

This establishes (3.17).  $\square$

Now, we return to the proof of Theorem 3.2. Lemma 3.2, (3.14b) and (3.15) yield that there exists a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that

$$(3.19a) \quad \sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2 \leq \|\tilde{X}^{-1}X\|_2^2 \|E\|_F^2,$$

$$(3.19b) \quad \sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2 \leq \|X^{-1}\|_2^2 \|F\|_F^2.$$

On the other hand,

$$(3.20) \quad \begin{aligned} \|E\|_F^2 + \|F\|_F^2 &= \left\| \begin{pmatrix} E & \\ & F^H \end{pmatrix} \right\|_F^2 \\ &\leq \left\| \begin{pmatrix} \widetilde{\Lambda}\tilde{X}^{-1} & \widetilde{\Omega}\tilde{X}^{-1} \\ \widetilde{\Omega}\tilde{X}^H & -\widetilde{\Lambda}\tilde{X}^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix} \right\|_F^2. \end{aligned}$$

We claim that

$$(3.21a) \quad \left\| \begin{pmatrix} \tilde{\Lambda}\tilde{X}^{-1} & \tilde{\Omega}\tilde{X}^{-1} \\ \tilde{\Omega}\tilde{X}^H & -\tilde{\Lambda}\tilde{X}^H \end{pmatrix} \right\|_2 \leq \max\{\|\tilde{X}^{-1}\|_2, \|\tilde{X}\|_2\},$$

$$(3.21b) \quad \left\| \begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix} \right\|_2 \leq \max\{\|X^{-1}\|_2, \|X\|_2\}.$$

In fact, it is easy to verify that

$$\begin{aligned} & \begin{pmatrix} \tilde{\Lambda}\tilde{X}^{-1} & \tilde{\Omega}\tilde{X}^{-1} \\ \tilde{\Omega}\tilde{X}^H & -\tilde{\Lambda}\tilde{X}^H \end{pmatrix} \begin{pmatrix} \tilde{\Lambda}\tilde{X}^{-1} & \tilde{\Omega}\tilde{X}^{-1} \\ \tilde{\Omega}\tilde{X}^H & -\tilde{\Lambda}\tilde{X}^H \end{pmatrix}^H \\ &= \begin{pmatrix} \Lambda_2\tilde{X}^{-1}\tilde{X}^{-H}\Lambda_2 + \Omega_2\tilde{X}^{-1}\tilde{X}^{-H}\Omega_2 & 0 \\ 0 & \Lambda_2\tilde{X}^H\tilde{X}\Lambda_2 + \Omega_2\tilde{X}^H\tilde{X}\Omega_2 \end{pmatrix} \end{aligned}$$

and

$$\begin{aligned} & \begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix}^H \begin{pmatrix} X\Omega & X^{-H}\Lambda \\ -X\Lambda & X^H\Omega \end{pmatrix} \\ &= \begin{pmatrix} \Lambda_1 X^H X \Lambda_1 + \Omega_1 X^H X \Omega_1 & 0 \\ 0 & \Lambda_1 X^{-H} X^{-1} \Lambda_1 + \Omega_1 X^{-H} X^{-1} \Omega_1 \end{pmatrix}. \end{aligned}$$

Thus, (3.21) follows from

$$\begin{aligned} \Lambda Q^H Q \Lambda + \Omega Q^H Q \Omega &\leq \|Q\|_2^2 (\Lambda \Lambda + \Omega \Omega) = \|Q\|_2^2 I, \\ \tilde{\Lambda} Q^H Q \tilde{\Lambda} + \tilde{\Omega} Q^H Q \tilde{\Omega} &\leq \|Q\|_2^2 (\tilde{\Lambda} \tilde{\Lambda} + \tilde{\Omega} \tilde{\Omega}) = \|Q\|_2^2 I, \end{aligned}$$

where  $Q = \tilde{X}^{-H}$ ,  $\tilde{X}$ ,  $X$ , or  $X^{-1}$ . Now, by (3.19)–(3.21), we have

$$\begin{aligned} & \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \\ & \leq \frac{1}{\sqrt{2}} \max\{\|\tilde{X}^{-1}X\|_2, \|X^{-1}\tilde{X}\|_2\} \sqrt{\|E\|_F^2 + \|F\|_F^2} \\ & \leq \max\{\|\tilde{X}^{-1}\|_2, \|\tilde{X}\|_2\}^2 \max\{\|X^{-1}\|_2, \|X\|_2\}^2 \frac{1}{\sqrt{2}} \|P_{Z^H} - P_{\tilde{Z}^H}\|_F. \end{aligned}$$

This establishes (3.3).  $\square$

**3.5. More bounds.** It may be unsatisfactory to some readers that our bounds are expressed by means of metrics on the Grassmann manifold of all  $n \times 2n$  matrices having full row rank, rather than our familiar metrics such as  $\|Z - \tilde{Z}\|_2$ . So in this subsection, we show how to deduce perturbation bounds containing only metrics such as  $\|Z - \tilde{Z}\|_2$  or  $\|Z - \tilde{Z}\|_F$  from our previous bounds. Also, we will illustrate how to establish more bounds along the lines of our proofs of Theorems 3.1, 3.2.

**Lemma 3.3.** *Let  $M, N \in \mathbb{C}^{q \times t}$  ( $q > t$ ) have full column rank; then*

$$P_M^\perp (N - M)(N^H N)^{-1/2}, P_N^\perp (N - M)(M^H M)^{-1/2},$$

and  $\sin \Theta(\mathcal{M}, \mathcal{N})$  have the same nonzero singular values, where  $\mathcal{M} = \mathcal{R}(M)$ ,  $\mathcal{N} = \mathcal{R}(N)$ ,  $P_M^\perp = I - P_M$ , and  $P_N^\perp = I - P_N$ . Moreover, if their nonzero



singular values are  $\sigma_1, \sigma_2, \dots$ , the nonzero singular values of  $P_N - P_M$  are  $\sigma_1, \sigma_1, \sigma_2, \sigma_2, \dots$ . Therefore,

$$\begin{aligned} \|P_N - P_M\|_2 &= \|P_M^\perp(N - M)(N^H N)^{-1/2}\|_2 = \|P_N^\perp(N - M)(M^H M)^{-1/2}\|_2 \\ &= \|\sin \Theta(\mathcal{M}, \mathcal{N})\|_2, \\ \frac{1}{\sqrt{2}}\|P_N - P_M\|_F &= \|P_M^\perp(N - M)(N^H N)^{-1/2}\|_F = \|P_N^\perp(N - M)(M^H M)^{-1/2}\|_F \\ &= \|\sin \Theta(\mathcal{M}, \mathcal{N})\|_F. \end{aligned}$$

The reader is referred to Li [16, Lemma 3.1] for a proof of this lemma, due originally to Sun.

The trivial combinations of this lemma with Theorems 3.1 and 3.2 produce some new bounds containing  $\|Z - \tilde{Z}\|_2$  or  $\|Z - \tilde{Z}\|_F$  immediately. To keep the paper reasonably short, we dispense with giving them explicitly.

Recall that identities such as (3.14) and (3.15) play an exclusive role in our proofs. In a similar way, one can get the following ones. The relation

$$(3.22a) \quad \tilde{A}X X^H B^H - \tilde{B}X X^H A^H = -(Z - \tilde{Z}) \begin{pmatrix} X X^H & \\ & X X^H \end{pmatrix} \begin{pmatrix} B^H \\ -A^H \end{pmatrix}$$

yields

$$(3.22b) \quad \tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda = -\tilde{Y}^H (Z - \tilde{Z}) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix}.$$

By treating  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  equally, (3.22b) produces

$$(3.23) \quad \Lambda X^{-1} \tilde{X} \tilde{\Omega} - \Omega X^{-1} \tilde{X} \tilde{\Lambda} = -Y^H (\tilde{Z} - Z) \begin{pmatrix} \tilde{X} & \\ & \tilde{X} \end{pmatrix} \begin{pmatrix} \tilde{\Omega} \\ -\tilde{\Lambda} \end{pmatrix}.$$

Now, one can easily prove the following theorem along the lines in the proofs of Theorems 3.1 and 3.2.

**Theorem 3.3.** *To the hypotheses of Theorem 3.1 add these:*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

*Then there exist permutations  $\sigma$  and  $\tau$  of  $\{1, 2, \dots, n\}$  such that*

$$(3.24) \quad \begin{aligned} &\max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \\ &\leq \min\{\kappa(X)\|\tilde{X}\|_2\|\tilde{Y}^H\|_2, \kappa(\tilde{X})\|X\|_2\|Y^H\|_2\}\|Z - \tilde{Z}\|_2 \end{aligned}$$

and

$$(3.25) \quad \begin{aligned} &\sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \\ &\leq \min\{\kappa(X)\|\tilde{X}\|_2\|\tilde{Y}^H\|_2, \kappa(\tilde{X})\|X\|_2\|Y^H\|_2\}\|Z - \tilde{Z}\|_F. \end{aligned}$$

It is very difficult to give a thorough comparison of (3.24) with (3.2), and of (3.25) with (3.3). However, we claim that the right-hand sides of (3.24) and (3.2) are both bounded by

$$\kappa(\tilde{X})\kappa(X) \min\{\|Z^+\|_2, \|\tilde{Z}^+\|_2\}\|Z - \tilde{Z}\|_2.$$

As to (3.2), according to Lemma 3.3 and (2.4), this is evident; and as to (3.24), it is guaranteed by

**Lemma 3.4.** *Suppose that  $A - \lambda B$  is as described in Theorem 3.1. If  $|\alpha_i|^2 + |\beta_i|^2 = 1$ ,  $i = 1, 2, \dots, n$ , then  $\|Y^H\|_2 \leq \|Z^+\|_2 \|X^{-1}\|_2$ .*

*Proof.* We have

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

as required.  $\square$

Another way of deriving perturbation bounds from the above bounds is described in the following subsection.

**3.6. Dual theorems.** It follows from Definition 2.1 that  $A - \lambda B \in \mathbf{D}_g(n)$  is equivalent to  $A^H - \lambda B^H \in \mathbf{D}_g(n)$ . Moreover, if  $A - \lambda B$  has only real generalized eigenvalues, so has  $A^H - \lambda B^H$ , and vice versa. A consequence of Theorem 3.1 is:

**Theorem 3.4.** *Under the conditions of Theorem 3.1, there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that*

$$\begin{aligned} &\max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \\ &\leq \|\tilde{Y}\|_2 \left\| \begin{pmatrix} \tilde{Y}^{-1} & \\ & \tilde{Y}^{-1} \end{pmatrix} (P_{Z_d} - P_{\tilde{Z}_d}) \begin{pmatrix} Y & \\ & Y \end{pmatrix} \right\|_2 \|Y^{-1}\|_2 \\ &\leq \kappa(Y) \kappa(\tilde{Y}) d_2(Z_d^H, \tilde{Z}_d^H), \end{aligned}$$

where

$$Z_d = \begin{pmatrix} A \\ B \end{pmatrix}, \quad \tilde{Z}_d = \begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix}.$$

We refer to the above theorem as the *dual theorem* of Theorem 3.1. In much the same way, we can derive theorems from other theorems such as Theorem 3.2, etc. We omit the details here.

Other perturbation theorems, e.g., Bauer and Fike-type theorems (Elsner and Sun [6], Li [11]) and Hoffman and Wielandt-type theorem (Sun [27]), for the generalized eigenvalue problem also have their own dual theorems just like Theorem 3.4 here. Readers who are interested in these are referred to Li [15].

#### 4. APPLICATIONS TO DEFINITE PENCILS

**Definition 4.1** ([3, 24]). Let  $A, B \in \mathbb{C}^{n \times n}$  be Hermitian matrices. The matrix  $A - \lambda B$  is said to be a *definite pencil of order  $n$* , if

$$(4.1) \quad c(A, B) \stackrel{\text{def}}{=} \min_{\substack{x \in \mathbb{C}^n \\ \|x\|_2 = 1}} |x^H (A + \sqrt{-1}B)x| > 0.$$

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

**Lemma 4.1.** *Let  $A - \lambda B \in \mathbf{D}(n)$ . Then there is a nonsingular matrix  $X \in \mathbb{C}^{n \times n}$  such that*

$$(4.2) \quad X^H A X = \Lambda \equiv \text{diag}(\alpha_1, \dots, \alpha_n), \quad X^H B X = \Omega \equiv \text{diag}(\beta_1, \dots, \beta_n).$$

In Lemma 4.1, it is easily verified that  $\alpha_i, \beta_i \in \mathbb{R}$ , and by an appropriate choice of  $X$ , we can make  $|\alpha_i|^2 + |\beta_i|^2 = 1$ . This lemma is then well known, and a proof of it can be found in Stewart [24] or in [25].

**Lemma 4.2.** *In Lemma 4.1, if  $|\alpha_i|^2 + |\beta_i|^2 = 1, i = 1, 2, \dots, n$ , then*

$$(4.3) \quad \|X\|_2 \leq \frac{1}{\sqrt{c(A, B)}}, \quad \|X^{-1}\|_2 \leq \frac{\|Z\|_2}{\sqrt{c(A, B)}}.$$

This lemma is implied in Elsner and Sun [6].

**Theorem 4.1.** *Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A} - \tilde{\lambda} \tilde{B} \in \mathbf{D}(n)$ , and let*

$$\begin{aligned} \lambda(A, B) &= \{(\alpha_i, \beta_i), i = 1, 2, \dots, n\}, \\ \lambda(\tilde{A}, \tilde{B}) &= \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = 1, 2, \dots, n\}, \end{aligned}$$

where  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}, i, j = 1, 2, \dots, n$ . Then there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that

$$(4.4a) \quad \max_{1 \leq j \leq n} \rho((\alpha_i, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq \frac{\|Z\|_2 \|\tilde{Z}\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} d_2(Z, \tilde{Z}),$$

$$(4.4b) \quad \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq \frac{\min\{\|Z\|_2, \|\tilde{Z}\|_2\}}{c(A, B)c(\tilde{A}, \tilde{B})} \|Z - \tilde{Z}\|_2.$$

*Proof.* By Lemma 4.1, we know that  $A - \lambda B$  and  $\tilde{A} - \tilde{\lambda} \tilde{B}$  admit decompositions

$$(4.5) \quad \begin{cases} X^H A X = \Lambda, \\ X^H B X = \Omega, \end{cases} \quad \text{and} \quad \begin{cases} \tilde{X}^H \tilde{A} \tilde{X} = \tilde{\Lambda}, \\ \tilde{X}^H \tilde{B} \tilde{X} = \tilde{\Omega}, \end{cases}$$

where  $X, \tilde{X} \in \mathbb{C}^{n \times n}$  are nonsingular matrices, and  $\Lambda, \Omega, \tilde{\Lambda}$ , and  $\tilde{\Omega}$  are of the form (3.1b) with  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$  and  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, i, j = 1, 2, \dots, n$ . So from Theorems 3.1 and 3.3 it follows that there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that

$$\begin{aligned} \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) &\leq \kappa(X)\kappa(\tilde{X})d_2(Z, \tilde{Z}), \\ \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) &\leq \min\{\kappa(X)\|\tilde{X}\|_2^2, \kappa(\tilde{X})\|X\|_2^2\} \|Z - \tilde{Z}\|_2. \end{aligned}$$

The assertions (4.4a) and (4.4b) now follow from Lemma 4.2.  $\square$

Analogously, Theorems 3.2 and 3.3, and Lemma 4.2, produce

**Theorem 4.2.** *Under the conditions of Theorem 4.1, there exists a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that*

$$(4.6a) \quad \begin{aligned} &\sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \\ &\leq \frac{\max\{1, \|Z\|_2^2\} \max\{1, \|\tilde{Z}\|_2^2\}}{c(A, B)c(\tilde{A}, \tilde{B})} d_F(Z, \tilde{Z}), \end{aligned}$$

$$(4.6b) \quad \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \leq \frac{\min\{\|Z\|_2, \|\tilde{Z}\|_2\}}{c(A, B)c(\tilde{A}, \tilde{B})} \|Z - \tilde{Z}\|_F.$$

Now, we are obliged to say something about the existing perturbation bounds for definite pencils in comparison with our bounds. Our conclusion is that our bounds and the bound in Stewart [24] and its improvement in Sun [28] are independent. Sometimes ours are much sharper than those in [24], as we will show, and sometimes they are weaker. The following is a slightly improved version, due to Li [18], of Stewart’s bound [24] (see Sun [28] for another improved version).

**Theorem 4.3.** *Let  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A}, \tilde{B} \in \mathbb{C}^{n \times n}$  be two Hermitian matrices. If  $\|\tilde{Z} - Z\|_2 < c(A, B)$ , then  $\tilde{A} - \lambda\tilde{B} \in \mathbf{D}(n)$ , and moreover, there exists a permutation  $\sigma$  of  $\{1, 2, \dots, n\}$  such that*

$$(4.7) \quad \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \leq \frac{\|\tilde{Z} - Z\|_2}{c(A, B)}.$$

**Example** (Sun [28]). Let  $A - \lambda B \in \mathbf{D}(n)$ . Consider  $\tilde{A} - \lambda\tilde{B} = (1 + r)(A - \lambda B)$  with  $r > 0$  such that

$$\frac{r\sqrt{\|A\|_2^2 + \|B\|_2^2}}{c(A, B)} < 1, \quad \frac{r\sqrt{\|A\|_2^2 + \|B\|_2^2}}{c(A, B)} \approx 1.$$

Obviously  $\tilde{A} - \lambda\tilde{B} \in \mathbf{D}(n)$  and  $\lambda(A, B) = \lambda(\tilde{A}, \tilde{B})$ , and the right-hand side of (4.4a) is 0, whereas the right-hand side of (4.7) is  $r\sqrt{\|A\|_2^2 + \|B\|_2^2}/c(A, B) \approx 1$ . This example shows that sometimes (4.4a) is much sharper than (4.7) (and than (4.4b) as well). The improved bound of [24] from Sun [28] also provides a best estimate for this example, just like ours.

The following observations, on the other hand, show that sometimes (4.4a) may be weaker than (4.7). By Lemma 3.3 and (2.4),

$$(4.8) \quad d_2(Z, \tilde{Z}) \leq \min\{\|(\tilde{Z}\tilde{Z}^H)^{-1/2}\|_2, \|(ZZ^H)^{-1/2}\|_2\}\|\tilde{Z} - Z\|_2$$

and

$$(4.9a) \quad \begin{aligned} \|(ZZ^H)^{-1/2}\|_2^{-1} &= \sqrt{\min_{\substack{x \in \mathbb{C}^n \\ \|x\|_2=1}} (x^H A^2 x + x^H B^2 x)} \\ &\geq \sqrt{\min_{\substack{x \in \mathbb{C}^n \\ \|x\|_2=1}} |x^H A x + \sqrt{-1} x^H B x|^2} = c(A, B), \end{aligned}$$

$$(4.9b) \quad \|(\tilde{Z}\tilde{Z}^H)^{-1/2}\|_2^{-1} \geq c(\tilde{A}, \tilde{B}).$$

The “ $\geq$ ” in the inequality (4.9a) holds because  $y^2$  is a convex function. As a matter of fact, suppose  $A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n)U^H$ ,  $U \in \mathcal{U}_n$ . (This is possible, since  $A$  is Hermitian.) We write  $Ux = (\xi_1, \dots, \xi_n)^T$ ; then

$$x^H A^2 x = \sum_{i=1}^n |\xi_i|^2 \lambda_i^2 \geq \left( \sum_{i=1}^n |\xi_i|^2 \lambda_i \right)^2 = (x^H A x)^2.$$

With the help of (4.8) and (4.9), we weaken the inequality (4.4a) by

$$(4.10) \quad \begin{aligned} & \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})) \\ & \leq \frac{\|Z\|_2 \|\tilde{Z}\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \frac{\|\tilde{Z} - Z\|_2}{\max\{c(A, B), c(\tilde{A}, \tilde{B})\}}. \end{aligned}$$

Now a rough comparison of (4.4a) with (4.7) may be done through the comparison of (4.10) with (4.7). First, we see that the difference between

$$\frac{\|\tilde{Z} - Z\|_2}{c(A, B)} \quad \text{and} \quad \frac{\|\tilde{Z} - Z\|_2}{\max\{c(A, B), c(\tilde{A}, \tilde{B})\}}$$

is negligible if  $\tilde{A} - \lambda\tilde{B}$  is close enough to  $A - \lambda B$ ; hence we ignore their contributions, since we are interested only in a rough comparison. Second, we note that the factor

$$(4.11) \quad \frac{\|Z\|_2 \|\tilde{Z}\|_2}{c(A, B)c(\tilde{A}, \tilde{B})}$$

appearing in the right-hand side of (4.10) is, generally, greater than 1. This is because  $\|Z\|_2 = \|(ZZ^H)^{1/2}\|_2 \geq \|(ZZ^H)^{-1/2}\|_2^{-1} \geq c(A, B)$  and  $\|\tilde{Z}\|_2 \geq c(\tilde{A}, \tilde{B})$  for the same reason. Thus, this factor allows us to conclude that (4.4a) is less sharp, especially when (4.11) is very large.

Also, it is very troublesome to compare other bounds, e.g., (4.4b) with (4.4a), (4.4b) with (4.10), and (4.4b) with (4.7), although a rough comparison can be made in a way similar as above.

## PART II. PERTURBATION BOUNDS FOR GENERALIZED EIGENSPACES

In this part, we prove several  $\sin \theta$  and  $\sin 2\theta$  theorems for diagonalizable pencils with real spectra in §5. Then in §6, we apply them to definite pencils, which lead to more perturbation bounds for definite pencils. This part is related to Davis and Kahan [4] and Sun [29]. All theorems in this part are generalizations of well-known perturbation theorems [4] for the standard eigenvalue problem of Hermitian matrices.

### 5. EIGENSPACE PERTURBATION OF A DIAGONALIZABLE PENCIL

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

$$(5.1a) \quad \begin{cases} Y^H A X = \Lambda, \\ Y^H B X = \Omega, \end{cases} \quad \text{and} \quad \begin{cases} \tilde{Y}^H \tilde{A} \tilde{X} = \tilde{\Lambda}, \\ \tilde{Y}^H \tilde{B} \tilde{X} = \tilde{\Omega}, \end{cases}$$

where

$$(5.1b) \quad \begin{cases} X = (X_1, X_2), \quad \tilde{X} = (\tilde{X}_1, \tilde{X}_2), \quad Y = (Y_1, Y_2), \quad \tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2), \\ X^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix}, \quad \tilde{X}^{-1} = \begin{pmatrix} \tilde{W}_1^H \\ \tilde{W}_2^H \end{pmatrix}, \quad Y^{-1} = \begin{pmatrix} S_1^H \\ S_2^H \end{pmatrix}, \quad \tilde{Y}^{-1} = \begin{pmatrix} \tilde{S}_1^H \\ \tilde{S}_2^H \end{pmatrix}, \\ X_1, \tilde{X}_1, Y_1, \tilde{Y}_1, W_1, \tilde{W}_1, S_1, \tilde{S}_1 \in \mathbb{C}^{n \times l}, \end{cases}$$

$$(5.2a) \quad \begin{cases} \Lambda = \text{diag}(\alpha_1, \dots, \alpha_n) \stackrel{\text{def}}{=} \text{diag}(\Lambda_1, \Lambda_2), \\ \Omega = \text{diag}(\beta_1, \dots, \beta_n) \stackrel{\text{def}}{=} \text{diag}(\Omega_1, \Omega_2), \\ \Lambda_1, \Omega_1 \in \mathbb{C}^{l \times l}, \quad \Lambda_2, \Omega_2 \in \mathbb{C}^{(n-l) \times (n-l)}, \end{cases}$$

$$(5.2b) \quad \begin{cases} \tilde{\Lambda} = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n) \stackrel{\text{def}}{=} \text{diag}(\tilde{\Lambda}_1, \tilde{\Lambda}_2), \\ \tilde{\Omega} = \text{diag}(\tilde{\beta}_1, \dots, \tilde{\beta}_n) \stackrel{\text{def}}{=} \text{diag}(\tilde{\Omega}_1, \tilde{\Omega}_2), \\ \tilde{\Lambda}_1, \tilde{\Omega}_1 \in \mathbb{C}^{l \times l}, \quad \tilde{\Lambda}_2, \tilde{\Omega}_2 \in \mathbb{C}^{(n-l) \times (n-l)}. \end{cases}$$

It is easy to see that two subspaces spanned by the first  $l$  column vectors of  $X$  and  $Y$  are a (right) eigenspace (left) eigenspace of  $A - \lambda B$  corresponding to  $(\alpha_i, \beta_i)$ ,  $i = 1, 2, \dots, l$ , respectively. Set

$$(5.3) \quad \mathcal{X}_1 = \mathcal{R}(X_1), \quad \tilde{\mathcal{X}}_1 = \mathcal{R}(\tilde{X}_1).$$

5.1. **sin  $\theta$  theorems.** The following theorem is one of our main results in this part.

**Theorem 5.1.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  both have only real generalized eigenvalues and admit decompositions (5.1)–(5.2). Let  $\mathbf{K}_1 = \{(\alpha_i, \beta_i), i = 1, 2, \dots, l\}$ ,  $\tilde{\mathbf{K}}_2 = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = l + 1, \dots, n\}$ . Assume that  $\Gamma_1$  and  $\Gamma_2$  are as described by one of Figures 3.5 and 3.6, and  $\eta$  is defined by (3.5a). If  $\mathbf{K}_1 \subset \Gamma_1$  and  $\tilde{\mathbf{K}}_2 \subset \Gamma_2$ , or vice versa, then for any unitarily invariant norm  $\|\cdot\|$  there holds*

$$(5.4) \quad \begin{aligned} & \|\sin \Theta(\mathcal{X}, \tilde{\mathcal{X}}_1)\| \\ & \leq \frac{\|X_1^+\|_2 \|\tilde{W}_2^+\|_2}{\eta} \left\| \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\|. \end{aligned}$$

Remark 5.1. From (5.4) it follows in a straightforward manner that

$$(5.4') \quad \begin{aligned} & \|\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\| \\ & \leq \begin{cases} \frac{\kappa(X_1)\kappa(\tilde{W}_2)}{\eta} \left\| \begin{pmatrix} \tilde{W}_{20}^H & \\ & \tilde{W}_{20}^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\|, \\ \frac{\kappa(X)\kappa(\tilde{X})}{\eta} \left\| \begin{pmatrix} \tilde{W}_{20}^H & \\ & \tilde{W}_{20}^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\|, \end{cases} \end{aligned}$$

where  $X_{10} = X_1(X_1^H X_1)^{-1/2}$ ,  $\tilde{W}_{20} = \tilde{W}_2(\tilde{W}_2^H \tilde{W}_2)^{-1/2}$ , and  $\kappa(X_1) = \|X_1\|_2 \|X_1^+\|_2$  is the spectral condition number of  $X_1$ , and similarly for  $\kappa(\tilde{W}_2)$ , etc.

*Proof of Theorem 5.1.* Without loss of generality, assume that (3.4) holds for all the  $(\alpha_i, \beta_i)$  and  $(\tilde{\alpha}_j, \tilde{\beta}_j)$ ; otherwise, let

$$D = \text{diag} \left( d_1/\sqrt{|\alpha_1|^2 + |\beta_1|^2}, \dots, d_n/\sqrt{|\alpha_n|^2 + |\beta_n|^2} \right),$$

$$\tilde{D} = \text{diag} \left( \tilde{d}_1/\sqrt{|\tilde{\alpha}_1|^2 + |\tilde{\beta}_1|^2}, \dots, \tilde{d}_n/\sqrt{|\tilde{\alpha}_n|^2 + |\tilde{\beta}_n|^2} \right),$$

and choose  $d_i, \tilde{d}_j$  with  $|d_i| = |\tilde{d}_j| = 1$  appropriately, so that the decompositions obtained by replacing  $Y, \Lambda, \Omega$ , and  $\tilde{Y}, \tilde{\Lambda}, \tilde{\Omega}$  in (5.1)–(5.2) by  $YD^H, D\Lambda, D\Omega$ , and  $\tilde{Y}\tilde{D}^H, \tilde{D}\tilde{\Lambda}, \tilde{D}\tilde{\Omega}$ , have the desired property.

Now (3.14) is obviously true. In what follows, we present another proof of (3.14b), which has the advantage that it is valid even for complex generalized eigenvalues, although all generalized eigenvalues encountered here are real.

From (5.1) it follows that

$$AXY^H B - BXY^H A = Y^{-H} \Lambda \Omega X^{-1} - Y^{-H} \Omega \Lambda X^{-1} = 0$$

$$\Rightarrow Z^+(AXY^H B - BXY^H A) = 0;$$

hence

$$-\tilde{Z}^+(\tilde{A}XY^H B - \tilde{B}XY^H A)$$

$$= -\tilde{Z}^+(\tilde{A}XY^H B - \tilde{B}XY^H A) + Z^+(AXY^H B - BXY^H A)$$

$$= Z^+(A, B) \begin{pmatrix} XY^H & \\ & XY^H \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix} - \tilde{Z}^+(\tilde{A}, \tilde{B}) \begin{pmatrix} XY^H & \\ & XY^H \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}$$

$$= (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} XY^H & \\ & XY^H \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}.$$

Since  $\text{rank } \tilde{Z} = n$ ,  $\tilde{Z}\tilde{Z}^+ = I^{(n)}$ , premultiplying the two sides of the above equation by  $\tilde{Z}$  leads to

$$(5.5a) \quad \tilde{A}XY^H B - \tilde{B}XY^H A = -\tilde{Z}(P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} XY^H & \\ & XY^H \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}.$$

Substituting (5.1) into this equation, we get

$$(5.5b) \quad \tilde{\Lambda}\tilde{X}^{-1}X\Omega - \tilde{\Omega}\tilde{X}^{-1}X\Lambda$$

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

This is nothing but (3.14b). On the other hand, by (5.1b) and (5.2), the right-hand side is

$$\tilde{\Lambda}\tilde{X}^{-1}X\Omega - \tilde{\Omega}\tilde{X}^{-1}X\Lambda$$

$$= \begin{pmatrix} \tilde{\Lambda}_1 \tilde{W}_1^H X_1 \Omega_1 - \tilde{\Omega}_1 \tilde{W}_1^H X_1 \Lambda_1 & \tilde{\Lambda}_1 \tilde{W}_1^H X_2 \Omega_2 - \tilde{\Omega}_1 \tilde{W}_1^H X_2 \Lambda_2 \\ \tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1 & \tilde{\Lambda}_2 \tilde{W}_2^H X_2 \Omega_2 - \tilde{\Omega}_2 \tilde{W}_2^H X_2 \Lambda_2 \end{pmatrix}.$$

So by (5.5b), we have

$$(5.6) \quad \tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1$$

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

Now we bound  $\| \|\tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1\| \|$  from below. This can be done in the same way as we did in the proof of Proposition 3.3. In fact, we have (refer to (3.10) for the first “ $\geq$ ”)

$$\begin{aligned}
 (5.7) \quad \| \|E\| \| &\geq \| \|\tilde{W}_2^H X_1\| \sin \theta \\
 &= \| \|(\tilde{W}_2^H \tilde{W}_2)^{1/2} (\tilde{W}_2^H \tilde{W}_2)^{-1/2} \tilde{W}_2^H X_1 (X_1^H X_1)^{-1/2} (X_1^H X_1)^{1/2}\| \| \sin \theta \\
 &\geq \eta \| \|(\tilde{W}_2^H \tilde{W}_2)^{-1/2}\|_2^{-1} \| \|\tilde{W}_{20}^H X_{10}\| \| \| \| (X_1^H X_1)^{-1/2}\|_2^{-1} \\
 &= \eta \| \|\tilde{W}_2^+\|_2^{-1} \| \|X_1^+\|_2^{-1} \| \| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\| \| .
 \end{aligned}$$

Besides Lemma 2.2, we have also employed the fact that  $\| \|(\tilde{W}_2^H \tilde{W}_2)^{-1/2}\|_2^{-1}$  equals the smallest singular value of  $\tilde{W}_2$ , which is  $\| \|\tilde{W}_2^+\|_2^{-1}$ , and similarly for  $\| \| (X_1^H X_1)^{-1/2}\|_2^{-1}$ .

On the other hand, from (5.6) it follows that

$$\| \|E\| \| \leq \| \| \left\| \begin{pmatrix} \tilde{W}_2^H & \\ & \tilde{W}_2^H \end{pmatrix} (P_{Z^H} - P_{\tilde{Z}^H}) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\| \| ,$$

since  $\tilde{\Lambda}_2^2 + \tilde{\Omega}_2^2 = I$ ,  $\Lambda_1^2 + \Omega_1^2 = I$ . This inequality together with (5.7) leads to (5.4).  $\square$

For  $\| \cdot \|_F$ , the inequality (5.4) holds even under a weaker condition imposed on the separation of  $\mathbf{K}_1$  and  $\tilde{\mathbf{K}}_2$ , and holds for complex generalized eigenvalues, as indicated in the following theorem.

**Theorem 5.2.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  admit the decompositions (5.1)–(5.2), but it is not necessary for them to have only real generalized eigenvalues. Set*

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

If  $\eta > 0$ , then (5.4) holds for the unitarily invariant norm  $\| \cdot \|_F$ .

*Proof.* We have argued that (5.5b) is true even for complex generalized eigenvalues; so is (5.6), from which it follows (denote  $\tilde{W}_2^H X_1 = (w_{ij}) \in \mathbb{C}^{(n-l) \times l}$ ) that

$$\begin{aligned}
 \| \|E\|_F^2 &= \| \|\tilde{\Lambda}_2 \tilde{W}_2^H X_1 \Omega_1 - \tilde{\Omega}_2 \tilde{W}_2^H X_1 \Lambda_1\|_F^2 \\
 &= \sum_{j=1}^{n-l} \sum_{i=1}^l |\tilde{\alpha}_{j+l} \beta_i - \tilde{\beta}_{j+l} \alpha_i|^2 |w_{ij}|^2 \geq \eta^2 \sum_{i,j} |w_{ij}|^2 = \eta^2 \| \|\tilde{W}_2^H X_1\|_F^2,
 \end{aligned}$$

implying

$$\| \|E\|_F \geq \eta \| \|\tilde{W}_2^+\|_2^{-1} \| \|X_1^+\|_2^{-1} \| \| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\|_F,$$

as claimed.  $\square$

The upper bounds in Theorems 5.1 and 5.2 contain  $P_{Z^H} - P_{\tilde{Z}^H}$  rather than  $\tilde{Z} - Z$  explicitly. However, the following theorems, obtained along the lines of the proofs of Theorem 5.1 and Theorem 5.2, have a different feature.



**Theorem 5.3.** *To the hypotheses of Theorem 5.1 add these:*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

Then

$$(5.8) \quad ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\|X_1^+\|_2 \|\tilde{W}_2^+\|_2}{\eta} \left\| \tilde{Y}^H (\tilde{Z} - Z) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\|.$$

*Remark 5.2.* More bounds may be deduced from (5.8) by using the following elementary inequalities:  $\|X_1^+\|_2 \leq \|X^{-1}\|_2$ ,  $\|\tilde{W}_2^+\|_2 \leq \|\tilde{X}\|_2$ ,  $\|\tilde{Y}\|_2 \leq \|\tilde{Y}\|_2$  and

$$(5.9) \quad \begin{aligned} \tilde{Y}^H \tilde{Z} \begin{pmatrix} \tilde{X} & \\ & \tilde{X} \end{pmatrix} = (\tilde{\Lambda}, \tilde{\Omega}) &\Rightarrow \tilde{Y}^{-H} = (\tilde{\Lambda}, \tilde{\Omega}) \begin{pmatrix} \tilde{X}^{-1} & \\ & \tilde{X}^{-1} \end{pmatrix} \tilde{Z}^H (\tilde{Z} \tilde{Z}^H)^{-1} \\ &\Rightarrow \|\tilde{Y}\|_2 \leq \|\tilde{X}^{-1}\|_2 \|\tilde{Z}^H (\tilde{Z} \tilde{Z}^H)^{-1}\|_2. \end{aligned}$$

As an example, one gets from (5.8)

$$||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\kappa(X) \|\tilde{X}\|_2 \|\tilde{Y}\|_2}{\eta} ||| \tilde{Z} - Z |||.$$

*Proof of Theorem 5.3.* By the proof of Theorem 5.1, it suffices to note that (refer to (3.22) and (3.23))

$$(5.10a) \quad \tilde{A} X Y^H B - \tilde{B} X Y^H A = (\tilde{Z} - Z) \begin{pmatrix} X Y^H & \\ & X Y^H \end{pmatrix} \begin{pmatrix} B \\ -A \end{pmatrix}$$

$$(5.10b) \quad \Rightarrow \tilde{\Lambda} \tilde{X}^{-1} X \Omega - \tilde{\Omega} \tilde{X}^{-1} X \Lambda = \tilde{Y}^H (\tilde{Z} - Z) \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Omega \\ -\Lambda \end{pmatrix},$$

which, just like (5.5), are both true, even for complex generalized eigenvalues.  $\square$

**Theorem 5.4.** *To the hypotheses of Theorem 5.2 add these:*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

Then (5.8) holds for the unitarily invariant norm  $||| \cdot ||| = \|\cdot\|_F$ .

**5.2.  $\sin 2\theta$  theorem.** As pointed out by Davis and Kahan [4],  $\sin 2\Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)$  is also a measure of the difference between the subspaces  $\mathcal{R}(X_1)$  and  $\mathcal{R}(\tilde{X}_1)$ .

**Theorem 5.5.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$  both have only real generalized eigenvalues and admit decompositions (5.1)–(5.2) with*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

Let  $\tilde{\mathbf{K}}_1 = \{(\tilde{\alpha}_i, \tilde{\beta}_i), i = 1, 2, \dots, l\}$ ,  $\tilde{\mathbf{K}}_2 = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = l + 1, \dots, n\}$ . Assume that  $\Gamma_1$  and  $\Gamma_2$  are described by one of the Figures 3.5 and 3.6, and  $\eta$  is defined by (3.5a). If  $\tilde{\mathbf{K}}_1 \subset \Gamma_1$  and  $\tilde{\mathbf{K}}_2 \subset \Gamma_2$ , or vice versa, then for any unitarily invariant norm  $||| \cdot |||$ , there holds

$$(5.11) \quad \begin{aligned} &||| \sin 2\Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| - 2\omega \|X_1\|_2 \|W_1\|_2 ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) |||^2 \\ &\leq \frac{\kappa(X)^3 \kappa(\tilde{X}) [1 + \kappa(Y) \kappa(X)] \|\tilde{X}\|_2 \|\tilde{Y}\|_2}{\eta} ||| \tilde{Z} - Z |||, \end{aligned}$$

where  $\omega = \| (W_1^H W_1)^{-1/2} W_1^H W_2 (W_2^H W_2)^{-1/2} \|_2$ .

*Remark 5.3.* More bounds may be deduced from (5.11) by using (5.9) and the following elementary inequalities:

$$\begin{aligned} \|X_1\|_2 &\leq \|X\|_2, \quad \|W_1\|_2 \leq \|X^{-1}\|_2 \Rightarrow \|X_1\|_2 \|W_1\|_2 \leq \kappa(X), \\ \| (W_1^H W_1)^{-1/2} W_1^H W_2 (W_2^H W_2)^{-1/2} \|_2 &\leq 1, \\ \|Y\|_2 &\leq \|X^{-1}\|_2 \|Z^H (Z Z^H)^{-1}\|_2, \quad \|Y^{-1}\|_2 \leq \|X\|_2 \|Z\|_2, \end{aligned}$$

since

$$Y^{-H} = Z \begin{pmatrix} X & \\ & X \end{pmatrix} \begin{pmatrix} \Lambda^H \\ \Omega^H \end{pmatrix}.$$

*Proof of Theorem 5.5.* Set

$$\begin{aligned} (5.12a) \quad R_X &= X \begin{pmatrix} I^{(l)} & \\ & -I^{(n-l)} \end{pmatrix} X^{-1} = (X_1, -X_2) \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix} \\ &= X_1 W_1^H - X_2 W_2^H, \end{aligned}$$

$$(5.12b) \quad R_Y = Y \begin{pmatrix} I^{(l)} & \\ & -I^{(n-l)} \end{pmatrix} Y^{-1} = (Y_1, -Y_2) Y^{-1},$$

from which we have

$$(5.13) \quad R_X^2 = R_Y^2 = I^{(n)}, \quad \|R_X\|_2 \leq \kappa(X), \quad \|R_Y\|_2 \leq \kappa(Y).$$

It follows from (5.1)–(5.2) that

$$\begin{aligned} Y^H (A + R_Y^H (\tilde{A} - A) R_X) X & \\ &= \Lambda + (Y_1, -Y_2)^H (\tilde{A} - A) (X_1, -X_2) \\ &= (Y_1, -Y_2)^H \tilde{A} (X_1, -X_2) \quad (\text{since } (Y_1, -Y_2)^H A (X_1, -X_2) = \Lambda) \\ &= (R_Y Y)^H \tilde{Y}^{-H} \tilde{\Lambda} \tilde{X}^{-1} R_X X, \end{aligned}$$

which implies

$$\hat{A} \stackrel{\text{def}}{=} A + R_Y^H (\tilde{A} - A) R_X = R_Y^H \tilde{Y}^{-H} \tilde{\Lambda} \tilde{X}^{-1} R_X.$$

Similarly, we define  $\hat{B} \stackrel{\text{def}}{=} B + R_Y^H (\tilde{B} - B) R_X = R_Y^H \tilde{Y}^{-H} \tilde{\Lambda} \tilde{X}^{-1} R_X$ . We see easily that  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}_g(n)$ . Hence by our hypotheses, the conditions of Theorem 5.3 are satisfied for the pencils  $\tilde{A} - \lambda \tilde{B}$  and  $\hat{A} - \lambda \hat{B}$ . Let  $\hat{X} \stackrel{\text{def}}{=} R_X^{-1} \tilde{X} = R_X \tilde{X} \equiv (\hat{X}_1, \hat{X}_2)$ , where  $\hat{X}_1 \in \mathbb{C}^{n \times l}$ , and let  $\hat{\mathcal{P}} = \mathcal{P}(\hat{X}_1)$ . By Lemma 2.2, we have

$$(5.14) \quad \|\sin \Theta(\hat{\mathcal{P}}, \tilde{\mathcal{P}})\| = \|(\tilde{W}_2^H, \tilde{W}_2)^{-1/2} \tilde{W}_2^H \hat{X}_1 (\hat{X}_1^H \hat{X}_1)^{-1/2}\|.$$

We attempt now to obtain an upper bound for the right-hand side of equation (5.14). Using  $\hat{X}_1 = R_X \tilde{X}_1$ , (5.12a), and

$$\begin{aligned} \tilde{W}_2^H (X_1 W_1^H + X_2 W_2^H) \tilde{X}_1 &= \tilde{W}_2^H X X^{-1} \tilde{X}_1 \\ &= 0 \Rightarrow \tilde{W}_2^H X_2 W_2^H \tilde{X}_1 = -\tilde{W}_2^H X_1 W_1^H \tilde{X}_1, \end{aligned}$$

we get

$$(5.15) \quad \tilde{W}_2^H \hat{X}_1 = \tilde{W}_2^H R_X \tilde{X}_1 = \tilde{W}_2^H (X_1 W_1^H - X_2 W_2^H) \tilde{X}_1 = 2 \tilde{W}_2^H X_1 W_1^H \tilde{X}_1.$$

On the other hand,  $(X_1(X_1^H X_1)^{-1/2}, W_2(W_2^H W_2)^{-1/2}) \in \mathcal{Z}_n$ ; therefore,

$$\begin{aligned} W_1 &= [X_1(X_1^H W_1)^{-1} X_1^H + W_2(W_2^H W_2)^{-1} W_2^H] W_1 \\ &= X_1(X_1^H X_1)^{-1} + W_2(W_2^H W_2)^{-1} W_2^H W_1 \quad (\text{since } X_1^H W_1 = I). \end{aligned}$$

Substituting the above equation into (5.15), we obtain

$$(5.16) \quad \widetilde{W}_2^H \widehat{X}_1 = 2\widetilde{W}_2^H X_1(X_1^H X_1)^{-1} X_1^H + 2\widetilde{W}_2^H X_1 W_1^H W_2(W_2^H W_2)^{-1} W_2^H \widetilde{X}_1.$$

Then, combining (5.14), (5.16) and (5.13) leads to

$$\begin{aligned} &||| \sin \Theta(\widehat{\mathcal{X}}_1, \widetilde{\mathcal{X}}_1) ||| \\ &= ||| (\widetilde{W}_2^H \widetilde{W}_2)^{-1/2} \widetilde{W}_2^H \widehat{X}_1 (\widetilde{X}_1^H R_X^H R_X \widetilde{X}_1)^{-1/2} ||| \\ &\geq ||| (\widetilde{W}_2^H \widetilde{W}_2)^{-1/2} \widetilde{W}_2^H \widehat{X}_1 (\widetilde{X}_1^H \widetilde{X}_1)^{-1/2} ||| \|R_X\|_2^{-1} \\ &= \|R_X\|_2^{-1} ||| \widetilde{W}_{20}^H X_{10} X_{10}^H \widetilde{X}_{10} + \widetilde{W}_{20}^H X_1 W_1^H W_{20} W_{20}^H \widetilde{X}_{10} |||; \end{aligned}$$

hence,

$$(5.17) \quad \begin{aligned} &\kappa(X) ||| \sin \Theta(\widehat{\mathcal{X}}_1, \widetilde{\mathcal{X}}_1) ||| \\ &\geq 2 ||| \widetilde{W}_{20}^H X_{10} X_{10}^H \widetilde{X}_{10} ||| \\ &\quad - 2 ||| \widetilde{W}_{20}^H X_{10} ||| \| (X_1^H X_1)^{1/2} \|_2 \| (W_1^H W_1)^{1/2} \|_2 \| W_{10}^H W_{20} \|_2 ||| W_{20}^H \widetilde{X}_{10} |||, \end{aligned}$$

where  $X_{10} = X_1(X_1^H X_1)^{-1/2}$ ,  $W_{10} = W_1(W_1^H W_1)^{-1/2}$ ,  $W_{20} = W_2(W_2^H W_2)^{-1/2}$ ,  $\widetilde{X}_{10} = \widetilde{X}_1(\widetilde{X}_1^H \widetilde{X}_1)^{-1/2}$ , and  $\widetilde{W}_{20} = \widetilde{W}_2(\widetilde{W}_2^H \widetilde{W}_2)^{-1/2}$ . The following fact is used for the first “ $\geq$ ”: for any matrix  $M$  with suitable dimension,

$$\begin{aligned} \widetilde{X}_1^H R_X^H R_X \widetilde{X}_1 &\leq \|R_X\|_2^2 \widetilde{X}_1^H \widetilde{X}_1 \\ &\Rightarrow (\widetilde{X}_1^H R_X^H R_X \widetilde{X}_1)^{-1} \geq \|R_X\|_2^{-2} (\widetilde{X}_1^H \widetilde{X}_1)^{-1} \\ &\Rightarrow M(\widetilde{X}_1^H R_X^H R_X \widetilde{X}_1)^{-1} M^H \geq \|R_X\|_2^{-2} M(\widetilde{X}_1^H \widetilde{X}_1)^{-1} M^H \\ &\Rightarrow ||| M(\widetilde{X}_1^H R_X^H R_X \widetilde{X}_1)^{-1/2} ||| \geq \|R_X\|_2^{-1} ||| M(\widetilde{X}_1^H \widetilde{X}_1)^{-1/2} |||. \end{aligned}$$

Now, we are able to embed  $\sin 2\Theta(\widehat{\mathcal{X}}_1, \widetilde{\mathcal{X}}_1)$  into the right-hand side of (5.17). It is easy to verify that  $(X_{10}, W_{20})$ ,  $(\widetilde{X}_{10}, \widetilde{W}_{20})$  are unitary matrices, and so is

$$(X_{10}, W_{20})^H (\widetilde{X}_{10}, \widetilde{W}_{20}) = \begin{pmatrix} X_{10}^H \widetilde{X}_{10} & X_{10}^H \widetilde{W}_{20} \\ W_{20}^H \widetilde{X}_{10} & W_{20}^H \widetilde{W}_{20} \end{pmatrix};$$

therefore, from Lemma 2.3 it follows that there exist unitary matrices with suitable dimensions  $U, V_1, V_2, V, U_1$ , and  $U_2$  such that

$$(5.18) \quad \begin{aligned} X_{10}^H \widetilde{X}_{10} &= U \Upsilon V_1^H, & X_{10}^H \widetilde{W}_{20} &= U \Sigma V^H, \\ \widetilde{X}_{10}^H X_{10} &= V \Upsilon^T U_1^H, & \widetilde{X}_{10}^H W_{20} &= V \Sigma^T U_2^H, \end{aligned}$$

where  $\Upsilon = \text{diag}(\gamma_1, \gamma_2, \dots)$ ,  $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots)$ ,  $\gamma_i = \cos \theta_i$ ,  $\sigma_i = \sin \theta_i$  for  $i = 1, 2, \dots$ , and  $\frac{\pi}{2} \geq \theta_1 \geq \theta_2 \geq \dots \geq 0$ . From (5.18) we get  $||| X_{10}^H \widetilde{W}_{20} ||| = ||| \Sigma ||| = ||| \widetilde{X}_{10}^H W_{20} |||$ . Therefore,

$$(5.19) \quad ||| \sin \Theta(\widehat{\mathcal{X}}_1, \widetilde{\mathcal{X}}_1) ||| = ||| X_{10}^H \widetilde{W}_{20} ||| = ||| \widetilde{X}_{10}^H W_{20} |||,$$

and

$$(5.20) \quad 2\|\|\widetilde{W}_{20}^H X_{10} X_{10}^H \widetilde{X}_{10}\|\| = 2\|\|V_1 \Upsilon^T \Sigma V_2^H\|\| = \|\|\text{diag}(\sin 2\theta_1, \sin 2\theta_2, \dots)\|\| \\ = \|\|\sin 2\Theta(\mathcal{X}_1, \widetilde{\mathcal{X}}_1)\|\|.$$

On the other hand, Theorem 5.3 tells us that (see Remark 5.2)

$$(5.21) \quad \|\|\sin \Theta(\widehat{\mathcal{X}}_1, \widetilde{\mathcal{X}}_1)\|\| \\ \leq \frac{1}{\eta} \kappa(R_X \widetilde{X}) \|\widetilde{X}\|_2 \|\widetilde{Y}\|_2 \|\|\widetilde{A} - \widehat{A}, \widetilde{B} - \widehat{B}\|\| \\ \leq \frac{1}{\eta} \kappa(X)^2 \kappa(\widetilde{X}) \|\widetilde{X}\|_2 \|\widetilde{Y}\|_2 \|\|\|(Z - \widetilde{Z}) - R_Y^H(Z - \widetilde{Z}) \begin{pmatrix} R_X & \\ & R_X \end{pmatrix}\|\|\| \\ \leq \frac{1}{\eta} \kappa(X)^2 \kappa(\widetilde{X}) \|\widetilde{X}\|_2 \|\widetilde{Y}\|_2 (1 + \|R_Y\|_2 \|R_X\|_2) \|Z - \widetilde{Z}\| \\ \leq \frac{1}{\eta} \kappa(X)^2 \kappa(\widetilde{X}) [1 + \kappa(X) \kappa(Y)] \|\widetilde{X}\|_2 \|\widetilde{Y}\|_2 \|Z - \widetilde{Z}\|.$$

Now substituting (5.19), (5.20), and (5.21) into (5.17) leads to (5.11).  $\square$

The following theorem indicates that the conditions of Theorem 5.5 can be weakened for  $\|\cdot\|_F$ .

**Theorem 5.6.** *Suppose  $A - \lambda B \in \mathbf{D}_g(n)$  and  $\widetilde{A} - \lambda \widetilde{B} \in \mathbf{D}_g(n)$  admit decompositions (5.1)–(5.2) with*

$$|\alpha_i|^2 + |\beta_i|^2 = |\widetilde{\alpha}_i|^2 + |\widetilde{\beta}_i|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

*(It is not necessary for them to have only real generalized eigenvalues.) Set*

$$\eta = \min_{\substack{1 \leq i \leq l \\ l+1 \leq j \leq n}} \rho((\widetilde{\alpha}_i, \widetilde{\beta}_i), (\widetilde{\alpha}_j, \widetilde{\beta}_j)).$$

*If  $\eta > 0$ , then (5.11) holds with  $\|\|\cdot\|\| = \|\cdot\|_F$ .*

A proof of Theorem 5.6 can be given along the lines of our proof of Theorem 5.5, with the help of Theorem 5.4.

**5.3. Dual theorems.** What we studied above is for the perturbation of a (right) eigenspace. As its dual case, we may study the perturbation of a left eigenspace by simply applying the above results to  $A^H - \lambda B^H$ . We omit the details here.

### 6. APPLICATIONS TO DEFINITE PENCILS

Suppose  $A - \lambda B \in \mathbf{D}(n)$  and  $\widetilde{A} - \lambda \widetilde{B} \in \mathbf{D}(n)$  admit the decompositions

$$(6.1a) \quad \begin{cases} X^H A X = \Lambda, \\ X^H B X = \Omega, \end{cases} \quad \text{and} \quad \begin{cases} \widetilde{X}^H \widetilde{A} \widetilde{X} = \widetilde{\Lambda}, \\ \widetilde{X}^H \widetilde{B} \widetilde{X} = \widetilde{\Omega}, \end{cases}$$

where  $\Lambda, \Omega, \dots, \widetilde{\alpha}_1, \widetilde{\beta}_1, \dots$  are of the form (5.2),

$$(6.1b) \quad \begin{cases} X = (X_1, X_2), & \widetilde{X} = (\widetilde{X}_1, \widetilde{X}_2), & X_1, \widetilde{X}_1 \in \mathbb{C}^{n \times l}, \\ X^{-1} = \begin{pmatrix} W_1^H \\ W_2^H \end{pmatrix}, & \widetilde{X}^{-1} = \begin{pmatrix} \widetilde{W}_1^H \\ \widetilde{W}_2^H \end{pmatrix}, & W_1, \widetilde{W}_1 \in \mathbb{C}^{n \times l}. \end{cases}$$

**Theorem 6.1.** *Besides the hypotheses of Theorem 5.1, we assume further that  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}(n)$  and that they admit the decompositions (6.1). Then for any unitarily invariant norm  $||| \cdot |||$ , there holds*

$$(6.2a) \quad ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\|Z\|_2 \|\tilde{Z}\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \cdot \frac{1}{\eta} \cdot \left\| \left( \begin{array}{c} \tilde{W}_{20}^H \\ \tilde{W}_{20}^H \end{array} \right) (P_{Z^H} - P_{\tilde{Z}^H}) \left( \begin{array}{cc} X_{10} & \\ & X_{10} \end{array} \right) \right\|,$$

where

$$(6.3) \quad X_{10} = X_1(X_1^H X_1)^{-1/2}, \quad W_{20} = \tilde{W}_2(\tilde{W}_2^H \tilde{W}_2)^{-1/2}.$$

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

$$(6.2b) \quad ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\|Z\|_2}{\sqrt{c(A, B)c(\tilde{A}, \tilde{B})}} \cdot \frac{1}{\eta} \cdot \left\| \left( \begin{array}{c} \tilde{W}_2^H \\ \tilde{W}_2^H \end{array} \right) (P_{Z^H} - P_{\tilde{Z}^H}) \left( \begin{array}{cc} X_1 & \\ & X_1 \end{array} \right) \right\|.$$

Here,  $\mathcal{X}_1$  and  $\tilde{\mathcal{X}}_1$  are as defined in (5.3).

*Proof.* From Theorem 5.1, Remark 5.1, and

$$(6.4) \quad \|X_1^+\|_2 \leq \|X^{-1}\|_2, \quad \|\tilde{W}_2^+\|_2 \leq \|\tilde{X}\|_2,$$

together with Lemma 4.2, one gets the inequalities (6.2).  $\square$

**Theorem 6.2.** *Besides the hypotheses of Theorem 5.2, we assume further that  $A - \lambda B \in \mathbf{D}(n)$  and  $\tilde{A} - \lambda \tilde{B} \in \mathbf{D}(n)$  and that they admit the decompositions (6.1). Then (6.2) holds for the unitarily invariant norm  $\|\cdot\|_F$ .*

Notice that one of the hypotheses of Theorem 5.2 which allows the pencils considered to have complex generalized eigenvalues is immaterial now, because definite pencils always have only real generalized eigenvalues. Therefore, the only difference between the conditions of Theorem 6.2, as well as Theorems 6.5 and 6.6 below, and those of other theorems in this section is the separation hypothesis imposed on the subsets of generalized eigenvalues of the two pencils.

**Theorem 6.3.** *To the hypotheses of the first part of Theorem 6.1 add these:*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

Then for any unitarily invariant norm  $||| \cdot |||$ , we have

$$(6.5) \quad ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\|Z\|_2}{\sqrt{c(A, B)c(\tilde{A}, \tilde{B})}} \cdot \frac{1}{\eta} \cdot \left\| \tilde{X}_2^H (\tilde{Z} - Z) \left( \begin{array}{cc} X_1 & \\ & X_1 \end{array} \right) \right\|.$$

*Proof.* The assertion (6.5) is a consequence of Theorem 5.3, (6.4), and Lemma 4.2.  $\square$

**Theorem 6.4.** *Under the conditions described in the first part of Theorem 6.1, we have, for any unitarily invariant norm  $||| \cdot |||$ ,*

$$(6.6) \quad ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| \leq \frac{\|Z\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \cdot \frac{1}{\eta} \cdot \left\| \tilde{X}_{20}^H (\tilde{Z} - Z) \left( \begin{array}{cc} X_{10} & \\ & X_{10} \end{array} \right) \right\|,$$

where  $X_{10}$  is defined by (6.3), and

$$(6.7) \quad \tilde{X}_{20} = \tilde{X}_2(\tilde{X}_2^H \tilde{X}_2)^{-1/2}.$$

*Proof.* Without loss of generality, we may assume that  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$ ,  $i, j = 1, 2, \dots, n$ . Since postmultiplying  $X, \tilde{X}$  by nonsingular matrices does not change the singular values of  $\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)$ , we have that (6.5) holds. Now, combining (6.5) with

$$\left\| \left\| \tilde{X}_2^H(\tilde{A} - A, \tilde{B} - B) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\| \right\| \leq \|\tilde{X}\|_2 \|X\|_2 \left\| \left\| \tilde{X}_{20}^H(\tilde{Z} - Z) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\| \right\|$$

and Lemma 4.2 leads to (6.6).  $\square$

Similarly, we have

**Theorem 6.5.** *To the hypotheses of Theorem 6.2 add these:*

$$|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1, \quad i, j = 1, 2, \dots, n.$$

*Then (6.5) holds for the unitarily invariant norm  $\|\cdot\|_F$ .*

**Theorem 6.6.** *Under the conditions of Theorem 6.2, the estimate (6.6) holds for the unitarily invariant norm  $\|\cdot\|_F$ .*

*Remark 6.1.* We omit the statement of the  $\sin 2\theta$  theorem for definite pencils to save space.

Now, we give a simple comparison of our bounds with those in Sun [29]. We have already noted the description (3.5b) and (3.5c), adopted by [29], for the separation of generalized eigenvalues, and have proved that they are equivalent to the description (3.5a). By adopting (3.5a) instead of (3.5b) and/or (3.5c), the main result of Sun [29] may be restated as: *under the conditions of the first part of Theorem 6.1, there holds, for any unitarily invariant norm  $\|\cdot\|$ ,*

$$(6.8) \quad \begin{aligned} \|\|\sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1)\|\| &\leq q \cdot \frac{\|Z\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \\ &\cdot \frac{1}{\eta} \cdot \|\|((\tilde{A} - A)X_{10}, (\tilde{B} - B)X_{10})\|\|_{\text{Sun}}, \end{aligned}$$

where the constant  $q$  satisfies  $1 \leq q \leq \sqrt{2}$ ,  $X_{10}$  is as defined in (6.3), and

$$\|\|((\tilde{A} - A)X_{10}, (\tilde{B} - B)X_{10})\|\|_{\text{Sun}} \stackrel{\text{def}}{=} (\|\|(\tilde{A} - A)X_{10}\|\|^2 + \|\|(\tilde{B} - B)X_{10}\|\|^2)^{1/2}.$$

Here the subscript Sun is used to distinguish the norm from

$$\left\| \left\| \tilde{X}_2^H(\tilde{A} - A, \tilde{B} - B) \begin{pmatrix} X_1 & \\ & X_1 \end{pmatrix} \right\| \right\|.$$

It is difficult to compare (6.8) with (6.2a) in a straightforward way. However, two indirect methods may be employed. One is to weaken (6.2a) just like we did at the end of §4. This leads us to the impression that (6.2a) is less sharp than (6.8). The other is to construct examples. Here is an example from Sun [28]. Let  $A - \lambda B \in \mathbf{D}(n)$ . Consider  $\tilde{A} - \lambda \tilde{B} = (1 + r)(A - \lambda B)$  with  $r > 0$  such that the right-hand side of (6.8) is less than 1 but almost 1. We assume that the spectrum of  $A - \lambda B$  consists of two parts with suitable properties such that (6.8) and (6.2a) hold. It is easy to see that the right-hand side of (6.2a) is

0, whereas the right-hand side of (6.8) is approximately 1. This example shows that sometimes (6.2a) is much sharper than (6.8).

As to (6.8) and (6.6), we may roughly regard them to differ only by a constant factor. To see this, we weaken (6.4) somewhat by writing

$$(6.9) \quad \begin{aligned} ||| \sin \Theta(\mathcal{X}_1, \tilde{\mathcal{X}}_1) ||| &\leq \frac{\|Z\|_2}{c(A, B)c(\tilde{A}, \tilde{B})} \\ &\cdot \frac{1}{\eta} \cdot \left\| \left( \tilde{A} - A, \tilde{B} - B \right) \begin{pmatrix} X_{10} & \\ & X_{10} \end{pmatrix} \right\| \end{aligned}$$

Ignoring the constant  $q \geq 1$ , we note that  $\frac{1}{\sqrt{2}} |||G||| \leq |||G|||_{\text{Sun}} \leq \sqrt{2} |||G|||$ , where  $G \stackrel{\text{def}}{=} ((\tilde{A} - A)X_{10}, (\tilde{B} - B)X_{10})$ . So (6.6) is sharper<sup>1</sup>.

7. CONCLUDING REMARKS

Starting with our theorems in both parts above, we can deduce some known results in the perturbation theory for the standard eigenvalue problem.

7.1. Suppose  $A, \tilde{A} \in \mathbb{C}^{n \times n}$  are both similar to Hermitian matrices, i.e., there exist nonsingular matrices  $X, \tilde{X} \in \mathbb{C}^{n \times n}$  such that

$$(7.1) \quad \tilde{X}^{-1} \tilde{A} \tilde{X} = \text{diag}(\tilde{\alpha}_1, \dots, \tilde{\alpha}_n), \quad X^{-1} A X = \text{diag}(\alpha_1, \dots, \alpha_n),$$

so that  $\lambda(A) = \{\alpha_i, i = 1, 2, \dots, n\}$ ,  $\lambda(\tilde{A}) = \{\tilde{\alpha}_j, j = 1, 2, \dots, n\}$  with all the  $\alpha_i, \tilde{\alpha}_j \in \mathbb{R}$ . For  $t > 0$  sufficiently large, both  $A - \lambda t I$  and  $\tilde{A} - \lambda t I$  belong to  $\mathbf{D}_g(n)$ , and moreover, we have  $\lambda(A, tI) = \{(\alpha_i, t), i = 1, 2, \dots, n\}$  and  $\lambda(\tilde{A}, tI) = \{(\tilde{\alpha}_j, t), j = 1, 2, \dots, n\}$ . Set  $Z_t = (A, tI)$  and  $\tilde{Z}_t = (\tilde{A}, tI)$ . From Theorem 3.1 it follows that there exists a fixed permutation  $\sigma$  and a sequence of increasing positive numbers  $t_1, t_2, \dots$  tending to  $+\infty$  such that  $\rho((\alpha_j, t_j), (\tilde{\alpha}_{\sigma(j)}, t_i)) \leq \kappa(X)\kappa(\tilde{X})d_2(Z_{t_i}, \tilde{Z}_{t_i})$  for  $i = 1, 2, \dots$ , i.e.,

$$\begin{aligned} &\frac{|\alpha_j - \tilde{\alpha}_{\sigma(j)}|t_i}{\sqrt{\alpha_j + t_i^2} \sqrt{\tilde{\alpha}_{\sigma(j)}^2 + t_i^2}} \\ &\leq \kappa(X)\kappa(\tilde{X}) \left\| \begin{pmatrix} A^H \\ t_i I \end{pmatrix} (AA^H + t_i^2 I)^{-1} (A, t_i I) \right. \\ &\quad \left. - \begin{pmatrix} \tilde{A}^H \\ t_i I \end{pmatrix} (\tilde{A}\tilde{A}^H + t_i^2 I)^{-1} (\tilde{A}, t_i I) \right\|_2. \end{aligned}$$

Multiplying the two sides by  $t_i$  and letting  $i$  tend to  $+\infty$  gives

$$(7.2a) \quad \begin{aligned} \max_{1 \leq j \leq n} |\alpha_j - \tilde{\alpha}_{\sigma(j)}| &\leq \kappa(X)\kappa(\tilde{X}) \left\| \begin{pmatrix} 0 & A^H - \tilde{A}^H \\ A - \tilde{A} & 0 \end{pmatrix} \right\|_2 \\ &= \kappa(X)\kappa(\tilde{X}) \|A - \tilde{A}\|_2. \end{aligned}$$

Now if  $A$  is Hermitian, then  $X \in \mathcal{U}_n$ ; thus (7.2a) becomes

$$(7.2b) \quad \max_{1 \leq j \leq n} |\alpha_j - \tilde{\alpha}_{\sigma(j)}| \leq \kappa(\tilde{X}) \|A - \tilde{A}\|_2.$$

<sup>1</sup>It is worth mentioning that Sun's proof of (6.8) can be refined to get (6.9).

This is a theorem of Kahan [9]. If, in addition,  $\tilde{A}$  is also Hermitian, i.e.,  $\tilde{X} \in \mathcal{U}_n$ , then (7.2a) becomes

$$(7.2c) \quad \max_{1 \leq j \leq n} |\alpha_j - \tilde{\alpha}_{\sigma(j)}| \leq \|A - \tilde{A}\|_2.$$

This is nothing but the celebrated Weyl-Lidskii theorem. Inequalities (7.2) can also be obtained by similar arguments starting with Theorem 3.4 instead of Theorem 3.1.

7.2. The Weyl-Lidskii theorem is also a consequence of Theorem 4.1. Let  $A, \tilde{A} \in \mathbb{C}^{n \times n}$  be two Hermitian matrices with  $\lambda(A) = \{\alpha_i, i = 1, 2, \dots, n\}$ ,  $\lambda(\tilde{A}) = \{\tilde{\alpha}_j, j = 1, 2, \dots, n\}$ , with all the  $\alpha_i, \tilde{\alpha}_j \in \mathbb{R}$ . For  $t > 0$  sufficiently large, both  $A - \lambda tI$  and  $\tilde{A} - \lambda tI$  belong to  $\mathbf{D}(n)$ , and moreover, we have  $\lambda(A, tI) = \{(\alpha_i, t), i = 1, 2, \dots, n\}$  and  $\lambda(\tilde{A}, tI) = \{(\tilde{\alpha}_j, t), j = 1, 2, \dots, n\}$ . Let  $Z_t, \tilde{Z}_t$  be as above. By Theorem 4.1 it follows that there exists a fixed permutation  $\sigma$  and a sequence of increasing positive numbers  $t_1, t_2, \dots$  tending to  $+\infty$  such that

$$\rho((\alpha_j, t_i), (\tilde{\alpha}_{\sigma(j)}, t_i)) \leq \frac{\|Z_{t_i}\|_2 \|\tilde{Z}_{t_i}\|_2}{c(A, t_i I) c(\tilde{A}, t_i I)} d_2(Z_{t_i}, \tilde{Z}_{t_i}),$$

for  $i = 1, 2, \dots$ . Note that all the  $\|Z_{t_i}\|_2, \|\tilde{Z}_{t_i}\|_2, c(A, t_i I), c(\tilde{A}, t_i I)$  can grow like  $t_i(1 + O(1/t_i))$ . This shows

$$\frac{\|Z_{t_i}\|_2 \|\tilde{Z}_{t_i}\|_2}{c(A, t_i I) c(\tilde{A}, t_i I)} = 1 + O\left(\frac{1}{t_i}\right).$$

Now repeating the arguments similar to those in §7.1, we finally get (7.2c).

7.3. Sun [30] proved that if  $A, \tilde{A} \in \mathbb{C}^{n \times n}$  are both similar to normal matrices, i.e., formally we have (7.1) but with  $\alpha_i, \tilde{\alpha}_j \in \mathbb{C}$ , then there is a permutation  $\tau$  of  $\{1, 2, \dots, n\}$  such that

$$\sqrt{\sum_{j=1}^n |\alpha_j - \tilde{\alpha}_{\tau(j)}|^2} \leq \kappa(X)\kappa(\tilde{X})\|A - \tilde{A}\|_F.$$

This is, obviously, a generalization of (1.1) to the case of the Frobenius norm. Sun [27] also proved that if  $A - \lambda B$  and  $\tilde{A} - \lambda \tilde{B}$  are two normal pencils with  $\lambda(A, B) = \{(\alpha_i, \beta_i), i = 1, 2, \dots, n\}$  and  $\lambda(\tilde{A}, \tilde{B}) = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = 1, 2, \dots, n\}$  (it is not necessary that all  $(\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)$  be real), then

$$(7.3) \quad \sqrt{\sum_{j=1}^n [\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))]^2} \leq d_F(Z, \tilde{Z}).$$

Therefore, it may be of interest to establish an inequality like (7.3) for diagonalizable pencils. Our inequality (3.3) is developed for this purpose. However, the right side of (3.3) does not have the desired form. We suspect that (3.3) might be true with its right-hand replaced by  $\kappa(X)\kappa(\tilde{X})d_F(Z, \tilde{Z})$ , but we are



unable to prove it. Liu [19] attempted to investigate this problem, but was not successful.

7.4. An interesting result on the Hadamard product of matrices can be deduced from Proposition 3.4. The conclusion is as follows.

**Proposition.** *Let  $\alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \in \mathbb{R}$ ,  $|\alpha_i|^2 + |\beta_i|^2 = |\tilde{\alpha}_j|^2 + |\tilde{\beta}_j|^2 = 1$ , for all  $i, j = 1, 2, \dots, n$ ; then*

$$\min_{U \in \mathbb{Z}_n} \|U \circ G\|_2 = \min_{\sigma} \max_{1 \leq j \leq n} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})),$$

where  $G \stackrel{\text{def}}{=} (g_{ij}) \stackrel{\text{def}}{=} (\tilde{\alpha}_i \beta_j - \tilde{\beta}_i \alpha_j) \in \mathbb{C}^{n \times n}$ , and “ $\circ$ ” denotes the Hadamard product, i.e.,  $U \circ G = (u_{ij} g_{ij})$ .

In fact,  $U \circ G = \tilde{\Lambda} U \Omega - \tilde{\Omega} U \Lambda$ , where  $\Lambda, \Omega, \tilde{\Lambda}$ , and  $\tilde{\Omega}$  are of the form (3.1b). It is easy to see from Proposition 3.4 that

$$\min_{U \in \mathbb{Z}_n} \|U \circ G\|_2 \geq \min_{\sigma} \max_{1 \leq j \leq n} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_{\sigma(j)}, \tilde{\beta}_{\sigma(j)})).$$

On the other hand, the same obviously holds with “ $\geq$ ” reversed, since permutation matrices are also unitary matrices.

7.5. Also, by a special limiting procedure like that in §§7.1 and 7.2, some known perturbation results for eigenspaces, such as the Davis-Kahan  $\sin \theta$  and  $\sin 2\theta$  theorems on the standard eigenvalue problem, can be deduced from our results in Part II. We omit the details here.

7.6. The most essential hypothesis throughout this paper is that matrix pencils considered are, generally, required to have only real generalized eigenvalues. However, it is worth mentioning that in a few theorems in Part II involving the Frobenius norm, this hypothesis can be removed. The reader is referred to Sun [29] for details.

7.7. Readers may wonder about the ad hoc example  $A - \lambda B$  (unperturbed) and  $(1 + r)A - \lambda(1 + r)B$  (perturbed), used in §§4 and 6 for our comparisons. Our choosing this example is no accident. In fact, according to Sun’s useful suggestion that all  $n \times n$  regular matrix pencils be embedded into the Grassmann manifold of all  $n \times 2n$  matrices having full row rank,  $A - \lambda B$  and  $(1 + r)A - \lambda(1 + r)B$  represent the same point on the manifold. Therefore, all bounds related to metrics on it (or  $P_{Z^H} - P_{\tilde{Z}^H}$  in Part II) should yield the best estimation, i.e., 0, for this kind of perturbations.

7.8. This paper deals with generalizations of well-known perturbation results for eigenvalues and eigenspaces of the standard eigenvalue problem. In Li [16], some known results for standard singular value variations are successfully generalized to perturbations of the generalized singular value problem.

#### ACKNOWLEDGMENT

I would like to thank Professor G. W. Stewart for many helpful remarks and for suggesting to combine two of my earlier manuscripts on the same topic into the present paper.

## BIBLIOGRAPHY

1. R. Bhatia and Ch. Davis, *A bound for the spectral variation of a unitary operator*, Linear and Multilinear Algebra **15** (1984), 71–76.
2. R. Bhatia, Ch. Davis, and A. McIntosh, *Perturbation of spectral subspaces and the solution of linear equations*, Linear Algebra Appl. **52/53** (1983), 45–67.
3. C. R. Crawford, *A stable generalized eigenvalue problem*, SIAM J. Numer. Anal. **8** (1976), 854–860.
4. Ch. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal. **7** (1970), 1–46.
5. L. Elsner and P. Lancaster, *The spectral variation of pencils of matrices*, J. Comput. Math. **3** (1985), 262–274.
6. L. Elsner and Ji-guang Sun, *Perturbation theorems for the generalized eigenvalue problem*, Linear Algebra Appl. **48** (1982), 341–357.
7. I. C. Gohberg and M. G. Krein, *Introduction to the theory of linear nonselfadjoint operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, RI, 1969.
8. A. J. Hoffman and H. W. Wielandt, *The variation of the spectrum of a normal matrix*, Duke Math. J. **20** (1953), 37–39.
9. W. M. Kahan, *Spectra of nearly Hermitian matrices*, Proc. Amer. Math. Soc. **48** (1975), 11–17.
10. T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, New York, 1966.
11. Ren-Cang Li, *On perturbation theorems for the generalized eigenvalues of regular matrix pencils*, Math. Numer. Sinica **11** (1989), 10–19 (Chinese); English transl., Chinese J. Numer. Math. Appl. **11**:2 (1989), 24–35.
12. —, *Perturbation bounds for generalized eigenvalues. I, II*, Math. Numer. Sinica **11**:2 (1989), 196–204, **11**:3 (1989), 239–247 (Chinese); English transl., Chinese J. Numer. Math. Appl. **11**:3 (1989), 34–43; **11**:4 (1989), 1–9.
13. —, *A converse to the Bauer-Fike type theorem*, Linear Algebra Appl. **109** (1988), 167–178.
14. —, *On the variation of the spectra of matrix pencils*, Linear Algebra Appl. **139** (1990), 147–164.
15. —, M.S. dissertation, Computing Center, Academia Sinica, 1987.
16. —, *Bounds on perturbations of generalized singular values and of associated subspaces*, SIAM J. Matrix Anal. Appl. **14** (1993), 195–234.
17. —, *Norms of certain matrices with applications to variations of the spectra of matrices and matrix pencils*, Linear Algebra Appl. **182** (1993), 199–234.
18. —, *A perturbation bound for definite pencils*, Linear Algebra Appl. **179** (1993), 191–202.
19. Xing-guo Liu, *The perturbation bounds of latent values of a class of matrix polynomials*, Math. Numer. Sinica **11**:1 (1989), 20–28. (Chinese)
20. L. Mirsky, *Symmetric gauge functions and unitarily invariant norms*, Quart. J. Math. Oxford **11** (1960), 50–59.
21. G. W. Stewart, *On the sensitivity of the eigenvalue problem  $Ax = \lambda Bx$* , SIAM J. Numer. Anal. **9** (1972), 669–686.
22. —, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Rev. **15** (1973), 727–764.
23. —, *On the perturbation of pseudo-inverses, projections and linear least squares problem*, SIAM Rev. **19** (1977), 634–662.
24. —, *Perturbation bounds for the definite generalized eigenvalue problem*, Linear Algebra Appl. **23** (1979), 69–83.
25. G. W. Stewart and Ji-guang Sun, *Matrix perturbation analysis*, Academic Press, New York, 1990.

26. Ji-guang Sun, *Invariant subspaces and generalized invariant subspaces. (I), (II)*, Math. Numer. Sinica **2** (1980), 1–13, 113–123. (Chinese)
27. ———, *The perturbation bounds of generalized eigenvalues of a class of matrix-pairs*, Math. Numer. Sinica **4** (1982), 23–29. (Chinese)
28. ———, *A note on Stewart's theorem for definite matrix pairs*, Linear Algebra Appl. **48** (1982), 331–339.
29. ———, *The perturbation bounds for eigenspaces of a definite matrix-pair*, Numer. Math. **41** (1983), 321–343.
30. ———, *On the perturbation of the eigenvalues of a normal matrix*, Math. Numer. Sinica **6** (1984), 334–336. (Chinese)
31. F. Uhlig, *A recurring theorem about pairs of quadratic forms and extensions: A survey*, Linear Algebra Appl. **25** (1979), 219–237.
32. H. Wielandt, *An extremum property of sums of eigenvalues*, Proc. Amer. Math. Soc. **6** (1955), 106–110.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT BERKELEY, BERKELEY, CALIFORNIA 94720

*E-mail address:* li@math.berkeley.edu