Numer. Math. 41, 321-343 (1983)

Numerische Mathematik

The Perturbation Bounds for Eigenspaces of a Definite Matrix-Pair*

Ji-guang Sun[1]

Computing Center, Academia Sinica, Peking, China

Summary. Let A and B be Hermitian matrices. The matrix-pair (A, B) is called "definite pair" and the corresponding eigenvalue problem $Ax=\lambda Bx$ is definite if $c(A,B)\equiv \min_{\|x\|=1} \{|x^H(A+iB)x|\}>0$. The perturbation bounds for eigenspaces of a definite pair on every unitary-invariant matrix norm were obtained by imposing additional restrictions on the location of the generalized eigenvalues. Thus it gives a positive answer for an open question proposed by Stewart [7]. The famous Davis-Kahan $\sin\theta$ theorems and $\sin 2\theta$ Theorem [2] can also be deduced from the present results.

Subject Classifications. AMS(MOS): 65F15; CR: 5.14.

Introduction

"Definite pair" is a class of important matrix-pairs (see [10]). Some results on the stability analysis of the definite generalized eigenvalue problem were obtained by Crawford [1], Stewart [7] and the author [9]. Stewart [7] has obtained perturbation bounds in the Frobenius norm for the eigenspaces of a definite pair under certain conditions and has pointed out: "For the Hermitian eigenvalue problem, Davis and Kahan have been able to obtain bounds on the spectral norm by imposing additional restrictions on the location of the eigenvalues. Whether such bounds can be obtained for the definite generalized eigenvalue problem is an open question."

The present work gives a positive answer for this open question. Perturbation bounds not only on the spectral norm but also on every unitary-invariant matrix norm are obtained (see Theorem 2.1, i.e. the $\sin\theta$ theorem for definite pairs). Moreover, under weaker conditions a perturbation bound for

* Dedicated to Gatlinburg VIII

[1] Visiting the University of Bielefeld, Fakultät für Mathematik, Universitätsstraße, D-4800 Bielefeld (Fed. Rep.)

the eigenspaces is developed (see Theorem 3.1). Besides, a part of the Davis-Kahan $\sin 2\theta$ theorem is generalized to definite pairs.

The abovementioned definitions and some basic results are given in §1. The $\sin\theta$ theorem, the generalized $\sin\theta$ theorem, the $\sin 2\theta$ theorem and the strengthened $\sin 2\theta$ theorem for definite pairs are proved in §2–§5, respectively. The last section points out that from our results one can deduce the famous Davis-Kahan $\sin\theta$ theorems and $\sin 2\theta$ theorem [2] by a limiting procedure.

Notation. Upper case letters are used for matrices and lower case Greek letters for scalars. The symbol $\mathbb{C}^{m\times n}$ denotes the set of complex $m\times n$ matrices. $\bar{A}$ and A^T stand for conjugate and transpose of A, respectively; $A^H=\bar{A}^T$. $I^{(n)}$ is the $n\times n$ identity matrix, and 0 is the null matrix. For a Hermitian matrix H with eigenvalues $\{\alpha_i\}$, $H>0$ ($H\geqq 0$) denotes that H is positive definite (semi-positive definite) and $\lambda_{\min}(H)=\min_i\{\alpha_i\}$. Let $\| \ \|$ denote the usual Euclidean vector norm, $\| \ \|_2$ the spectral norm and $\| \ \|_F$ the Frobenius matrix norm. The column space of A is denoted by $\mathscr{R}(A)$. $\mathscr{S}_1\cap\mathscr{S}_2$ and $\mathscr{S}_1\cup\mathscr{S}_2$ stand for the intersection and union of two sets $\mathscr{S}_1$ and $\mathscr{S}_2$, respectively, and $\emptyset$ for the empty set. The chordal distance between the points (α,β) and $(\tilde{\alpha},\tilde{\beta})$ in the complex projective plane $\mathscr{P}(1,1)$ is

$$\rho((\alpha,\beta),(\tilde{\alpha},\tilde{\beta}))=|\alpha\tilde{\beta}-\beta\tilde{\alpha}|/\sqrt{(|\alpha|^2+|\beta|^2)(|\tilde{\alpha}|^2+|\tilde{\beta}|^2)}.$$

§1. Preliminaries

Definition 1.1. Let Hermitian matrices $A, B\in\mathbb{C}^{n\times n}$. (A,B) is a "definite pair", if

$$c(A,B)\equiv\min_{\|x\|=1}\{|x^H(A+iB)x|\}>0. \tag{1.1}$$

$\mathbb{D}(n)$ denotes the set of all definite pairs of $n\times n$ matrices. The following Theorem 1.1 and Theorem 1.2 are well known (see [7]).

Theorem 1.1. *Let $(A,B)\in\mathbb{D}(n)$. Then there is a nonsingular matrix $Q\in\mathbb{C}^{n\times n}$ such that*

$$Q^HAQ=\Lambda,\quad Q^HBQ=\Omega,\quad \Lambda=\operatorname{diag}(\alpha_i),\quad \Omega=\operatorname{diag}(\beta_i). \tag{1.2}$$

Theorem 1.2. *Let $(A,B)\in\mathbb{D}(n)$ and*

$$A_\varphi=\cos\varphi A-\sin\varphi B,\qquad B_\varphi=\sin\varphi A+\cos\varphi B, \tag{1.3}$$

where φ is a real number. Then there is a $\varphi\in[0,2\pi]$ such that $B_\varphi>0$ and $c(A,B)=\lambda_{\min}(B_\varphi)$.

Definition 1.2 [3]. Let $A,B\in\mathbb{C}^{n\times n}$. A vector $x\in\mathbb{C}^n$, $x\neq 0$ is an eigenvector of (A,B) corresponding to the generalized eigenvalues (α,β), if

$$(\alpha,\beta)\neq(0,0)\quad\text{and}\quad \beta Ax=\alpha Bx.$$

If $\beta\neq 0$, then $\lambda=\frac{\alpha}{\beta}$ is a finite generalized eigenvalue of (A,B), and $\frac{\alpha}{\beta}$ is called the non-homogeneous coordinate of the point (α,β) in the complex projective plane.

$\lambda(A, B)$ denotes the set of all generalized eigenvalues of (A, B).

Eigenspaces of a definite pair, as a generalization of the eigenvector concept, have been defined by Stewart as follows.

Definition 1.3 [7]. Let $(A, B) \in \mathbb{D}(n)$. A subspace $\mathscr{X}$ is an eigenspace of (A, B) if

$$\dim(A\mathscr{X}+B\mathscr{X}) \leqq \dim(\mathscr{X}).$$

According to [7] (see [7], 79–80), we can adopt the following decompositions in order to study perturbation bounds of any ℓ-dimensional eigenspace for $(A, B) \in \mathbb{D}(n)$:

$$Z^H A Z = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad Z^H B Z = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \tag{1.4}$$

where $A_1, B_1 \in \mathbb{C}^{\ell \times \ell}$, and

$$Z = (Z_1, Z_2),\ Z_1^H Z_1 = I^{(\ell)},\ Z_2^H Z_2 = I^{(n-\ell)},\ 0 < \ell < n. \tag{1.5}$$

Obviously, $\mathscr{R}(Z_1)$ is an ℓ-dimensional eigenspace for (A, B), $(A_1, B_1) \in \mathbb{D}(\ell)$ and $(A_2, B_2) \in \mathbb{D}(n-\ell)$. We shall use the same notation for perturbed pairs $(\tilde{A}, \tilde{B}) \in \mathbb{D}(n)$, expect that all quantities will be marked with tildes. Let

$$U = (Z_1, W_2), \quad \tilde{U} = (\tilde{Z}_1, \tilde{W}_2) \tag{1.6}$$

be $n \times n$ unitary matrices, $Z_1, \tilde{Z}_1 \in \mathbb{C}^{n \times \ell}$, and let

$$\Theta_1 \equiv \arccos (Z_1^H \tilde{Z}_1 \tilde{Z}_1^H Z_1)^{\frac{1}{2}} \geqq 0. \tag{1.7}$$

Now we discuss the relationship between the matrix $Z_1^H \tilde{W}_2$ and the rotation of $\mathscr{R}(Z_1)$ to $\mathscr{R}(\tilde{Z}_1)$ and explain the geometric significance of the quantity $\|\sin \Theta_1\|$ for every unitary-invariant matrix norm.

Let

$$\begin{aligned} P_0 &= Z_1 Z_1^H, \quad Q_0 = I - P_0 = I - Z_1 Z_1^H = W_2 W_2^H, \\ \tilde{P}_0 &= \tilde{Z}_1 \tilde{Z}_1^H, \quad \tilde{Q}_0 = I - \tilde{P}_0 = I - \tilde{Z}_1 \tilde{Z}_1^H = \tilde{W}_2 \tilde{W}_2^H. \end{aligned} \tag{1.8}$$

Evidently, if there exists a unitary matrix V, such that

$$V P_0 V^H = \tilde{P}_0, \tag{1.9}$$

then from the relations

$$\mathscr{R}(\tilde{Z}_1) = \tilde{P}_0 \mathbb{C}^n = V P_0 V^H \mathbb{C}^n = V P_0 \mathbb{C}^n = V \mathscr{R}(Z_1)$$

we know that V is indeed a rotation of $\mathscr{R}(Z_1)$ to $\mathscr{R}(\tilde{Z}_1)$. Now we seek the representation of V.

First, we write

$$V = \tilde{U}(\tilde{U}^H V U) U^H = (\tilde{Z}_1, \tilde{W}_2)\, T \begin{pmatrix} Z_1^H \\ W_2^H \end{pmatrix}, \tag{1.10}$$

where

$$T = \begin{pmatrix} \tilde{Z}_1^H V Z_1 & \tilde{Z}_1^H V W_2 \\ \tilde{W}_2^H V Z_1 & \tilde{W}_2^H V W_2 \end{pmatrix} = \begin{pmatrix} T_1 & R_1 \\ R_2 & T_2 \end{pmatrix}$$

is a unitary matrix. From (1.8) and (1.9) we obtain $VP_0=\tilde{P}_0V$, $VQ_0=\tilde{Q}_0V$ and thus $T_1^HT_1=I^{(\ell)}$, $T_2^HT_2=I^{(n-\ell)}$, consequently $R_1=0$, $R_2=0$. Substituting such T into (1.10), we obtain

$$V=(\tilde{Z}_1,\tilde{W}_2)\begin{pmatrix}T_1 & 0\\ 0 & T_2\end{pmatrix}\begin{pmatrix}Z_1^H\\ W_2^H\end{pmatrix}=(Z_1,W_2)\begin{pmatrix}Z_1^H\\ W_2^H\end{pmatrix}(\tilde{Z}_1,\tilde{W}_2)\begin{pmatrix}T_1 & 0\\ 0 & T_2\end{pmatrix}\begin{pmatrix}Z_1^H\\ W_2^H\end{pmatrix}$$

$$=U\begin{pmatrix}C_1 & S_1\\ S_2 & C_2\end{pmatrix}U^H,\qquad C_1=Z_1^H\tilde{Z}_1T_1,\ S_1=Z_1^H\tilde{W}_2T_2.$$

Suppose that the singular values of $Z_1^H\tilde{Z}_1$ are $\gamma_1,\dots,\gamma_\ell$, then $\theta_i=\cos^{-1}\gamma_i$ ($i=1,\dots,\ell$) are exactly the angles between the corresponding base vectors of $\mathscr{R}(Z_1)$ and $\mathscr{R}(\tilde{Z}_1)$ by a suitable selecting of their base vectors. Hence from (1.7) it follows that

$$\Theta_1\equiv\arccos(C_1C_1^H)^{\frac{1}{2}}\geqq 0,$$

and thus

$$S_1S_1^H=I-C_1C_1^H=I-\cos^2\Theta_1=\sin^2\Theta_1.$$

Then we have

$$\|\sin\Theta_1\|=\|S_1\|=\|Z_1^H\tilde{W}_2\| \tag{1.11}$$

for any unitary-invariant matrix norm $\|\ \ \|$. Therefore $\|\sin\Theta_1\|$ is a measure of the difference between the subspaces $\mathscr{R}(Z_1)$ and $\mathscr{R}(\tilde{Z}_1)$ (Ref. [2], 9–10; [5], 733–736).

In addition $\|\sin 2\Theta_1\|$ is also a measure of the difference between the subspaces $\mathscr{R}(Z_1)$ and $\mathscr{R}(\tilde{Z}_1)$ (Ref. [2], 8–11).

In [8] and [3] the author used the generalized chordal metrics

$$d_F(Z_1,\tilde{Z}_1)=[\operatorname{tr}(I-Z_1^H\tilde{Z}_1\tilde{Z}_1^HZ_1)]^{\frac{1}{2}}$$

and

$$d_2(Z_1,\tilde{Z}_1)=[\|I-Z_1^H\tilde{Z}_1\tilde{Z}_1^HZ_1\|_2]^{\frac{1}{2}}$$

to characterize the distance between the subspaces $\mathscr{X}=\mathscr{R}(Z_1)$ and $\tilde{\mathscr{X}}=\mathscr{R}(\tilde{Z}_1)$; here we assume $Z_1^HZ_1=\tilde{Z}_1^H\tilde{Z}_1=I$ without loss of generality. Obviously, the relations

$$d_F(Z_1,\tilde{Z}_1)=\|\sin\Theta_1\|_F,\qquad d_2(Z_1,\tilde{Z}_1)=\|\sin\Theta_1\|_2$$

are valid.

Hence, according to [3], if $P_{\mathscr{X}}$ and $P_{\tilde{\mathscr{X}}}$ are the respective projectors onto $\mathscr{X}$ and $\tilde{\mathscr{X}}$, then

$$\|\sin\Theta_1\|_F=\frac{1}{\sqrt{2}}\|P_{\mathscr{X}}-P_{\tilde{\mathscr{X}}}\|_F,\qquad \|\sin\Theta_1\|_2=\|P_{\mathscr{X}}-P_{\tilde{\mathscr{X}}}\|_2. \tag{1.12}$$

In § 4 and § 5 we shall use the following lemmas.

Lemma 1.1. *Suppose that* $Z=(Z_1,Z_2)\in\mathbb{C}^{n\times n}$, $Z_1^HZ_1=I^{(\ell)}$ *and* $Z_2^HZ_2=I^{(n-\ell)}$, $0<\ell<n$. *Then* Z *is non-singular iff* $\|Z_1^HZ_2\|_2<1$.

Proof. There exist matrices U, Σ and V such that $Z_1^HZ_2=U\Sigma V^H$. $U\Sigma V^H$ is the singular value decomposition (SVD) of $Z_1^HZ_2$, where $U\in\mathbb{C}^{\ell\times\ell}$ and $V\in\mathbb{C}^{(n-\ell)\times(n-\ell)}$ are unitary matrices, and $\Sigma=\operatorname{diag}(\sigma_1,\sigma_2,\dots,)\in\mathbb{C}^{\ell\times(n-\ell)}$ in which

$\sigma_1 \geqq \sigma_2 \geqq \ldots \geqq 0$. And thus we have $|\det Z|^2 = \prod_i (1-\sigma_i^2)$ and $\|Z_1^H Z_2\|_2 = \sigma_1 \leqq 1$. Therefore $\det Z \neq 0$ iff $\|Z_1^H Z_2\|_2 < 1$. □

Lemma 1.2. *Suppose* $A = \begin{pmatrix} I & B \\ B^H & I \end{pmatrix}$ *and* $\|B\|_2 < 1$. *Then*

$$\|A\|_2 = 1 + \|B\|_2 \tag{1.13}$$

and

$$\|A^{-1}\|_2 = \frac{1}{1-\|B\|_2}. \tag{1.14}$$

Proof. Utilizing the SVD of B: $B = U\Sigma V^H$, U and V are unitary matrices, $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots)$ with $\sigma_1 \geqq \sigma_2 \geqq \ldots \geqq 0$, we get (1.13) and (1.14) at once. □

Lemma 1.3. *Suppose* $A_1 \in \mathbb{C}^{m \times n_1}$, $A_2 \in \mathbb{C}^{m \times n_2}$, $m \leqq n_1$ *and* $A_1 A_1^H + A_2 A_2^H = I$, *then there exist unitary matrices* U, V_1 *and* V_2: $U \in \mathbb{C}^{m \times m}$, $V_1 \in \mathbb{C}^{n_1 \times n_1}$ *and* $V_2 \in \mathbb{C}^{n_2 \times n_2}$, *such that*

$$A_1 = U\Sigma_1 V_1^H, \quad A_2 = U\Sigma_2 V_2^H n, \tag{1.15}$$

where $\Sigma_1 = \mathrm{diag}(\alpha_1, \alpha_2, \ldots)$ *and* $\Sigma_2 = \mathrm{diag}(\beta_1, \beta_2, \ldots)$ *satisfying*

$$\alpha_i \geqq 0, \quad \beta_i \geqq 0, \quad \alpha_i^2 + \beta_i^2 = 1, \quad i = 1, 2, \ldots.$$

Proof. Using Theorem 2 in [11], there are a non-singular matrix U and unitary matrices V_1 and V_2 such that (1.15) holds, where $\Sigma_1 = \mathrm{diag}(\alpha_1, \alpha_2, \ldots)$ and $\Sigma_2 = \mathrm{diag}(\beta_1, \beta_2, \ldots)$ satisfying $\alpha_i, \beta_i \geqq 0$, $i = 1, 2, \ldots$. From $A_1 A_1^H + B_1 B_1^H = I$ we deduce $\alpha_i^2 + \beta_i^2 = 1$ for $i = 1, 2, \ldots$, and $UU^H = I$ by a suitable selection of those $\{\alpha_i\}$ and $\{\beta_i\}$. □

§ 2. The sin θ Theorem

The following theorem is the main result in this paper.

Theorem 2.1. *(The* $\sin\theta$ *theorem for definite pairs). Let* (A, B), $(\tilde{A}, \tilde{B}) = (A+E, B+F) \in \mathbb{D}(n)$ *with the decompositions given in* (1.4) *and* (1.5). *Assume that there are* $\alpha \geqq 0$ *and* $\delta > 0$ *satisfying* $\alpha + \delta \leqq 1$, *and a real number* γ, *such that*

$$\rho((\gamma, 1), (\alpha_i, \beta_i)) \leqq \alpha, \quad \forall (\alpha_i, \beta_i) \in \lambda(A_1, B_1) \tag{2.1}$$

and

$$\rho((\gamma, 1), (\tilde{\alpha}_j, \tilde{\beta}_j)) \geqq \alpha + \delta, \quad \forall (\tilde{\alpha}_j, \tilde{\beta}_j) \in \lambda(\tilde{A}_2, \tilde{B}_2) \tag{2.2}$$

(or vice-versa). Then for every unitary-invariant matrix norm,

$$\|\sin \Theta_1\| \leqq \frac{p(\alpha, \delta; \gamma)\, \|(A, B)\|_2}{c(A, B)\, c(\tilde{A}, \tilde{B})} \cdot \frac{\|(EZ_1, FZ_1)\|}{\delta}, \tag{2.3}$$

where Θ_1 *is defined by* (1-7),

$$\begin{gathered} p(\alpha, \delta; \gamma) = \frac{q(\gamma)\,[(\alpha+\delta)\sqrt{1-\alpha^2} + \alpha\sqrt{1-(\alpha+\delta)^2}\,]}{2\alpha+\delta}, \\ q(\gamma) = \sqrt{2} \ \text{for}\ \gamma \neq 0 \ \text{and}\ q(0) = 1 \end{gathered} \tag{2.4}$$

and

$$\|(A,B)\|_2=\sqrt{\|A^2+B^2\|_2}, \qquad \|(EZ_1,FZ_1)\|=\sqrt{\|EZ_1\|^2+\|FZ_1\|^2}. \tag{2.5}$$

Proof. This theorem is proved by the following steps 2.1–2.4:

2.1. *The Perturbation Equations*

First, set

$$W'=Z^{-H}=(W_1',W_2'), \qquad \tilde W'=\tilde Z^{-H}=(\tilde W_1',\tilde W_2'), \qquad W_1' \text{ and } \tilde W_1'\in\mathbb{C}^{n\times\ell} \tag{2.1.1}$$

in which Z and $\tilde Z$ were given in (1.4) and (1.5). Moreover, set

$$W=W'\begin{pmatrix}(W_1'^H W_1')^{-\frac12} & 0\\ 0 & (W_2'^H W_2')^{-\frac12}\end{pmatrix}, \qquad \tilde W=\tilde W'\begin{pmatrix}(\tilde W_1'^H \tilde W_1')^{-\frac12} & 0\\ 0 & (\tilde W_2'^H \tilde W_2')^{-\frac12}\end{pmatrix} \tag{2.1.2}$$

and

$$(W_j'^H W_j')^{\frac12}(A_j,B_j)=(A_j',B_j'), \qquad (\tilde W_j'^H \tilde W_j')^{\frac12}(\tilde A_j,\tilde B_j)=(\tilde A_j',\tilde B_j'), \qquad j=1,2. \tag{2.1.3}$$

Then (1.4) can be written as

$$(AZ,BZ)=W\left(\begin{pmatrix}A_1' & 0\\ 0 & A_2'\end{pmatrix},\begin{pmatrix}B_1' & 0\\ 0 & B_2'\end{pmatrix}\right), \qquad (\tilde A\tilde Z,\tilde B\tilde Z)=\tilde W\left(\begin{pmatrix}\tilde A_1' & 0\\ 0 & \tilde A_2'\end{pmatrix},\begin{pmatrix}\tilde B_1' & 0\\ 0 & \tilde B_2'\end{pmatrix}\right), \tag{2.1.4}$$

where $Z=(Z_1,Z_2)$, $W=(W_1,W_2)$, $\tilde Z=(\tilde Z_1,\tilde Z_2)$ and $\tilde W=(\tilde W_1,\tilde W_2)$ satisfy

$$\begin{aligned} &Z_1^H Z_1=\tilde Z_1^H\tilde Z_1=W_1^H W_1=\tilde W_1^H\tilde W_1=I^{(\ell)}\\ &Z_2^H Z_2=\tilde Z_2^H\tilde Z_2=W_2^H W_2=\tilde W_2^H\tilde W_2=I^{(n-\ell)},\\ &Z_1^H W_2=\tilde Z_1^H\tilde W_2=0, \qquad Z_2^H W_1=\tilde Z_2^H\tilde W_1=0.\end{aligned} \tag{2.1.5}$$

From (2.1.4) it follows that $AZ_1=W_1A_1'$, $BZ_1=W_1B_1'$, so that we define the residuals

$$R_A=\tilde AZ_1-W_1A_1', \qquad R_B=\tilde BZ_1-W_1B_1'. \tag{2.1.6}$$

Obviously, the relations

$$R_A=EZ_1, \qquad R_B=FZ_1 \tag{2.1.7}$$

are valid.

Utilizing (2.1.4) and (2.1.1), we know

$$\tilde A=\tilde W_1\tilde A_1'\tilde W_1'^H+\tilde W_2\tilde A'\tilde W_2'^H, \qquad \tilde B=\tilde W_1\tilde B_1'\tilde W_1'^H+\tilde W_2\tilde B_2'\tilde W_2'^H.$$

Substituting these relations into (2.1.6) and taking the transpose conjugate, we get

$$\begin{aligned} R_A^H&=Z_1^H(\tilde W_1'\tilde A_1'^H\tilde W_1^H+\tilde W_2'\tilde A_2'^H\tilde W_2^H)-A_1'^H W_1^H\\ R_B^H&=Z_1^H(\tilde W_1'\tilde B_1'^H\tilde W_1^H+\tilde W_2'\tilde B_2'^H\tilde W_2^H)-B_1'^H W_1^H.\end{aligned} \tag{2.1.8}$$

Moreover, utilizing (2.1.5), (2.1.2), (2.1.1) and (2.1.8), we obtain

$$\begin{aligned} R_A^H \tilde{Z}_2 &= Z_1^H \tilde{W}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} \tilde{A}_2'^H (\tilde{W}_2'^H \tilde{W}_2')^{-\frac{1}{2}} - A_1'^H W_1^H \tilde{Z}_2 \\ R_B^H \tilde{Z}_2 &= Z_1^H \tilde{W}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} \tilde{B}_2'^H (\tilde{W}_2'^H \tilde{W}_2')^{-\frac{1}{2}} - B_1'^H W_1^H \tilde{Z}_2. \end{aligned} \tag{2.1.9}$$

Let

$$\begin{aligned} \hat{A}_2 &= (\tilde{W}_2'^H \tilde{W}_2')^{-\frac{1}{2}} \tilde{A}_2' (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} = \tilde{A}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} \\ \hat{B}_2 &= (\tilde{W}_2'^H \tilde{W}_2')^{-\frac{1}{2}} \tilde{B}_2' (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} = \tilde{B}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}}. \end{aligned} \tag{2.1.10}$$

Then (2.1.9) becomes

$$\begin{aligned} R_A^H \tilde{Z}_2 &= Z_1^H \tilde{W}_2 \hat{A}_2^H - A_1'^H W_1^H \tilde{Z}_2 \\ R_B^H \tilde{Z}_2 &= Z_1^H \tilde{W}_2 \hat{B}_2^H - B_1'^H W_1^H \tilde{Z}_2. \end{aligned} \tag{2.1.11}$$

Let

$$W_1^H \tilde{Z}_2 = X, \qquad Z_1^H \tilde{W}_2 = Y, \qquad -R_A^H \tilde{Z}_2 = C, \qquad -R_B^H \tilde{Z}_2 = D, \tag{2.1.12}$$

then the Eqs. (2.1.11) can be written as

$$A_1'^H X - Y \hat{A}_2^H = C, \qquad B_1'^H X - Y \hat{B}_2^H = D, \tag{2.1.13}$$

where $A_1', B_1' \in \mathbb{C}^{\ell \times \ell}$, $\hat{A}_2, \hat{B}_2 \in \mathbb{C}^{(n-\ell)\times(n-\ell)}$, and $X, Y \in \mathbb{C}^{\ell \times (n-\ell)}$ are the unknowns.

From (1.11) and (2.1.12) it follows that

$$\|Y\| = \|\sin \Theta_1\|$$

for every unitary-invariant matrix norm. Hence for the proof of inequality (2.3) it is sufficient to establish (2.3) for $\|Y\|$.

2.2. The Simplification of the Perturbation Equations

Now we give suitable representations for A_1', B_1', $\hat{A}_2$ and $\hat{B}_2$ in the Eqs. (2.1.13) and transform the equations into a simpler form.

Since (A_i, B_i) $(i=1,2)$ in (1.4) are definite pairs, there are non-singular matrices P_i and real diagonal matrices Λ_i and Ω_i $(i=1,2)$ such that

$$\begin{aligned} &\begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} = \operatorname{diag}(\alpha_k), \qquad \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix} = \operatorname{diag}(\beta_k), \\ &(A_i, B_i) = P_i^H(\Lambda_i P_i, \Omega_i P_i), \qquad \Lambda_i^2 + \Omega_i^2 = I, \qquad i=1,2. \end{aligned} \tag{2.2.1}$$

Let

$$Q = Z \begin{pmatrix} P_1^{-1} & 0 \\ 0 & P_2^{-1} \end{pmatrix} = (Q_1, Q_2), \qquad Q_1 \in \mathbb{C}^{n \times \ell}; \tag{2.2.2}$$

evidently Q satisfies

$$Q^H A Q = \Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, \qquad Q^H B Q = \Omega = \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix}, \qquad \Lambda^2 + \Omega^2 = I. \tag{2.2.3}$$

Substituting $Z_i = Q_i P_i$ into $Z_i^H Z_i = I$, we get

$$P_i P_i^H = (Q_i^H Q_i)^{-1}, \qquad i=1,2.$$

Moreover, substituting the unique decompositions $P_i = H_i V_i$ ($H_i > 0$ and V_i unitary) into the above relations we have $H_i = (Q_i^H Q_i)^{-\frac{1}{2}}$, $i = 1, 2$. Hence

$$P_i = (Q_i^H Q_i)^{-\frac{1}{2}} V_i \tag{2.2.4}$$

and

$$Z_i = Q_i (Q_i^H Q_i)^{-\frac{1}{2}} V_i, \qquad V_i \text{ unitary}, \ i = 1, 2. \tag{2.2.5}$$

Further from (2.1.1) it follows that

$$W'^H W' = (Z^H Z)^{-1} = \begin{pmatrix} I & Z_1^H Z_2 \\ Z_2^H Z_1 & I \end{pmatrix}^{-1}$$
$$= \begin{pmatrix} I + Z_1^H Z_2 (I - Z_2^H Z_1 Z_1^H Z_2)^{-1} Z_2^H Z_1 & * \\ * & (I - Z_2^H Z_1 Z_1^H Z_2)^{-1} \end{pmatrix},$$

combining this relation with (2.2.5) we obtain

$$W_1'^H W_1' = I + Z_1^H Z_2 (I - Z_2^H Z_1 Z_1^H Z_2)^{-1} Z_2^H Z_1 = (I - Z_1^H Z_2 Z_2^H Z_1)^{-1}. \tag{2.2.6}$$

Therefore using (2.1.3) and (2.2.1) we get

$$A_1' = (W_1'^H W_1')^{\frac{1}{2}} P_1^H \Lambda_1 P_1, \qquad B_1' = (W_1'^H W_1')^{\frac{1}{2}} P_1^H \Omega_1 P_1, \tag{2.2.7}$$

where P_1 and $W_1'^H W_1'$ are given in (2.2.4) and (2.2.6), respectively.

Similarly, since $(\tilde{A}_i, \tilde{B}_i)$ $(i = 1, 2)$ are definite pairs, there are non-singular matrices $\tilde{P}_i$ and real diagonal matrices $\tilde{\Lambda}_i$ and $\tilde{\Omega}_i$ $(i = 1, 2)$ such that

$$\begin{gathered} \begin{pmatrix} \tilde{\Lambda}_1 & 0 \\ 0 & \tilde{\Lambda}_2 \end{pmatrix} = \operatorname{diag}(\tilde{\alpha}_k), \qquad \begin{pmatrix} \tilde{\Omega}_1 & 0 \\ 0 & \tilde{\Omega}_2 \end{pmatrix} = \operatorname{diag}(\tilde{\beta}_k), \\ (\tilde{A}_i, \tilde{B}_i) = \tilde{P}_i^H (\tilde{\Lambda}_i \tilde{P}_i, \tilde{\Omega}_i \tilde{P}_i), \qquad \tilde{\Lambda}_i^2 + \tilde{\Omega}_i^2 = I, \qquad i = 1, 2. \end{gathered} \tag{2.2.8}$$

Let

$$\tilde{Q} = \tilde{Z} \begin{pmatrix} \tilde{P}_1^{-1} & 0 \\ 0 & \tilde{P}_2^{-1} \end{pmatrix} = (\tilde{Q}_1, \tilde{Q}_2), \qquad \tilde{Q}_1 \in \mathbb{C}^{n \times \ell};$$

evidently $\tilde{Q}$ satisfies

$$\tilde{Q}^H \tilde{A} \tilde{Q} = \tilde{\Lambda} = \begin{pmatrix} \tilde{\Lambda}_1 & 0 \\ 0 & \tilde{\Lambda}_2 \end{pmatrix}, \qquad \tilde{Q}^H \tilde{B} \tilde{Q} = \tilde{\Omega} = \begin{pmatrix} \tilde{\Omega}_1 & 0 \\ 0 & \tilde{\Omega}_2 \end{pmatrix}, \qquad \tilde{\Lambda}^2 + \tilde{\Omega}^2 = I. \tag{2.2.9}$$

In the same way as above we have

$$\tilde{P}_i = (\tilde{Q}_i^H \tilde{Q}_i)^{-\frac{1}{2}} \tilde{V}_i \tag{2.2.10}$$

and

$$\tilde{Z}_i = \tilde{Q}_i (\tilde{Q}_i^H \tilde{Q}_i)^{-\frac{1}{2}} \tilde{V}_i, \qquad \tilde{V}_i \text{ unitary}, \ i = 1, 2, \tag{2.2.11}$$

and

$$\tilde{W}_2'^H \tilde{W}_2' = (I - \tilde{Z}_2^H \tilde{Z}_1 \tilde{Z}_1^H \tilde{Z}_2)^{-1}. \tag{2.2.12}$$

Hence by (2.1.10) and (2.2.8) we obtain

$$\hat{A}_2 = \tilde{P}_2^H \tilde{\Lambda}_2 \tilde{P}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}}, \qquad \hat{B}_2 = \tilde{P}_2^H \tilde{\Omega}_2 \tilde{P}_2 (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}}, \tag{2.2.13}$$

where $\tilde{P}_2$ and $\tilde{W}_2'^H \tilde{W}_2'$ are given in (2.2.10) and (2.2.12), respectively.

Substituting (2.2.7) and (2.2.13) into the Eqs. (2.1.13), we obtain the simplified perturbation equations

$$\Lambda_1 X_1 - Y_1 \tilde{\Lambda}_2 = C_1, \qquad \Omega_1 X_1 - Y_1 \tilde{\Omega}_2 = D_1, \tag{2.2.14}$$

where

$$X_1 = P_1 (W_1'^H W_1')^{\frac{1}{2}} X \tilde{P}_2^{-1}, \qquad Y_1 = P_1^{-H} Y (\tilde{W}_2'^H \tilde{W}_2')^{\frac{1}{2}} \tilde{P}_2^H \tag{2.2.15}$$

and

$$C_1 = P_1^{-H} C \tilde{P}_2^{-1}, \qquad D_1 = P_1^{-H} D \tilde{P}_2^{-1}. \tag{2.2.16}$$

2.3. The Proof of the Inequality (2.3) *for* $\gamma = 0$

From (2.1) and $\alpha < 1$ we know that $\beta_i \neq 0$ and

$$\frac{|\alpha_i/\beta_i|}{\sqrt{1+(\alpha_i/\beta_i)^2}} \leqq \alpha \qquad \text{for } 1 \leqq i \leqq \ell, \tag{2.3.1}$$

and thus

$$\left(\frac{\alpha_i}{\beta_i}\right)^2 \leqq \frac{\alpha^2}{1-\alpha^2}, \qquad 1 \leqq i \leqq \ell.$$

Combining the above inequalities with $\alpha_i^2 + \beta_i^2 = 1$ (see (2.2.1)) we obtain

$$|\alpha_i| \leqq \alpha, \qquad \frac{1}{|\beta_i|} \leqq \frac{1}{\sqrt{1-\alpha^2}}, \qquad 1 \leqq i \leqq \ell,$$

i.e.

$$\|\Lambda_1\|_2 \leqq \alpha, \qquad \|\Omega_1^{-1}\|_2 \leqq \frac{1}{\sqrt{1-\alpha^2}}. \tag{2.3.2}$$

Similarly from (2.2) and $\alpha + \delta > 0$ we know that $\tilde{\alpha}_j \neq 0$ and

$$\frac{1}{\sqrt{1+(\tilde{\beta}_j/\tilde{\alpha}_j)^2}} \geqq \alpha + \delta \qquad \text{for } \ell + 1 \leqq j \leqq n, \tag{2.3.3}$$

and thus

$$\left(\frac{\tilde{\beta}_j}{\tilde{\alpha}_j}\right)^2 \leqq \frac{1-(\alpha+\delta)^2}{(\alpha+\delta)^2}, \qquad \ell + 1 \leqq j \leqq n.$$

Combining the above inequalities with $\tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$ (see (2.2.8)) we obtain

$$|\tilde{\beta}_j| \leqq \sqrt{1-(\alpha+\delta)^2}, \qquad \frac{1}{|\tilde{\alpha}_j|} \leqq \frac{1}{\alpha+\delta}, \qquad \ell + 1 \leqq j \leqq n,$$

i.e.

$$\|\tilde{\Lambda}_2^{-1}\|_2 \leqq \frac{1}{\alpha+\delta}, \qquad \|\tilde{\Omega}_2\|_2 \leqq \sqrt{1-(\alpha+\delta)^2}. \tag{2.3.4}$$

Combine (2.3.2) and (2.3.4) with the Eqs. (2.2.14) and remember that every unitary-invariant norm $\|\ \|$ is compatible with the spectral norm (Ref. [2], 23; [6], 638). Then from

$$\|Y_1\| \leqq \|Y_1 \tilde{\Lambda}_2\| \, \|\tilde{\Lambda}_2^{-1}\|_2 \leqq \frac{1}{\alpha+\delta} \|Y_1 \tilde{\Lambda}_2\|$$

and

$$\|X_1\| \leqq \|\Omega_1^{-1}\|_2 \, \|\Omega_1 X_1\| \leqq \frac{1}{\sqrt{1-\alpha^2}} \|\Omega_1 X_1\|,$$

we get

$$\|C_1\| \geqq \|Y_1 \tilde{A}_2\| - \|A_1 X_1\| \geqq (\alpha+\delta) \, \|Y_1\| - \alpha \, \|X_1\|$$

and

$$\|D_1\| \geqq \|\Omega_1 X_1\| - \|Y_1 \tilde{\Omega}_2\| \geqq \sqrt{1-\alpha^2} \, \|X_1\| - \sqrt{1-(\alpha+\delta)^2} \, \|Y_1\|,$$

so that

$$\|Y_1\| \leqq \frac{\|C_1\| + \alpha \, \|X_1\|}{\alpha+\delta} \leqq \frac{\|C_1\| + \alpha(\|D_1\| + \sqrt{1-(\alpha+\delta)^2} \, \|Y_1\|)/\sqrt{1-\alpha^2}}{\alpha+\delta}.$$

Therefore

$$\begin{aligned} \|Y_1\| &\leqq \frac{(\sqrt{1-\alpha^2} \, \|C_1\| + \alpha \, \|D_1\|) \, [(\alpha+\delta)\sqrt{1-\alpha^2} + \alpha\sqrt{1-(\alpha+\delta)^2}]}{(\alpha+\delta)^2 - \alpha^2} \\ &\leqq p(\alpha, \delta; 0) \sqrt{\|C_1\|^2 + \|D_1\|^2}/\delta. \end{aligned} \tag{2.3.5}$$

Observe that $(\tilde{W}_2'^H \tilde{W}_2')^{-1} \leqq I^{(n-\ell)}$ (see (2.2.12)); then from (2.2.15), (2.2.16), (2.1.12) and (2.1.7) we have

$$\|Y\| \leqq \|P_1\|_2 \, \|\tilde{P}_2^{-1}\|_2 \, \|Y_1\| \tag{2.3.6}$$

and

$$\begin{aligned} \sqrt{\|C_1\|^2 + \|D_1\|^2} &\leqq \|P_1^{-1}\|_2 \, \|\tilde{P}_2^{-1}\|_2 \sqrt{\|C\|^2 + \|D\|^2} \\ &\leqq \|P_1^{-1}\|_2 \, \|\tilde{P}_2^{-1}\|_2 \, \|(EZ_1, FZ_1)\|. \end{aligned} \tag{2.3.7}$$

Substituting (2.3.5) and (2.3.7) into (2.3.6) we obtain

$$\|Y\| \leqq p(\alpha, \delta; 0) \, \|P_1\|_2 \, \|P_1^{-1}\|_2 \, \|\tilde{P}_2^{-1}\|_2^2 \, \|(EZ_1, FZ_1)/\delta. \tag{2.3.8}$$

Now we estimate $\|P_1\|_2$, $\|P_2^{-1}\|_2$ and $\|\tilde{P}_2^{-1}\|_2^2$.

First, it follows from (2.2.4) and (2.2.10) that

$$\begin{aligned} &\|P_1\|_2 \leqq \|(Q_1^H Q_1)^{-1}\|_2^{\frac{1}{2}}, \qquad \|P_1^{-1}\|_2 \leqq \|(Q_1^H Q_1)\|_2^{\frac{1}{2}}, \\ &\|\tilde{P}_2^{-1}\|_2^2 \leqq \|(\tilde{Q}_2^H \tilde{Q}_2)\|_2. \end{aligned} \tag{2.3.9}$$

Secondly, by Theorem 1.2, for the definite pair (A, B) there exists a real number φ such that $B_\varphi = \sin\varphi \, A + \cos\varphi \, B > 0$ and $\lambda_{\min}(B_\varphi) = c(A, B)$. Utilizing the decomposition (2.2.3) and writing

$$\Omega_\varphi = \sin\varphi \, \Lambda + \cos\varphi \, \Omega = \begin{pmatrix} \Omega_{1\varphi} & 0 \\ 0 & \Omega_{2\varphi} \end{pmatrix},$$

where Ω_φ is a real diagonal matrix, $0 < \Omega_\varphi \leqq I$ and $\Omega_{1\varphi} \in \mathbb{C}^{\ell \times \ell}$, then we have

$$c(A, B) \, Q_1^H Q_1 \leqq Q_1^H B_\varphi Q_1 = \Omega_{1\varphi}$$

i.e.

$$Q_1^H Q_1 \leqq \frac{\Omega_{1\varphi}}{c(A, B)} \leqq \frac{1}{c(A, B)} I^{(\ell)}, \tag{2.3.10}$$

and thus

$$\|(Q_1^H Q_1)\|_2^{\frac{1}{2}} \leqq 1/\sqrt{c(A,B)}. \tag{2.3.11}$$

Similarly we have

$$\tilde{Q}_2^H \tilde{Q}_2 \leqq \frac{1}{c(\tilde{A},\tilde{B})} I^{(n-\ell)}, \qquad \|\tilde{Q}_2^H \tilde{Q}_2\|_2 \leqq 1/c(\tilde{A},\tilde{B}). \tag{2.3.12}$$

Thirdly, from (2.2.3),

$$(Q_1^H A Q_1)^2 + (Q_1^H B Q_1)^2 = I^{(\ell)}. \tag{2.3.13}$$

Decomposing

$$Q_1 = V_1 K_1, \quad V_1 \in \mathbb{C}^{n\times \ell}, \quad V_1^H V_1 = I^{(\ell)}, \quad \text{non-singular } K_1 \in \mathbb{C}^{\ell\times\ell}, \tag{2.3.14}$$

substituting this decomposition into (2.3.13), and writing

$$V_1^H A V_1 = M, \qquad V_1^H B V_1 = N, \tag{2.3.15}$$

we get

$$(K_1 K_1^H)^{-1} = M K_1 K_1^H M + N K_1 K_1^H N. \tag{2.3.16}$$

Therefore

$$\begin{aligned}\|(Q_1^H Q_1)^{-1}\|_2 &= \|(K_1^H K_1)^{-1}\|_2 = \|(K_1 K_1^H)^{-1}\|_2 \\ &\leqq \|M^2+N^2\|_2 \, \|(K_1 K_1^H)\|_2. \end{aligned} \tag{2.3.17}$$

Moreover, it follows from (2.3.14) and (2.3.11) that

$$\|(K_1 K_1^H)\|_2 = \|(K_1^H K_1)\|_2 = \|(Q_1^H Q_1)\|_2 \leqq \frac{1}{c(A,B)}, \tag{2.3.18}$$

and from (2.3.14) and (2.3.15),

$$\|M^2+N^2\|_2 \leqq \|A^2+B^2\|_2. \tag{2.3.19}$$

Hence from (2.3.17)–(2.3.19) we get

$$\|(Q_1^H Q_1)^{-1}\|_2 \leqq \frac{\|A^2+B^2\|_2}{c(A,B)}. \tag{2.3.20}$$

Fourthly, substituting (2.3.11), (2.3.12) and (2.3.20) into (2.3.9), we get

$$\|P_1\|_2 \leqq \frac{\|(A,B)\|_2}{\sqrt{c(A,B)}}, \qquad \|P_1^{-1}\|_2 \leqq \frac{1}{\sqrt{c(A,B)}}, \qquad \|\tilde{P}_2^{-1}\|_2^2 \leqq \frac{1}{c(\tilde{A},\tilde{B})}. \tag{2.3.21}$$

Finally, substituting (2.3.21) into (2.3.8) we obtain (2.3) for $\gamma=0$.

2.4. The Proof of the Inequality (2.3) *for* $\gamma \neq 0$

Let $c=1/\sqrt{1+\gamma^2}$, $s=\gamma/\sqrt{1+\gamma^2}$,

$$\Lambda_{10} = \operatorname{diag}(\alpha_{i0}) = c\,\Lambda_1 - s\,\Omega_1, \qquad \Omega_{10} = \operatorname{diag}(\beta_{i0}) = s\,\Lambda_1 + c\,\Omega_1$$

and

$$\tilde{\Lambda}_{20} = \operatorname{diag}(\tilde{\alpha}_{j0}) = c\,\tilde{\Lambda}_2 - s\,\Omega_2, \qquad \tilde{\Omega}_{20} = \operatorname{diag}(\tilde{\beta}_{j0}) = s\,\tilde{\Lambda}_2 + c\,\tilde{\Omega}_2.$$

We have

$$\alpha_{i0}=\frac{\alpha_i-\gamma\beta_i}{\sqrt{1+\gamma^2}},\quad \beta_{i0}=\frac{\gamma\alpha_i+\beta_i}{\sqrt{1+\gamma^2}}\quad \text{for } 1\leqq i\leqq \ell$$

and

$$\tilde{\alpha}_{j0}=\frac{\tilde{\alpha}_j-\gamma\tilde{\beta}_j}{\sqrt{1+\gamma^2}},\quad \tilde{\beta}_{j0}=\frac{\gamma\tilde{\alpha}_j+\tilde{\beta}_j}{\sqrt{1+\gamma^2}}\quad \text{for } \ell+1\leqq j\leqq n$$

which satisfy

$$\rho((0,1),(\alpha_{i0},\beta_{i0}))\leqq\alpha \qquad 1\leqq i\leqq \ell$$

and

$$\rho((0,1),(\tilde{\alpha}_{j0},\tilde{\beta}_{j0}))\geqq\alpha+\delta \qquad \ell+1\leqq j\leqq n.$$

At the same time the Eqs. (2.2.14) become

$$A_{10}X_1-Y_1\tilde{A}_{20}=C_{10},\qquad \Omega_{10}X_1-Y_1\tilde{\Omega}_{20}=D_{10},$$

where

$$C_{10}=c\,C_1-s\,D_1,\qquad D_{10}=s\,C_1+c\,D_1.$$

According to the proof in 2.3 (see (2.3.5) and (2.3.6)) we have

$$\|Y\|\leqq p(\alpha,\delta;0)\,\|P_1\|_2\,\|\tilde{P}_2^{-1}\|_2\,(\sqrt{1-\alpha^2}\,\|C_{10}\|+\alpha\,\|D_{10}\|)/\delta$$

for every unitary-invariant norm. However,

$$\begin{aligned}\sqrt{1-\alpha^2}\,\|C_{10}\|+\alpha\,\|D_{10}\|&\leqq\sqrt{\|C_{10}\|^2+\|D_{10}\|^2}\\&\leqq\sqrt{2(\|C_1\|^2+\|D_1\|^2)},\end{aligned}\tag{2.4.1}$$

hence the perturbation bound in Theorem 2.1 contains the factor $\sqrt{2}$ for $\gamma\neq 0$. The proof is complete.

§ 3. The Generalized sin θ Theorem

We notice that for the norm $\|\ \|_F$ the inequalities (2.4.1) become

$$\sqrt{1-\alpha^2}\,\|C_{10}\|_F+\alpha\,\|D_{10}\|_F\leqq\sqrt{\|C_{10}\|_F^2+\|D_{10}\|_F^2}=\sqrt{\|C_1\|_F^2+\|D_1\|_F^2},$$

and thus $q(\gamma)\equiv 1$ in (2.4). Hence from (2.3) we obtain

$$\|\sin\Theta_1\|_F\leqq\frac{\|(A,B)\|_2}{c(A,B)\,c(\tilde{A},\tilde{B})}\cdot\frac{\|(EZ_1,FZ_1)\|_F}{\delta},\tag{3.1}$$

where

$$\|(EZ_1,FZ_1)\|_F=\sqrt{\|EZ_1\|_F^2+\|FZ_1\|_F^2}.$$

In this section we prove that under weaker conditions the inequality (3.1) is also valid.

Theorem 3.1. *(The generalized* $\sin\theta$ *theorem for definite pairs.) Let* (A, B), $(\tilde{A}, \tilde{B})$, Z *and* $\tilde{Z}$ *be the same as in Theorem* 2.1. *Set*

$$\delta \equiv \min_{i,j}\{\rho((\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)) \colon (\alpha_i, \beta_i) \in \lambda(A_1, B_1), (\tilde{\alpha}_j, \tilde{\beta}_j) \in \lambda(\tilde{A}_2, \tilde{B}_2)\}.$$

If $\delta > 0$, *then the inequality* (3.1) *is valid.*

Proof. We associate the matrices X_1 and $Y_1 \in \mathbb{C}^{\ell \times (n-\ell)}$ with the $\ell(n-\ell)$-vectors x and y which are the direct sums of the column vectors of X_1 and Y_1, respectively. Similarly, C_1 and $D_1 \in \mathbb{C}^{\ell \times (n-\ell)}$ with the $\ell(n-\ell)$-vectors c and d, so that the Eqs. (2.2.14) take the form

$$\begin{aligned} (I \otimes \Lambda_1)\, x - (\tilde{\Lambda}_2 \otimes I)\, y &= c \\ (I \otimes \Omega_1)\, x - (\tilde{\Omega}_2 \otimes I)\, y &= d, \end{aligned} \tag{3.2}$$

where $\otimes$ is the Kronecker product symbol (see [4], 8–9).

From (3.2) we obtain

$$(\tilde{\Lambda}_2 \otimes \Omega_1 - \tilde{\Omega}_2 \otimes \Lambda_1)\, y = -(I \otimes \Omega_1)\, c + (I \otimes \Lambda_1)\, d$$

and

$$\min_{\substack{1 \leq i \leq \ell \\ \ell+1 \leq j \leq n}} |\alpha_i \tilde{\beta}_j - \beta_i \tilde{\alpha}_j| \, \|y\| = \sqrt{\|c\|^2 + \|d\|^2}$$

i.e.

$$\|Y_1\|_F \leq \frac{\sqrt{\|C_1\|_F^2 + \|D_1\|_F^2}}{\delta}. \tag{3.3}$$

Moreover, according to (2.2.16), (2.1.12) and (2.1.7) we have

$$\begin{aligned} \sqrt{\|C_1\|_F^2 + \|D_1\|_F^2} &= \sqrt{\|P_1^{-H} R_A^H \tilde{Z}_2 \tilde{P}_1^{-1}\|_F^2 + \|P_1^{-H} R_B^H \tilde{Z}_2 \tilde{P}_2^{-1}\|_F^2} \\ &\leq \|P_1^{-1}\|_2 \, \|\tilde{P}_2^{-1}\|_2 \sqrt{\|R_A\|_F^2 + \|R_B\|_F^2} \\ &\leq \|P_1^{-1}\|_2 \, \|\tilde{P}_2^{-1}\|_2 \, \|(EZ_1, FZ_1)\|_F, \end{aligned}$$

and so it follows from (2.2.15) and (3.3) that

$$\|Y\|_F \leq \|P_1\|_2 \|P_1^{-1}\|_2 \|\tilde{P}_2^{-1}\|_2^2 \, \|(EZ_1, FZ_1)\|_F / \delta. \tag{3.4}$$

Substituting (2.3.21) into (3.4), we obtain (3.1). The proof is complete.

§ 4. The $\sin 2\theta$ Theorem

In this section we hypothesize a gap between $\lambda(\tilde{A}_1, \tilde{B}_1)$ and $\lambda(\tilde{A}_2, \tilde{B}_2)$, and thus obtain the $\sin 2\theta$ theorem for definite pairs.

Theorem 4.1. *(The* $\sin 2\theta$ *theorem for definite pairs). Let* (A, B), $(\tilde{A}, \tilde{B}) = (A + E, B + F) \in \mathbb{D}(n)$ *with decompositions given in* (1.4) *and* (1.5), $\lambda(\tilde{A}_1, \tilde{B}_1) = \{(\tilde{\alpha}_i, \tilde{\beta}_i)\}_{i=1}^{\ell}$ *and* $\lambda(\tilde{A}_2, \tilde{B}_2) = \{(\tilde{\alpha}_i, \tilde{\beta}_i)\}_{i=\ell+1}^{n}$. *Assume that there are* $\alpha \geq 0$ *and*

$\delta>0$ satisfying $\alpha+\delta\leqq 1$, and γ a real number, such that

$$\rho((\gamma,1),(\tilde{\alpha}_i,\tilde{\beta}_i))\begin{cases}\leqq\alpha & \forall(\tilde{\alpha}_i,\tilde{\beta}_i)\in\lambda(\tilde{A}_1,\tilde{B}_1)\\ \geqq\alpha+\delta & \forall(\tilde{\alpha}_i,\tilde{\beta}_i)\in\lambda(\tilde{A}_2,\tilde{B}_2)\end{cases} \tag{4.1}$$

(or vice-versa). Then for every unitary-invariant matrix norm,

$$\|\sin 2\Theta_1\|-\frac{2\xi}{1-\xi^2}\|\sin\Theta_1\|^2\leqq\frac{r\,\|(E,F)\|}{\delta}, \tag{4.2}$$

where Θ_1 is defined by (1.7) and $\|(E,F)\|$ by (2.5),

$$\xi=\|Z_1^H Z_2\|_2,\qquad \eta=\frac{1+\xi}{1-\xi} \tag{4.3}$$

and

$$r\equiv r(\tilde{A},\tilde{B};\alpha,\delta;\gamma;\eta)=\frac{p(\alpha,\delta;\gamma)\,\eta^2(\eta+1)\,\|(\tilde{A},\tilde{B})\|_2}{(c(\tilde{A},\tilde{B}))^2}; \tag{4.4}$$

here $p(\alpha,\delta;\gamma)$ is given in (2.4) and $\|(\tilde{A},\tilde{B})\|_2$ is defined by (2.5).

Proof. 1. Let

$$W'=Z^{-H}=(W_1',W_2'),\qquad \tilde{W}'=\tilde{Z}^{-H}=(\tilde{W}_1',\tilde{W}_2'),\qquad W_1' \text{ and } \tilde{W}_1'\in\mathbb{C}^{n\times\ell} \tag{4.5}$$

and

$$X=W_1'Z_1^H-W_2'Z_2^H. \tag{4.6}$$

The matrix X satisfies

$$X^2=I, \tag{4.7}$$

and from the Lemmas 1.1–1.2 and (1.5) we get

$$\begin{aligned}\|X\|_2&=\left\|W'\begin{pmatrix}I^{(\ell)} & 0\\ 0 & -I\end{pmatrix}Z^H\right\|_2\leqq\|W'\|_2\,\|Z\|_2=\sqrt{\|(Z^HZ)^{-1}\|_2\,\|Z^HZ\|_2}\\ &\leqq\sqrt{\frac{1+\|Z_1^HZ_2\|_2}{1-\|Z_1^HZ_2\|_2}}=\sqrt{\eta}.\end{aligned} \tag{4.8}$$

Writing

$$Z^HEZ=\begin{pmatrix}E_1 & P^H\\ P & E_2\end{pmatrix},\qquad Z^HFZ=\begin{pmatrix}F_1 & Q^H\\ Q & F_2\end{pmatrix}, \tag{4.9}$$

then it follows from (1.2), (4.5) and (4.6) that

$$\begin{aligned}A+XEX^H&=W'\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}W'^H+(W_1'Z_1^H-W_2'Z_2^H)\,W'\begin{pmatrix}E_1 & P^H\\ P & E_2\end{pmatrix}W'(Z_1W_1'^H-Z_2W_2'^H)\\ &=(W_1',-W_2')\begin{pmatrix}A_1+E_1 & P^H\\ P & A_2+E_2\end{pmatrix}\begin{pmatrix}W_1'^H\\ -W_2'^H\end{pmatrix}\\ &=XW'\begin{pmatrix}A_1+E_1 & P^H\\ P & A_2+E_2\end{pmatrix}W'^HX^H\\ &=X\tilde{A}X^H=X\tilde{W}'\begin{pmatrix}\tilde{A}_1 & 0\\ 0 & \tilde{A}_2\end{pmatrix}\tilde{W}'^HX^H.\end{aligned} \tag{4.10}$$

Similarly,

$$B+XFX^H=X\tilde{B}X^H=X\tilde{W}'\begin{pmatrix}\tilde{B}_1 & 0\\ 0 & \tilde{B}_2\end{pmatrix}\tilde{W}'^H X^H. \tag{4.11}$$

Let

$$\hat{A}=A+XEX^H,\qquad \hat{B}=B+XFX^H \tag{4.12}$$

and

$$\hat{Z}_1=X^H\tilde{Z}_1(\tilde{Z}_1^H XX^H\tilde{Z}_1)^{-\frac{1}{2}},\qquad \hat{Z}_2=X^H\tilde{Z}_2(\tilde{Z}_2^H XX^H\tilde{Z}_2)^{-\frac{1}{2}},$$
$$\hat{Z}=(\hat{Z}_1,\hat{Z}_2). \tag{4.13}$$

It is observed that

$$(\tilde{W}'^H X^H)^{-1}=X^H\tilde{Z}, \tag{4.14}$$

then we can write (4.10) and (4.11) as

$$\hat{Z}^H\hat{A}\hat{Z}=\begin{pmatrix}\hat{A}_1 & 0\\ 0 & \hat{A}_2\end{pmatrix},\qquad \hat{Z}^H\hat{B}\hat{Z}=\begin{pmatrix}\hat{B}_1 & 0\\ 0 & \hat{B}_2\end{pmatrix}, \tag{4.15}$$

where $\hat{Z}_1^H\hat{Z}_1=I^{(\ell)}$, $\hat{Z}_2^H\hat{Z}_2=I^{(n-\ell)}$ and

$$\hat{A}_i=H_i\tilde{A}_iH_i,\qquad \hat{B}_i=H_i\tilde{B}_iH_i,\qquad H_i=(\tilde{Z}_i^H XX^H\tilde{Z}_i)^{-\frac{1}{2}},\qquad i=1,2. \tag{4.16}$$

From (4.14), (4.12) and (4.10) we have

$$\begin{aligned}c(\hat{A},\hat{B})&\equiv\inf_{x\neq 0}\left\{\frac{(x^H X\tilde{A}X^H x)^2+(x^H X\tilde{B}X^H x)^2}{x^H x}\right\}^{\frac{1}{2}}\\ &=\inf_{y\neq 0}\left\{\frac{(y^H\tilde{A}y)^2+(y^H\tilde{B}y)^2}{y^H XX^H y}\right\}^{\frac{1}{2}}\\ &\geqq\frac{1}{\|X\|_2}c(\tilde{A},\tilde{B})\geqq\frac{1}{\sqrt{\eta}}c(\tilde{A},\tilde{B})>0,\end{aligned} \tag{4.17}$$

which shows that the Hermitian matrix-pair $(\hat{A},\hat{B})$ is in $\mathbb{D}(n)$. Moreover, from (4.16) we have

$$\lambda(\hat{A}_2,\hat{B}_2)=\lambda(\tilde{A}_2,\tilde{B}_2). \tag{4.18}$$

Therefore the decompositions (4.15) and the location of the generalized eigenvalues of the definite pairs $(\hat{A},\hat{B})$ and $(\tilde{A},\tilde{B})$ satisfy the hypotheses in Theorem 2.1, hence for

$$\hat{\Theta}_1\equiv\arccos(\hat{Z}_1^H\tilde{Z}_1\tilde{Z}_1^H\hat{Z}_1)^{\frac{1}{2}}\geqq 0 \tag{4.19}$$

and

$$\hat{E}=\tilde{A}-\hat{A}=E-XEX^H,\qquad \hat{F}=\tilde{B}-\hat{B}=F-XFX^H, \tag{4.20}$$

and for every unitary-invariant norm we have

$$\|\sin\hat{\Theta}_1\|\leqq\frac{p(\alpha,\delta;\gamma)\,\|(\hat{A},\hat{B})\|_2}{c(\hat{A},\hat{B})\,c(\tilde{A},\tilde{B})}\cdot\frac{\|(\hat{E}\hat{Z}_1,\hat{F}\hat{Z}_1)\|}{\delta}. \tag{4.21}$$

2. Let $W, \tilde{W}$ be defined by (2.1.1) and (2.1.2). Then obviously $Z=(Z_1, Z_2)$, $W=(W_1, W_2)$, $\tilde{Z}=(\tilde{Z}_1, \tilde{Z}_2)$ and $\tilde{W}=(\tilde{W}_1, \tilde{W}_2)$ satisfy (2.1.5).

By the definition of $\hat{\Theta}_1$ (see (4.19)) we have (Ref. (1.11))

$$\|\sin \hat{\Theta}_1\| = \|\hat{Z}_1^H \tilde{W}_2\| = \|(\tilde{Z}_1^H X X^H \tilde{Z}_1)^{-\frac{1}{2}} \tilde{Z}_1^H X \tilde{W}_2\|,$$

and thus

$$\|\tilde{Z}_1^H X \tilde{W}_2\| \leqq \|(\tilde{Z}_1^H X X^H \tilde{Z}_1)^{\frac{1}{2}}\|_2 \, \|\sin \hat{\Theta}_1\|. \tag{4.22}$$

Utilizing (4.6) we get

$$\tilde{Z}_1^H X \tilde{W}_2 = \tilde{Z}_1^H W_1' Z_1^H \tilde{W}_2 - \tilde{Z}_1^H W_2' Z_2^H \tilde{W}_2,$$

but from (4.5) and (2.1.5) it follows that

$$\tilde{Z}_1^H W_1' Z_1^H \tilde{W}_2 + \tilde{Z}_1^H W_2' Z_2^H \tilde{W}_2 = \tilde{Z}_1^H (W' Z^H) \tilde{W}_2 = \tilde{Z}_1^H \tilde{W}_2 = 0,$$

therefore

$$\tilde{Z}_1^H X \tilde{W}_2 = 2 \tilde{Z}_1^H W_1' Z_1^H \tilde{W}_2. \tag{4.23}$$

Moreover, observe that the matrix (Z_1, W_2) is unitary, and $Z^H W' = I$; therefore we obtain

$$W_1' = (Z_1 Z_1^H + W_2 W_2^H) W_1' = Z_1 + W_2 W_2^H W_1'.$$

Substituting the above relation into (4.23), we get

$$\tilde{Z}_1^H X \tilde{W}_2 = 2 \tilde{Z}_1^H Z_1 Z_1^H \tilde{W}_2 + 2 \tilde{Z}_1^H W_2 W_2^H W_1' Z_1^H \tilde{W}_2.$$

Then combining this equality with (4.22) we obtain

$$\begin{aligned} &2\|Z_1^H \tilde{Z}_1 Z_1^H \tilde{W}_2\| - 2\|W_2^H W_1'\|_2 \, \|\tilde{Z}_1^H W_2\| \, \|Z_1^H \tilde{W}_2\| \\ &\leqq \|(\tilde{Z}_1^H X X^H \tilde{Z}_1)^{\frac{1}{2}}\|_2 \, \|\sin \hat{\Theta}_1\|. \end{aligned} \tag{4.24}$$

3. Now we deduce the inequation (4.2) from (4.21) and (4.24). First, from (2.1.5) we know that the matrix

$$(Z_1, W_2)^H (\tilde{Z}_1, \tilde{W}_2) = \begin{pmatrix} Z_1^H \tilde{Z}_1 & Z_1^H \tilde{W}_2 \\ W_2^H \tilde{Z}_1 & W_2^H \tilde{W}_2 \end{pmatrix}$$

is unitary. By Lemma 1.3 there are unitary matrices U, V_1, V_2, V, U_1 and U_2, such that

$$Z_1^H \tilde{Z}_1 = U \Gamma V_1^H, \quad Z_1^H \tilde{W}_2 = U \Sigma V_2^H, \quad \tilde{Z}_1^H Z_1 = V \Gamma^T U_1^H, \quad \tilde{Z}_1^H W_2 = V \Sigma^T U_2^H, \tag{4.25}$$

where

$$\Gamma = \operatorname{diag}(\gamma_1, \gamma_2, \ldots), \qquad \Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots),$$

$$\gamma_i = \cos \theta_i, \qquad \sigma_i = \sin \theta_i, \qquad i = 1, 2, \ldots,$$

and $\frac{\pi}{2} \geqq \theta_1 \geqq \theta_2 \geqq \ldots \geqq 0$. From (4.25) we get

$$\|Z_1^H \tilde{W}_2\| = \|\Sigma\| = \|\tilde{Z}_1^H W_2\|.$$

But according to the definition of Θ_1 (see (1.7)) we have the relation (1.11), therefore

$$\|\tilde{Z}_1^H W_2\| = \|\sin \Theta_1\|. \tag{4.26}$$

Further, from (4.25) we get

$$2\|\tilde{Z}^H Z_1 Z_1^H \tilde{W}_2\| = 2\| V_1 \Gamma^T \Sigma V_2^H\| = \|\mathrm{diag}(\sin 2\theta_i)\| = \|\sin 2\Theta_1\|. \tag{4.27}$$

Secondly, it follows from (2.1.2) that

$$\| W_2^H W_1'\|_2 \leqq \|(W_2'^H W_2')^{-1}\|_2^{\frac{1}{2}} \| W_2'^H W_1'\|_2. \tag{4.28}$$

Utilizing

$$W'^H W' = (Z^H Z)^{-1} = \begin{pmatrix} I & Z_1^H Z_2 \\ Z_2^H Z_1 & I \end{pmatrix}^{-1} = \begin{pmatrix} * & * \\ -(I - Z_2^H Z_1 Z_1^H Z_2)^{-1} Z_2^H Z_1 & (I - Z_2^H Z_1 Z_1^H Z_2)^{-1} \end{pmatrix}$$

and $\|Z_1^H Z_2\|_2 < 1$ (see Lemma 1.1), we get

$$\|(W_2'^H W_2')^{-1}\|_2 = \|I - Z_2^H Z_1 Z_1^H Z_2\|_2 \leqq 1$$

and

$$\| W_2'^H W_1'\|_2 = \|(I - Z_2^H Z_1 Z_1^H Z_2)^{-1} Z_2^H Z_1\|_2 \leqq \frac{\|Z_2^H Z_1\|_2}{1 - \|Z_2^H Z_1\|_2^2} = \frac{\xi}{1-\xi^2}.$$

Substituting the above inequalities into (4.28), we obtain

$$\| W_2^H W_1'\|_2 \leqq \frac{\xi}{1-\xi^2}. \tag{4.29}$$

Thirdly, from $\tilde{Z}_1^H \tilde{Z}_1 = I$ and (4.8),

$$\|(\tilde{Z}_1^H X X^H \tilde{Z}_1)^{\frac{1}{2}}\|_2 \leqq \|X\|_2 \leqq \sqrt{\eta}. \tag{4.30}$$

Substituting (1.11), (4.26), (4.27), (4.29) and (4.30) into (4.24), we obtain

$$\|\sin 2\Theta_1\| - \frac{2\xi}{1-\xi^2} \|\sin \Theta_1\|^2 \leqq \sqrt{\eta}\, \|\sin \hat{\Theta}_1\|. \tag{4.31}$$

Finally, from (4.10)–(4.12) and (4.8) we get

$$\|\hat{A}^2 + \hat{B}^2\|_2 \leqq \eta^2 \, \|\tilde{A}^2 + \tilde{B}^2\|_2, \tag{4.32}$$

and from $\| Y_1\|_2 = 1$, (4.20) and (4.8) we get

$$\|\hat{E}\hat{Z}_1\| \leqq (1+\eta)\,\|E\|, \qquad \|\hat{E}\hat{Z}_1\| \leqq (1+\eta)\,\|F\|. \tag{4.33}$$

Substituting (4.17), (4.32) and (4.33) into (4.21), and combining with (4.31), we obtain the inequation (4.2). The proof is complete.

Corollary 4.1. *To the hypotheses of Theorem* 2.1 *add this: Z is a unitary matrix. Then for every unitary-invariant matrix norm* $\| \ \|$, *we have*

$$\|\sin 2\Theta_1\| \leqq \frac{2p(\alpha, \delta; \gamma)\,\|\tilde{A}, \tilde{B})\|_2}{(c(\tilde{A}, \tilde{B}))^2} \cdot \frac{\|(E, F)\|}{\delta}.$$

§ 5. The Strengthened $\sin 2\theta$ Theorem

In this section we give some conditions under which Θ_1 and $\sin 2\Theta_1$ will be small, and we obtain explicit expressions of the perturbation bounds for eigenspaces of a definite pair on the spectral norm.

Theorem 5.1. *To the hypotheses of Theorem* 4.1 *add these:*

$$\omega \equiv \frac{\varepsilon}{c(\tilde{A}, \tilde{B})} < \frac{\delta}{2}, \tag{5.1}$$

$$\rho((\gamma, 1), (\alpha_i, \beta_i)) \leqq \alpha + \frac{\delta}{2} \quad \forall (\alpha_i, \beta_i) \in \lambda(A_1, B_1) \tag{5.2}$$

and

$$\frac{\xi}{1-\xi^2} + \frac{r\varepsilon}{\delta} < 1 - \frac{2\varphi}{\pi}, \tag{5.3}$$

where

$$\varphi = \arctan\left(\frac{\xi}{1-\xi^2}\right), \qquad \varepsilon = \|(E, F)\|_2. \tag{5.4}$$

Then beside (4.2) *we also have*

$$\Theta_1 < \left(\frac{\pi}{4} - \frac{\varphi}{2}\right) I. \tag{5.5}$$

Proof. 1. We consider the family of definite pairs

$$(A(t), B(t)) = (\tilde{A} - tE, \tilde{B} - tF), \qquad 0 \leqq t \leqq 1 \tag{5.6}$$

(on account of the continuity of the function $c(A, B)$ [7] and the condition (5.1) the Hermitian pairs $(A(t), B(t))$ are definite pairs). Let $\lambda(\tilde{A}, \tilde{B}) = \{(\tilde{\alpha}_i, \tilde{\beta}_i)\}_{i=1}^n$ and $\lambda(A(t), B(t)) = \{(\alpha_i(t), \beta_i(t))\}_{i=1}^n$. According to Stewart's inequality ([7], Theorem 3.2) and (5.1) we have

$$\rho((\alpha_i, \beta_i), (\alpha_i(t), \beta_i(t))) \leqq \frac{t\varepsilon}{c(\tilde{A}, \tilde{B})} \leqq \omega < \frac{\delta}{2}, \qquad 1 \leqq i \leqq n$$

for a suitable numbering of $\{(\alpha_i(t), \beta_i(t))\}$. And thus we obtain

$$\begin{aligned} \rho((\gamma, 1), (\alpha_i(t), \beta_i(t)) &\leqq \rho((\gamma, 1), (\tilde{\alpha}_i, \tilde{\beta}_i)) + \rho((\tilde{\alpha}_i, \tilde{\beta}_i), (\alpha_i(t), \beta_i(t)) \\ &\leqq \alpha + \omega < \alpha + \frac{\delta}{2}, \qquad 1 \leqq i \leqq \ell \end{aligned}$$

and

$$\rho((\gamma, 1), (\alpha_i(t), \beta_i(t))) \geqq \rho((\gamma, 1), (\tilde{\alpha}_i, \tilde{\beta}_i)) - ((\tilde{\alpha}_i, \tilde{\beta}_i), (\alpha_i(t), \beta_i(t)))$$
$$\geqq \alpha + \delta - \omega > \alpha + \frac{\delta}{2}, \qquad \ell + 1 \leqq i \leqq n.$$

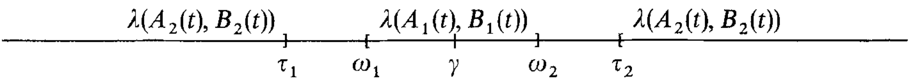
But $\rho((\gamma, 1), (\omega_i, 1)) = \alpha + \omega$ and $\rho((\gamma, 1), (\tau_i, 1)) = \alpha + \delta - \omega$ for $i = 1, 2$. Therefore $\lambda(A(t), B(t))$ is disjoint from $]\tau_1, \omega_1[$ and from $]\omega_2, \tau_2[$.

2. Let $Z(t) = (Z_1(t), Z_2(t))$ be non-singular matrices satisfying

$$Z_1(t)^H Z_1(t) = I^{(\ell)}, \quad Z_2(t)^H Z_2(t) = I^{(n-\ell)},$$
$$Z(t)^H A(t)\, Z(t) = \begin{pmatrix} A_1(t) & 0 \\ 0 & A_2(t) \end{pmatrix}, \quad Z(t)^H B(t)\, Z(t) = \begin{pmatrix} B_1(t) & 0 \\ 0 & B_2(t) \end{pmatrix},$$
$$\lambda(A_1(t), B_1(t)) \subseteq \{(\alpha, \beta) \in \mathscr{P}(1, 1) \colon \rho((\gamma, 1), (\alpha, \beta)) \leqq \alpha + \omega\}$$

and

$$\lambda(A_2(t), B_2(t)) \subseteq \{(\alpha, \beta) \in \mathscr{P}(1, 1) \colon \rho((\gamma, 1), (\alpha, \beta)) \geqq \alpha + \delta - \omega\}.$$



We consider the family of projectors

$$\tilde{P}(t) = Z_1(t)\, Z_1(t)^H, \qquad 0 \leqq t \leqq 1. \tag{5.7}$$

For an arbitrary constant $t \in [0, 1]$ and a small real number Δt which satisfies $0 \leqq t + \Delta t \leqq 1$, on account of the location of the generalized eigenvalues of $(A_1(t), B_1(t))$ and $(A_2(t + \Delta t), B_2(t + \Delta t))$ we can use Theorem 2.1 for

$$\tilde{\Theta}_1(t) \equiv \arccos (Z_1(t)^H Z_1(t + \Delta t)\, Z_1(t + \Delta t)^H Z_1(t))^{\frac{1}{2}} \geqq 0.$$

It is observed that (see (1.12))

$$\|\sin \tilde{\Theta}_1(t)\|_2 = \|\tilde{P}(t) - \tilde{P}(t + \Delta t)\|_2;$$

then according to Theorem 2.1 we get

$$\|\tilde{P}(t) - \tilde{P}(t + \Delta t)\|_2 \leqq \frac{\sqrt{2}\sqrt{\|A(t)\|_2^2 + \|B(t)\|_2^2}}{c(A(t), B(t))\, c(A(t + \Delta t), B(t + \Delta t))}$$
$$\cdot \frac{\sqrt{\|A(t + \Delta t) - A(t)\|_2^2 + \|B(t + \Delta t) - B(t)\|_2^2}}{(\alpha + \delta - \omega) - (\alpha + \omega)}$$
$$\leqq \frac{\sqrt{2}\, \varepsilon \sqrt{(\|A\|_2 + \|E\|_2)^2 + (\|B\|_2 + \|F\|_2)^2}}{(c(\tilde{A}, \tilde{B}) - \varepsilon)^2\, (\delta - 2\omega)} \cdot |\Delta t| \to 0 \qquad (|\Delta t| \to 0).$$

This shows that $\tilde{P}(t)$ is continuous in the spectral norm topology. Now we define

$$\theta(t) = \sin^{-1} \|\tilde{P}(t) - \tilde{P}(0)\|_2, \qquad 0 \leqq t \leqq 1. \tag{5.8}$$

Obviously $\theta(t)$ is also a continuous function on $[0,1]$ with $\theta(0)=0$, and it follows from (1.12) that

$$\theta(t)=\sin^{-1}\|\sin\Theta_1(t)\|_2, \qquad \Theta_1(t)\equiv\arccos(Z_1(t)^H\tilde{Z}_1\tilde{Z}_1^H Z_1(t))^{\frac{1}{2}}\geqq 0. \tag{5.9}$$

Since the hypothesis (5.2) implies $\tilde{P}(1)=Z_1 Z_1^H$, we have

$$\theta(1)=\sin^{-1}\|Z_1Z_1^H-\tilde{Z}_1\tilde{Z}_1^H\|_2=\sin^{-1}\|\sin\Theta_1\|_2=\|\Theta_1\|_2,$$

and thus

$$\theta(1)\,I\geqq\Theta_1. \tag{5.10}$$

3. Now we prove

$$\theta(1)<\frac{\pi}{4}-\frac{\varphi}{2}. \tag{5.11}$$

Because of the continuity of $\theta(t)$ and $\theta(0)=0$, there is a number $t\in[0,1)$, such that $\theta(t)\leqq\frac{\pi}{4}-\frac{\varphi}{2}$. Using Theorem 4.1 for $\Theta_1(t)$ (see (5.9)), we obtain

$$\begin{aligned}\sin 2\theta(t)&\leqq\|\sin 2\Theta_1(t)\|_2\\&\leqq\frac{2\xi}{1-\xi^2}\|\sin\Theta_1(t)\|_2^2+\frac{r\varepsilon}{\delta}\\&=\frac{2\xi}{1-\xi^2}\sin^2\theta(t)+\frac{r\varepsilon}{\delta}\leqq\frac{\xi}{1-\xi^2}+\frac{r\varepsilon}{\delta},\end{aligned}$$

and thus

$$\begin{aligned}\theta(t)&\leqq\frac{1}{2}\sin^{-1}\left(\frac{\xi}{1-\xi^2}+\frac{r\varepsilon}{\delta}\right)\leqq\frac{\pi}{4}\left(\frac{\xi}{1-\xi^2}+\frac{r\varepsilon}{\delta}\right)\\&<\frac{\pi}{4}\left(1-\frac{2\varphi}{\pi}\right)=\frac{\pi}{4}-\frac{\varphi}{2}.\end{aligned}$$

Utilizing the continuity of $\theta(t)$ we obtain (5.11). The proof is complete.

Theorem 5.2. *Assume that the hypotheses in Theorem* 4.1 *and Theorem* 5.1 *are valid, and let* $\theta_1=\|\Theta_1\|_2$ *(i.e.* $\theta(1)$*). Then*

$$\sin(2\theta_1+\varphi)\leqq r\cos\varphi\cdot\frac{\varepsilon}{\delta}+\sin\varphi. \tag{5.12}$$

Proof. Let $\lambda(\Theta_1)=\{\Theta_i\}$. According to Theorem 5.1 we have

$$\frac{\pi}{4}\geqq\frac{\pi}{4}-\frac{\varphi}{2}>\theta_1\geqq\theta_2\geqq\theta_2\geqq\ldots\geqq 0,$$

and thus

$$\frac{\sqrt{2}}{2}>\sin\theta_1\geqq\sin\theta_2\geqq\ldots\geqq 0, \qquad 1>\sin 2\theta_1\geqq\sin 2\theta_2\geqq\ldots\geqq 0.$$

Therefore

$$\|\sin 2\Theta_1\|_2 - \frac{2\xi}{1-\xi^2}\|\sin\Theta_1\|_2^2 = 2\sin\theta_1\left(\cos\theta_1 - \frac{\xi}{1-\xi^2}\sin\theta_1\right)$$
$$= \sec\varphi[\sin(2\theta_1+\varphi) - \sin\varphi].$$

Substituting this relation into (4.2) (on the spectral norm), we obtain (5.12) at once. The proof is complete.

§6. Final Remarks

6.1. In our notation we can state the famous Davis-Kahan $\sin\theta$ Theorem [2] as follows:

Suppose that A and $\tilde{A} = A + E$ are $n\times n$ Hermitian matrices, $Z = (Z_1, Z_2)$ and $\tilde{Z} = (\tilde{Z}_1, \tilde{Z}_2)$ are $n\times n$ unitary matrices such that

$$Z^H A Z = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad \tilde{Z}^H \tilde{A} \tilde{Z} = \begin{pmatrix} \tilde{A}_1 & 0 \\ 0 & \tilde{A}_2 \end{pmatrix}, \tag{6.1.1}$$

where Z_1 and $\tilde{Z}_1 \in \mathbb{C}^{n\times\ell}$, $0<\ell<n$, A_1 and $\tilde{A}_1 \in \mathbb{C}^{\ell\times\ell}$. Let $\lambda(A_1) = \{\alpha_i\}_{i=1}^{\ell}$ and $\lambda(\tilde{A}_2) = \{\tilde{\alpha}_j\}_{j=\ell+1}^{n}$ be the eigenvalues of A_1 and $\tilde{A}_2$ respectively. Assume that there is an interval $[\beta_0, \alpha_0]$ and a number $\delta_0 > 0$ such that $\lambda(A_1) \subseteq [\beta_0, \alpha_0]$ and $\lambda(\tilde{A}_2) \subseteq (-\infty, \beta_0 - \delta_0] \cup [\alpha_0 + \delta_0, +\infty)$ (or vice-versa). Then for every unitary-invariant norm,

$$\|\sin\Theta_1\| \leqq \frac{\|EZ_1\|}{\delta_0}, \tag{6.1.2}$$

where Θ_1 is defined by (1.7).

Without loss of generality we can assume that $0 \leqq \alpha_0 = -\beta_0$ in the hypotheses of the $\sin\theta$ theorem, because the translation of the spectrum of A_1 and $\tilde{A}_2$ $\left(\text{by translating } A \to A - \frac{\alpha_0+\beta_0}{2} I \text{ and } \tilde{A} \to \tilde{A} - \frac{\alpha_0+\beta_0}{2} I\right)$ do not effect EZ_1.

It is worth-while to point out that the above mentioned $\sin\theta$ theorem can be deduced from Theorem 2.1. We consider (A, rI) and $(\tilde{A}, rI) \in \mathbb{D}(n)$ for $r>0$, $\tilde{A} = A + E$ and $F = 0$. Obviously

$$(Z^H A Z, Z^H(rI)Z) = \left(\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \begin{pmatrix} rI & 0 \\ 0 & rI \end{pmatrix}\right),$$

$$(\tilde{Z}^H \tilde{A} \tilde{Z}, \tilde{Z}^H(rI)\tilde{Z}) = \left(\begin{pmatrix} \tilde{A}_1 & 0 \\ 0 & A_2 \end{pmatrix}, \begin{pmatrix} rI & 0 \\ 0 & rI \end{pmatrix}\right),$$

$\lambda(A_1, rI) = \{(\alpha_i, r)\}_{i=1}^{\ell}$ and $\lambda(\tilde{A}_2, rI) = \{(\tilde{\alpha}_j, r)\}_{j=\ell+1}^{n}$. According to the hypotheses we have

$$\rho((0,1),(\alpha_i, r)) = \frac{\left|\frac{\alpha_i}{r}\right|}{\sqrt{1+\left(\frac{\alpha_i}{r}\right)^2}} \leqq \frac{\frac{\alpha_0}{r}}{\sqrt{1+\left(\frac{\alpha_0}{r}\right)^2}}, \qquad 1 \leqq i \leqq \ell$$

and

$$\rho((0,1),(\tilde{\alpha}_j,r))=\frac{1}{\sqrt{1+\left(\frac{r}{\tilde{\alpha}_j}\right)^2}}\geq\frac{\frac{\alpha_0+\delta_0}{r}}{\sqrt{1+\left(\frac{\alpha_0+\delta_0}{r}\right)^2}},\qquad \ell+1\leqq j\leqq n;$$

here

$$\frac{\frac{\alpha_0}{r}}{\sqrt{1+\left(\frac{\alpha_0}{r}\right)^2}}=\alpha \quad\text{and}\quad \frac{\frac{\alpha_0+\delta_0}{r}}{\sqrt{1+\left(\frac{\alpha_0+\delta_0}{r}\right)^2}}-\frac{\frac{\alpha_0}{r}}{\sqrt{1+\left(\frac{\alpha_0}{r}\right)^2}}=\delta$$

satisfy the hypotheses of Theorem 2.1: $\alpha\geqq 0$, $\delta>0$ and $\alpha+\delta\leqq 1$. Hence if we set

$$t_1(r)=\sqrt{1+\left(\frac{\alpha_0}{r}\right)^2},\quad t_2(r)=\sqrt{1+\left(\frac{\alpha_0+\delta_0}{r}\right)^2},\quad t_3(r)=\sqrt{1+\left(\frac{\|A\|_2}{r}\right)^2}$$

and

$$S_1(r)=\inf_{\|x\|=1}\left\{\sqrt{1+\left(\frac{x^HAx}{r}\right)^2}\right\},\qquad S_2(r)=\inf_{\|x\|=1}\left\{\sqrt{1+\left(\frac{x^H\tilde{A}x}{r}\right)^2}\right\},$$

then from (2.3) and after some calculations, for every unitary-invariant matrix norm, we have

$$\|\sin\Theta_1\|\leqq\frac{(2\alpha_0+\delta_0)\,t_1(r)\,t_2(r)\,t_3(r)\,\|EZ_1\|}{S_1(r)\,S_2(r)\,[(\alpha_0+\delta_0)\,t_1(r)+\alpha_0 t_2(r)]\,[(\alpha_0+\delta_0)\,t_1(r)-\alpha_0 t_2(r)]},$$

which gives the inequality (6.1.2) when $r\to+\infty$.

6.2. Similarly, from Theorem 3.1 and Corollary 4.1 by a limiting procedure, we can derive the generalized $\sin\theta$ theorem (see (6.2.1)) and the $\sin 2\theta$ theorem (see (6.2.2)) for Hermitian matrices which are due to Davis and Kahan [2]:

Suppose that A, $\begin{pmatrix}A_1 & 0\\ 0 & A_2\end{pmatrix}$, $\tilde{A}$, $\begin{pmatrix}\tilde{A}_1 & 0\\ 0 & \tilde{A}_2\end{pmatrix}$, Z and $\tilde{Z}$ are the same as in 6.1. If

$$\delta\equiv\min|\lambda(A_1)-\lambda(\tilde{A}_2)|>0,$$

then we have

$$\|\sin\Theta_1\|_F\leqq\frac{\|EZ_1\|_F}{\delta}; \tag{6.2.1}$$

Moreover, if there is an interval $[\beta,\alpha]$ and a $\delta>0$ such that $\lambda(\tilde{A}_1)\subset[\beta,\alpha]$ and $\lambda(\tilde{A}_2)\subset(-\infty,\beta-\delta]\cup[\alpha+\delta,+\infty)$ (or vice-versa), then for every unitary-invariant matrix norm we have

$$\|\sin 2\Theta_1\|\leqq\frac{2\,\|E\|}{\delta}. \tag{6.2.2}$$

6.3. Finally we notice that the upper bound in (3.1) is independent of the dimension ℓ of the subspace $\mathscr{R}(Z_1)$, but the upper bound obtained by G.W. Stewart (see [7], Theorem 4.4 and Corollary 4.5) contains a factor $\sqrt{\ell}$.

Acknowledgement. I would like to thank Professor F.L. Bauer, Professor Chandler Davis and Professor G.W. Stewart for their encouragement and Professor L. Elsner for his support. I am grateful for the assistance of the Alexander von Humboldt Foundation in the Federal Republic of Germany.

References

1. Crawford, C.R.: A stable generalized eigenvalue problem. SIAM J. Numer. Anal. **6**, 854–860 (1976)
2. Davis, C., Kahan, W.M.: The rotation of eigenvectors by a perturbation. III. SIAM J. Numer. Anal. **7**, 1–46 (1970)
3. Elsner, L., Sun, J.G.: Perturbation theorems for the generalized eigenvalue problem. Linear Algebra Appl. (to appear)
4. Marcus, M., Minc, H.: A Survey of Matrix Theory and Matrix Inequalities. Boston: Allyn and Bacon, 1964
5. Stewart, G.W.: Error and perturbation bounds for subspaces associated with certain eigenvalue problems. SIAM Rev. **15**, 727–764 (1973)
6. Stewart, G.W.: On the perturbation of pseudo-inverses, projections and linear least squares problems. SIAM Rev. **19**, 634–662 (1977)
7. Stewart, G.W.: Perturbation bounds for the definite generalized eigenvalue problem. Linear Algebra Appl. **23**, 69–85 (1979)
8. Sun, J.G.: Invariant subspaces and generalized invariant subspaces. II. Perturbation theorems. Math. Numer. Sinica **2**, 113–123 (1980)
9. Sun, J.G.: A note on Stewart's theorem for definite matrix-pairs. Linear Algebra Appl. (to appear)
10. Uhlig, F.: A recurring theorem about pairs of quadratic forms and extensions: A survey. Linear Algebra Appl. **25**, 219–237 (1979)
11. Van Loan, C.: Generalizing the singular value decomposition. SIAM J. Numer. Anal. **13**, 76–83 (1976)

Received August 7, 1981/June 30, 1982