THE SPECTRAL VARIATION OF PENCILS OF MATRICES*

L. ELSNER
(*University of Bielefeld, Bielefeld, F. R. Germany*)

P. LANCASTER
(*University of Calgary, Calgary, Alberta, Canada*)

Abstract

Perturbation theorems for the spectrum of a regular matrix pencil $\lambda A-B$ are given. As it may include points near or at infinity the Euclidean distance is not appropriate. We use the chordal metric and the distances $w(\lambda, \tilde{\lambda})=\min\{|\lambda-\tilde{\lambda}|, |\lambda^{-1}-\tilde{\lambda}^{-1}|\}$ and $v(\lambda, \tilde{\lambda})=\{|\lambda-\tilde{\lambda}|$ if $|\lambda|\leqslant 1$ and $|\lambda^{-1}-\tilde{\lambda}^{-1}|$ if $|\lambda|>1\}$. For those purposes we develop here an algebraic treatment of matrix pairs, with special reference to diagonable and definite pairs, using ideas from the theory of matrix polynomials.

§ 1. Introduction

Throughout this paper (A, B) will denote a regular pair of $n\times n$ complex matrices. That is $\det(\lambda A-B)\not\equiv 0$, where λ is a complex parameter. Thus, there is a discrete set of complex numbers, the eigenvalues of the pair, for which $\det(\lambda A-B)=0$. Denote this set by $\sigma(\lambda A-B)$, the spectrum of the pencil $\lambda A-B$, and note that it may include the point at infinity.

We are concerned with a perturbation problem: To find bounds for the variation in the eigenvalues when (A, B) is perturbed to $(A+E, B+F)$ in terms of norms of E and F. It is well recognized that, when A is singular, or "nearly" so, the Euclidean metric is not appropriate for measuring the eigenvalue variations. This has led to investigations in terms of a homogeneous problem: Consider the set of complex pairs λ, μ for which $\det(\lambda A-\mu B)=0$ and measure the distance between pairs in terms of the chordal metric, ρ. This is because the chordal metric has the homogeneity property:

$$\rho((k\lambda, k\mu), (\alpha, \beta))=\rho((\lambda, \mu), (\alpha, \beta)).$$

This formulation also has the merit of treating A and B in a symmetrical way. This line of attack has been studied by Stewart[7], Elsner and Sun[2], et al.

The investigations of this paper are based on a rather different idea. The chordal metric suggests that an eigenvalue λ is better viewed as a representative $(\lambda, 1)$ of a class of equivalent number pairs for the eigenvalue problem in homogeneous form. Since the difficulties of perturbation theory arise when $|\lambda|$ is large compared to 1, we propose that, when $|\lambda|>1$ we consider the "reversed" eigenvalue pair $(1, \lambda^{-1})$. In this way we retain the symmetrical treatment of A and B and avoid

* Received January 22, 1985.

some of the complications associated with working with the chordal metric.

When considering perturbation of an eigenvalue λ of (A, B) with $|\lambda|\leqslant 1$ it is convenient to measure the distance from $\sigma(\lambda\tilde{A}-\tilde{B})$ by also writing all eigenvalues of $(\tilde{A}, \tilde{B})$ in the direct form $(\tilde{\lambda}, 1)$. Similarly, if $|\lambda|>1$, $\tilde{\lambda}$ is written in the reverse form $(1, \tilde{\lambda}^{-1})$ for all $\tilde{\lambda}\in\sigma(\lambda\tilde{A}-\tilde{B})$. Thus, we are led to the measure of distance

$$v(\lambda, \tilde{\lambda})=\begin{cases}|\lambda-\tilde{\lambda}| & \text{if } |\lambda|\leqslant 1,\\ |\lambda^{-1}-\tilde{\lambda}^{-1}| & \text{if } |\lambda|>1.\end{cases} \tag{1.1}$$

Note that v is not a metric. For example, $v(\lambda, \tilde{\lambda})\neq v(\tilde{\lambda}, \lambda)$, in general. Nevertheless, it is a useful and convenient measure in this context.

We shall also employ the following measures of distance (the latter is the chordal metric and will admit comparisons with the analysis of [2] and [7], for example):

$$w(\lambda, \tilde{\lambda})=\min(|\lambda-\tilde{\lambda}|, |\lambda^{-1}-\tilde{\lambda}^{-1}|), \tag{1.2}$$

$$\rho(\lambda, \tilde{\lambda})=\frac{|\lambda-\tilde{\lambda}|}{(1+|\lambda|^2)^{1/2}(1+|\tilde{\lambda}|^2)^{1/2}}. \tag{1.3}$$

Note that $w(\lambda, \tilde{\lambda})=w(\lambda^{-1}, \tilde{\lambda}^{-1})$ and $\rho(\lambda, \tilde{\lambda})=\rho(\lambda^{-1}, \tilde{\lambda}^{-1})$. Also, these measures of distance are related as follows:

$$\rho\leqslant w\leqslant v, \quad w\leqslant\frac{2\rho}{1-\rho} \tag{1.4}$$

and, if $w<1$, then $v\leqslant w/(1-w)$.

Let (A, B) and $(\tilde{A}, \tilde{B})$ be regular pairs of $n\times n$ matrices with (possibly infinite) eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ and $\tilde{\lambda}_1, \cdots, \tilde{\lambda}_n$, respectively. A spectral variation is defined for these pairs in terms of v, w or ρ by

$$S_{(A,B)}(\tilde{A}, \tilde{B})=\max_j \min_i v(\tilde{\lambda}_j, \lambda_i), \tag{1.5}$$

$$S^{(1)}_{(A,B)}(\tilde{A}, \tilde{B})=\max_j \min_i w(\tilde{\lambda}_j, \lambda_i), \tag{1.6}$$

$$S^{(2)}_{(A,B)}(\tilde{A}, \tilde{B})=\max_j \min_i \rho(\tilde{\lambda}_j, \lambda_i). \tag{1.7}$$

The primary objective of this paper is to obtain bounds for these spectral variations in terms of $\|\tilde{A}-A\|$ and $\|\tilde{B}-B\|$. This is achieved in Theorem 3.3 for diagonable pairs and in Theorem 5.3 for definite hermitian pairs. In Theorem 5.2 we have a result for a more general class of hermitian pairs. In the case of $S^{(2)}$ our results give some improvement on the results of Elsner and Sun[2]. The contributions of this paper also include improved proofs, and (possibly more convenient) measures S and $S^{(1)}$ of the spectral variation, and the more general result on hermitian pairs just cited.

The crux of our analysis is the division of the eigenvalues of (A, B) into those which are "small" and "large" in an appropriate sense. This is suggestive of an algebraic analysis of matrix polynomials presented by Gohberg, Lancaster and Rodman in Chapter 7 of [3]. Taking advantage of these ideas we need, and develop here, an algebraic treatment of matrix pairs, with special reference to diagonable and definite pairs; a treatment that seems to be missing in the literature on this problem area.

§ 2. The Algebraic Structure of Regular Pencils

The concepts and results of this section are closely related to the development of Chapter 7 of [3]. Here, they are tailored to the specific needs of this paper.

Let $A, B\in \mathbb{C}^{n\times n}$ and assume that the pencil $\lambda A-B$ is regular, i.e. $\det(\lambda A-B)\neq 0$. A pair (X, T) is a decomposable pair for $\lambda A-B$ (with parameter m) if the following three conditions are satisfied:

$$T\in \mathbb{C}^{n\times n} \text{ and for some integer } m\leqslant n,\ T=T_1\oplus T_2, \text{ where } T_1\in \mathbb{C}^{m\times m}. \tag{2.1}$$

$$X\in \mathbb{C}^{n\times n},\quad \det X\neq 0. \tag{2.2}$$

If $X=[X_1, X_2]$, where $X_1\in \mathbb{C}^{n\times m}$, then

$$AX_1T_1-BX_1=0,\quad AX_2-BX_2T_2=0. \tag{2.3}$$

The integer m is called the parameter of the pair (X, T).

For any matrix $M\in \mathbb{C}^{n\times n}$, let $r(M)$ denote the spectral radius of M. Thus, $r(M)=\max\limits_{\lambda\in\sigma}(|\lambda|)$, where $\sigma=\sigma(\lambda I-M)$.

The pair (X, T) is called strictly decomposable, if it is decomposable and also

$$\rho_1<\rho_2^{-1}, \tag{2.4}$$

where $\rho_i=r(T_i)$, $i=1, 2$.

In our applications m will likely be chosen so that $\rho_1=1$. To avoid trivialities we generally assume tacitly that $0<m<n$. For the existence of decomposable pairs we refer to Chapter 7 of [3].

Given a decomposable pair (X, T) we define another matrix pencil by writing

$$T(\lambda)=\begin{bmatrix}\lambda I_m-T_1 & 0\\ 0 & \lambda T_2-I_{n-m}\end{bmatrix}. \tag{2.5}$$

The following lemma shows that, in particular, $\lambda A-B$ and $T(\lambda)$ are strictly equivalent.

Lemma 2.1. *If (X, T) is a decomposable pair for the regular pencil $\lambda A-B$, then the matrix $V=[AX_1, BX_2]$ is nonsingular. If*

$$Y=V^{-1}=\begin{bmatrix}Y_1\\ Y_2\end{bmatrix}, \tag{2.6}$$

where $Y_1\in \mathbb{C}^{m\times n}$, then for all $\lambda\in \mathbb{C}$

$$Y(\lambda A-B)X=T(\lambda) \tag{2.7}$$

and, for $\lambda\notin\sigma(\lambda A-B)$

$$(\lambda A-B)^{-1}=X_1(\lambda I-T_1)^{-1}Y_1+X_2(\lambda T_2-I)^{-1}Y_2. \tag{2.8}$$

Proof. Observe that

$$VT(\lambda)=[AX_1, BX_2]\begin{bmatrix}\lambda I-T_1 & 0\\ 0 & \lambda T_2-I\end{bmatrix}=[\lambda AX_1-AX_1T_1, \lambda BX_2T_2-BX_2].$$

Using (2.3) we find

$$VT(\lambda)=[(\lambda A-B)X_1, (\lambda A-B)X_2]=(\lambda A-B)X.$$

Taking determinants we see $\det V\neq 0$. Multiplying on the left by $Y=V^{-1}$ gives

(2.7). Equation (2.8) follows from the rearrangement $(\lambda A-B)^{-1}=XT(\lambda)^{-1}Y$. ∎

The triple (X, T, Y) with Y defined as in Lemma 2.1 is called a decomposable triple for $\lambda A-B$. There is a dual representation for X in terms of Y.

Lemma 2.2. *If (X, T, Y) is a decomposable triple for the regular pair $\lambda A-B$ then*

$$X=\begin{bmatrix} Y_1A \\ Y_2B \end{bmatrix}^{-1}. \tag{2.9}$$

Proof. Comparing the coefficients of λ in the first m rows of (2.7) gives

$$Y_1AX=[I_m, 0],$$

and in the last $n-m$ rows the constant terms give

$$Y_2BX=[0, I_{n-m}].$$

Hence

$$\begin{bmatrix} Y_1A \\ Y_2B \end{bmatrix}X=I_n. ∎$$

Observe that we can just as well start out by defining decomposable triples (X, T, Y) as triples of matrices in $\mathbb{C}^{n\times n}$ such that (2.1) and (2.7) hold, where $T(\lambda)$ is defined as in (2.5). Then (2.2), (2.3), (2.6) and (2.9) follow.

Two pairs of matrices (X, T) and $(\tilde{X}, \tilde{T})$ satisfying (2.1) and (2.2) with the same parameter m are said to be similar if there are nonsingular matrices R_1, R_2 (of sizes m and $n-m$, respectively), such that

$$\begin{cases} X_1=\tilde{X}_1R_1, & T_1=R_1^{-1}\tilde{T}_1R_1, \\ X_2=\tilde{X}_2R_2, & T_2=R_2^{-1}\tilde{T}_2R_2. \end{cases} \tag{2.10}$$

It is easily verified that if (X, T) is a decomposable pair then every matrix pair similar to (X, T) is also a decomposable pair. A converse statement is true for strictly decomposable pairs.

Theorem 2.3. *Any two strictly decomposable pairs with the same parameter are similar and the transforming matrices defining this similarity are unique.*

Proof. Let (X, T) and $(\tilde{X}, \tilde{T})$ be strictly decomposable pairs for $\lambda A-B$ with the same parameter. Thus, we may write

$$X=[X_1, X_2], \quad T=T_1\oplus T_2, \quad \tilde{X}=[\tilde{X}_1, \tilde{X}_2], \quad \tilde{T}=\tilde{T}_1\oplus\tilde{T}_2.$$

Let Y, $\tilde{Y}$ be defined as in Lemma 2.1. Then from (2.7)

$$Y(\lambda A-B)X=\begin{bmatrix} \lambda I-T_1 & 0 \\ 0 & \lambda T_2-I \end{bmatrix}, \quad \tilde{Y}(\lambda A-B)\tilde{X}=\begin{bmatrix} \lambda I-\tilde{T}_1 & 0 \\ 0 & \lambda\tilde{T}_2-I \end{bmatrix}. \tag{2.11}$$

Choose α such that $\hat{B}=B-\alpha A$ is invertible. Then it follows from (2.11) that

$$\begin{aligned} \hat{B}^{-1}A &= X\begin{bmatrix} (T_1-\alpha I)^{-1} & 0 \\ 0 & (I-\alpha T_2)^{-1} \end{bmatrix} YY^{-1}\begin{bmatrix} I & 0 \\ 0 & T_2 \end{bmatrix}X^{-1} \\ &= X\begin{bmatrix} (T_1-\alpha I)^{-1} & 0 \\ 0 & (I-\alpha T_2)^{-1}T_2 \end{bmatrix}X^{-1} \end{aligned} \tag{2.12}$$

and similarly

$$\hat{B}^{-1}A=\tilde{X}\begin{bmatrix}(\tilde{T}_1-\alpha I)^{-1} & 0\\ 0 & (I-\alpha\tilde{T}_2)^{-1}\tilde{T}_2\end{bmatrix}\tilde{X}^{-1}. \tag{2.13}$$

This shows that $(T_1-\alpha I)^{-1}\oplus(I-\alpha T_2)^{-1}T_2$ and $(\tilde{T}_1-\alpha I)\oplus(I-\alpha\tilde{T}_2)^{-1}\tilde{T}_2$ are similar. We may assume that $r(T_1)\geqslant r(\tilde{T}_1)$, and claim now that $\sigma((I-\alpha T_2)^{-1}T_2)\cap\sigma((\tilde{T}_1-\alpha I)^{-1})=\phi$. Otherwise there is $\mu_1\in\sigma(\tilde{T}_1)$, $\mu_2\in\sigma(T_2)$ such that $\mu_2/(1-\alpha\mu_2)=(\mu_1-\alpha)^{-1}$, which implies $\mu_1\mu_2=1$. This contradicts $r(\tilde{T}_1)\leqslant r(T_1)<r(T_2)^{-1}$.

We may now deduce that $(\tilde{T}_1-\alpha I)^{-1}$ and $(T_1-\alpha I)^{-1}$ are similar and so are $(I-\alpha T_2)^{-1}T_2$ and $(I-\alpha\tilde{T}_2)^{-1}\tilde{T}_2$. Equations (2.12) and (2.13) now tell us that the columns of X_1 and $\tilde{X}_1$ are both bases of the invariant subspace of $\hat{B}^{-1}A$ belonging to the spectrum of $(T_1-\alpha I)^{-1}$. Hence there is a unique $R_1\in\mathbb{C}^{m\times m}$ such that $X_1=\tilde{X}_1R_1$ and similarly there is a unique R_2 with $X_2=\tilde{X}_2R_2$. From Theorem 2.3 we infer $AX_1R_1^{-1}\tilde{T}_1R_1=BX_1=AX_1T_1$. This implies $T_1=R_1^{-1}\tilde{T}_1R_1$, as (by Lemma 2.1) AX_1 is of full rank. Similarly $T_2=R_2^{-1}\tilde{T}_2R_2$. ∎

The idea of similarity of pairs carries over to triples in a natural way. Two decomposable triples (X, T, Y), $(\tilde{X}, \tilde{T}, \tilde{Y})$ for the same regular pencil $\lambda A-B$ are said to be similar if there are nonsingular matrices R_1, R_2 such that

$$X_i=\tilde{X}_iR_i,\quad T_i=R_i^{-1}\tilde{T}_iR_i,\quad Y_i=R_i^{-1}\tilde{Y}_i,\quad i=1,2.$$

It is clear from the definition of Y in Lemma 2.1 that there is an exact analogue of Theorem 2.3 for triples.

Corollary 2.4. Any two strictly decomposable triples with the same parameter are similar and the transforming matrices defining this similarity are unique.

The next corollary follows on making an appropriate choice of parameter m followed by determination of a decomposable triple (X, T, Y), and reduction of T_1 and T_2 to Jordan form by choice of R_1 and R_2.

Corollary 2.5. For a given $\varkappa>0$ there exists a strictly decomposable triple for the regular pencil $\lambda A-B$ of the form

$$[X_1, X_2],\ \begin{bmatrix}J_0 & 0\\ 0 & J_\infty\end{bmatrix},\ \begin{bmatrix}Y_1\\ Y_2\end{bmatrix}, \tag{2.14}$$

where J_0 and J_∞ are Jordan matrices such that

$$r(J_0)\leqslant\varkappa<r(J_\infty)^{-1}.$$

Here, the constant $\varkappa$ is chosen to divide the "small" from the "large" eigenvalues of $\lambda A-B$. The spectral information concerning the small and large eigenvalues is summarized in the triples (X_1, J_0, Y_1) and (X_2, J_∞, Y_2), respectively. The eigenvalues of J_∞ are, of course, the reciprocals of eigenvalues of $\lambda A-B$. In particular, J_∞ has a zero eigenvalue if and only if $\lambda A-B$ has an infinite eigenvalue.

Note that (2.7) implies

$$YAX=I_m\oplus J_\infty,\quad YBX=J_0\oplus I_{n-m} \tag{2.15}$$

and that the parameter m might be equal to 0 or n, so that some of the matrices in (2.14) do not appear. This can only happen if either B is nonsingular and $r(B^{-1}A)\leqslant\varkappa^{-1}$ or A is nonsingular and $r(A^{-1}B)\leqslant\varkappa$.

§ 3. Spectral Variation of Diagonable Pairs

A regular pair (A, B) is said to be diagonable if the matrices J_0 and J_∞ of Corollary 2.5 are diagonal. This property is clearly independent of the choice of $\varkappa$. Equivalent statements are that the pencil $\lambda A-B$ has only linear elementary divisors, or that $\lambda A-B$ has a complete set of eigenvectors. We now use the structure developed in section 2 in an analysis of the spectral variation for diagonable pairs. For bounding the spectral variation we need the following lemma. Here, and throughout this paper we use the euclidean vector norm and the corresponding induced, (or operator) norm on matrices. We also use corresponding lower bounds for a matrix as developed in [6], for example.

Lemma 3.1. *Let (A, B) be diagonable, and Y, X be any nonsingular matrices such that*

$$\begin{cases} YAX=\Lambda=\mathrm{diag}[\alpha_1, \cdots, \alpha_n], \\ YBX=\Omega=\mathrm{diag}[\beta_1, \cdots, \beta_n]. \end{cases} \tag{3.1}$$

If $\tilde\lambda$ is an eigenvalue of the pair $\tilde A$, $\tilde B$, where $\tilde A=A+E$, $\tilde B=B+F$ then

$$\min_i |\tilde\lambda\alpha_i-\beta_i|\leqslant\|X\|\|Y\|\|[E, F]\|(1+|\tilde\lambda|^2)^{1/2}. \tag{3.2}$$

Proof. Let $(\tilde\lambda\tilde A-\tilde B)x=0$ with $\|x\|=1$. Then

$$(\tilde\lambda A-B)x=(A-\tilde A)\tilde\lambda x-(B-\tilde B)x=-[E, F]\begin{bmatrix}\tilde\lambda x\\ -x\end{bmatrix},$$

whence

$$\|(\tilde\lambda A-B)x\|\leqslant\|[E, F]\|(1+|\tilde\lambda|^2)^{1/2}.$$

From (3.1) $\tilde\lambda A-B=Y^{-1}(\tilde\lambda\Lambda-\Omega)X^{-1}$ and so

$$\|Y\|^{-1}glb(\tilde\lambda\Lambda-\Omega)\|X\|^{-1}\leqslant\|[E, F]\|(1+|\tilde\lambda|^2)^{1/2},$$

which implies (3.2).∎

Now we may use Corollary 2.5 to be more specific about the diagonal matrices Λ and Ω. Choose $\varkappa=1$ in that corollary and we find that there is a strictly decomposable triple (X, J, Y) with the form (2.14) and for which

$$r(J_0)\leqslant1<r(J_\infty)^{-1}, \tag{3.3}$$

$$YAX=I_m\oplus J_\infty,\quad YBX=J_0\oplus I_{n-m}. \tag{3.4}$$

Thus in (3.1) we may take $\Lambda=I_m\oplus J_\infty$, $\Omega=J_0\oplus I_{n-m}$ and observe that, in view of (3.3),

$$\|\Lambda\|=\|\Omega\|=1. \tag{3.5}$$

In our estimation of spectral variations we shall also need the following bound for the size of the left eigenvectors.

Lemma 3.2. *Let (A, B) be a diagonable pair and let X, J, Y be a strictly decomposable triple for $\lambda A-B$ where $J=J_0\oplus J_\infty$, as in Corollary 2.5, with $\varkappa=1$. Then*

$$\|Y\|\leqslant\sqrt{2}\|X^{-1}\|\|K^{-1}\|^{1/2}, \tag{3.6}$$

where $K=AA^*+BB^*$.

Proof. Observe first that, since (A, B) is regular, the matrix $[A, B]$ has rank n, and hence K is nonsingular. From equations (3.4)

$$YKY^* = (YA)(YA)^* + (YB)(YB)^* = (\Lambda X^{-1})(\Lambda X^{-1})^* + (\Omega X^{-1})(\Omega X^{-1})^*.$$

Hence, for any $y \in \mathbb{C}^n$,

$$y^* YKY^* y = \|X^{*-1}\Lambda^* y\|^2 + \|X^{*-1}\Omega^* y\|^2,$$

and

$$glb(K)\|Y^* y\|^2 \leqslant \|X^{-1}\|^2 (\|\Lambda^* y\|^2 + \|\Omega^* y\|^2).$$

Using (3.5) and the fact that $(glb(K))^{-1} = \|K^{-1}\|$ we find that if $y \neq 0$,

$$\frac{\|Y^* y\|^2}{\|y\|^2} \leqslant 2\|X^{-1}\|^2 \|K^{-1}\|^2$$

and (3.6) follows.∎

Theorem 3.3. *Let (A, B) be a diagonable pair and let X, J be a strictly decomposable pair for $\lambda A - B$ where $J = J_0 \oplus J_\infty$, as in Corollary 2.5, with $\varkappa = 1$. Let $\tilde{A} = A + E$, $\tilde{B} = B + F$, and assume that*

$$\eta \triangleq 2\|X\|\|X^{-1}\|\|K^{-1}\|^{1/2}\|[E, F]\| < 1, \tag{3.7}$$

where $K = AA^* + BB^*$. Then

$$S_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \frac{\eta}{1-\eta}, \quad S^{(1)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \frac{\eta}{(1-\eta)^{1/2}}, \quad S^{(2)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \eta/2. \tag{3.8}$$

Recall that the spectral variations are defined in (1.4), (1.5) and (1.6).

Proof. Let (X, J) be the strictly decomposable pair which determines the triple of Lemma 3.2. Let $\tilde{\lambda}$ be an eigenvalue of $\lambda\tilde{A} - \tilde{B}$ with $|\tilde{\lambda}| \leqslant 1$. Combine equations (3.2) and (3.6) to get

$$\min_i |\tilde{\lambda}\alpha_i - \beta_i| \leqslant \eta, \tag{3.9}$$

where η is defined in (3.7) and

$$\mathrm{diag}[\alpha_1, \cdots, \alpha_n] = I_m \oplus J_\infty, \quad \mathrm{diag}[\beta_1, \cdots, \beta_n] = J_0 \oplus I_{n-m}. \tag{3.10}$$

Hence there is a j for which

$$|\tilde{\lambda} - \lambda_j| \leqslant \eta \quad \text{or} \quad |\tilde{\lambda}\mu_j - 1| \leqslant \eta$$

depending on whether the minimum of (3.9) is attained for $i \leqslant m$ or $i > m$. Here, $\lambda_1, \cdots, \lambda_m, \mu_{m+1}^{-1}, \cdots, \mu_n^{-1}$ are, of course, the eigenvalues of $\lambda A - B$. If $j > m$ then $1 - |\mu_j||\tilde{\lambda}| \leqslant \eta$. Hence

$$|\mu_j| \geqslant \frac{1-\eta}{|\tilde{\lambda}|} \geqslant 1 - \eta,$$

and

$$\left|\tilde{\lambda} - \frac{1}{\mu_j}\right| \leqslant \frac{\eta}{|\mu_j|} \leqslant \frac{\eta}{1-\eta}.$$

Thus, wherever the minimum is attained, the inequality $|\tilde{\lambda} - \lambda_j| \leqslant \frac{\eta}{1-\eta}$ holds for some eigenvalue λ_j of $\lambda A - B$.

If $|\tilde{\lambda}| > 1$ we view $\tilde{\mu} = \tilde{\lambda}^{-1}$ as an eigenvalue of the pencil $\tilde{A} - \mu\tilde{B}$: with $|\tilde{\mu}| < 1$. Since the right-hand side of (3.2) is symmetric in A and B, we have, as in (3.9),

$$\min_i |\alpha_i - \tilde{\mu}\beta_i| \leqslant \eta. \tag{3.11}$$

Then continue the argument as above and it is found that

$$|\tilde{\lambda}^{-1} - \lambda_j^{-1}| \leqslant \frac{\eta}{1-\eta}$$

for some eigenvalue λ_j of $\lambda A - B$.

The first conclusion of (3.8) now follows from the definitions (1.1) and (1.5).

We turn now to the chordal metric. Referring to (3.1), define

$$D = (\Lambda\Lambda^* + \Omega\Omega^*)^{-1/2}, \quad \hat{Y} = DY, \quad \hat{\Lambda} = D\Lambda = \operatorname{diag}[\hat{\alpha}_1, \cdots, \hat{\alpha}_n],$$
$$\hat{\Omega} = D\Omega = \operatorname{diag}[\hat{\beta}_1, \cdots, \hat{\beta}_n]$$

so that

$$\hat{\Lambda}\hat{\Lambda}^* + \hat{\Omega}\hat{\Omega}^* = I_n. \tag{3.12}$$

Going through the proof of Lemma 3.2, we obtain $glb(K)\|\hat{Y}^* y\|^2 \leqslant \|X^{-1}\|^2(\|\hat{\Lambda}^* y\|^2 + \|\hat{\Omega}^* y\|^2) = \|X^{-1}\|^2\|y\|^2$ by (3.12) and hence

$$\|\hat{Y}^*\| \leqslant \|X^{-1}\| \, \|K^{-1}\|^{1/2}.$$

From (3.2) we get

$$\min_i \frac{|\tilde{\lambda}\hat{\alpha}_i - \hat{\beta}_i|}{\sqrt{1+|\tilde{\lambda}|^2}} \leqslant \|X\| \, \|X^{-1}\| \, \|K^{-1}\|^{1/2} \|[E, F]\| = \eta/2.$$

As $|\hat{\alpha}_i|^2 + |\hat{\beta}_i|^2 = 1$, we find that the left-hand side is equal to $\rho(\tilde{\lambda}, \hat{\alpha}_i/\hat{\beta}_i)$ for a suitable i, hence

$$S^{(2)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \eta/2. \tag{3.13}$$

To obtain the inequality concerning the distance function w, we use the relation $w \leqslant 2\rho/(1-\rho)$ of (1.4) which, together with (3.13), implies

$$S^{(1)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \frac{\eta}{1-\eta/2} \leqslant \frac{\eta}{\sqrt{1-\eta}}. \blacksquare \tag{3.14}$$

We remark that (3.13) can also be obtained from Theorem 2.1 and (1.13) of [2], but our approach is more elementary and direct. Eq. (3.14) can also be proved starting out from (3.9).

§ 4. Hermitian Pencils; Algebraic Structure

The first results of this section are familiar in the context of more general hermitian matrix-valued functions. See Chapters II. 2 and II.3 of [4], for example. Nevertheless, a careful discussion in this context is worthwhile. We start with:

Theorem 4.1. *Let the regular pencil $\lambda A - B$ have a strictly decomposable triple*

$$[X_1, X_2], \quad \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}, \quad \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}.$$

Then the pencil $\lambda A^* - B^*$ has the strictly decomposable triple

$$[Y_1^*, Y_2^*], \quad \begin{bmatrix} T_1^* & 0 \\ 0 & T_2^* \end{bmatrix}, \quad \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix}.$$

Proof. The properties (2.1), (2.2), (2.4) are immediate. From (2.7) we obtain

$$0=Y(\lambda A-B)-T(\lambda)X^{-1}=Y(\lambda A-B)-T(\lambda)\begin{bmatrix} Y_1A \\ Y_2B \end{bmatrix}=\begin{bmatrix} T_1Y_1A-Y_1B \\ \lambda(Y_2A-T_2Y_2B) \end{bmatrix}.$$

Thus $T_1Y_1A=Y_1B$ and $Y_2A=T_2Y_2B$. Taking adjoints in these equations establishes property (2.3) for the pencil λA^*-B^*. ∎

Lemma 4.2. *Let $\lambda A-B$ be a regular pencil with $A=A^*$, $B=B^*$, and let*

$$[X_1,\ X_2],\ \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix},\ \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \tag{4.1}$$

be a strictly decomposable triple for $\lambda A-B$. Then there are unique nonsingular hermitian matrices H_1, H_2 such that

$$\begin{aligned} X_1&=Y_1^*H_1, \quad T_1=H_1^{-1}T_1^*H_1, \quad Y_1=H_1^{-1}X_1^*, \\ X_2&=Y_2^*H_2, \quad T_2=H_2^{-1}T_2^*H_2, \quad Y_2=H_2^{-1}X_2^*. \end{aligned} \tag{4.2}$$

Proof. Since $A^*=A$, $B^*=B$ we have $\lambda A-B=\lambda A^*-B^*$ for $\lambda\in\mathbb{R}$. By Theorem 4.1 we have a second strictly decomposable triple for $\lambda A-B$:

$$[Y_1^*,\ Y_2^*],\ \begin{bmatrix} T_1^* & 0 \\ 0 & T_2^* \end{bmatrix},\ \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix}.$$

Now use Corollary 2.4 to assert the existence of nonsingular matrices H_1, H_2 such that equations (4.2) hold. But taking adjoints in the system (4.2) it is found that the similarity can also be achieved with the transforming matrices H_1^* and H_2^*. Hence $H_1=H_1^*$, $H_2=H_2^*$. ∎

Equations (4.2) show that, for $j=1, 2$, T_j is H_j-selfadjoint. Hence, by Theorem 3.3 of [4] there are nonsingular matrices S_j such that

$$\begin{aligned} S_1^{-1}T_1S_1&=J_0, \quad S_1^*H_1S_1=P_0, \\ S_2^{-1}T_2S_2&=J_\infty, \quad S_2^*H_2S_2=P_\infty, \end{aligned} \tag{4.3}$$

where $(J_0,\ P_0)$ is a canonical pair for $(T_1,\ H_1)$ and includes a sign-characteristic associated with the part of the finite real spectrum of $\lambda A-B$ belonging to J_0, i.e. real eigenvalues λ with $|\lambda|<\rho_1=r(T_1)$. Also, $(J_\infty,\ P_\infty)$ is a canonical pair for $(T_2,\ H_2)$ and includes a sign-characteristic associated with the real eigenvalues of $\lambda A-B$ which exceed $\rho_2=r(T_2)$ in absolute value. Note, in particular, that if $\lambda A-B$ is diagonable, then P_0 and P_∞ are diagonal matrices with only $+1$ or -1 in the main diagonal positions.

Theorem 4.3. *Let $\lambda A-B$ be a regular pencil with $A^*=A$, $B^*=B$, and having a strictly decomposable triple of parameter m. Then there is a strictly decomposable triple for $\lambda A-B$ of the form*

$$[X_1,\ X_2],\ \begin{bmatrix} J_0 & 0 \\ 0 & J_\infty \end{bmatrix}\begin{bmatrix} P_0X_1^* \\ P_\infty X_2^* \end{bmatrix}, \tag{4.4}$$

where $(J_0,\ P_0)$ and $(J_\infty,\ P_\infty)$ are canonical pairs of sizes m and $n-m$.

Proof. Take the triple in (4.1) and transform it to the similar triple $(X_jS_j,\ S_j^{-1}T_jS_j,\ S_j^{-1}Y_j)$, $j=1, 2$ where S_1, S_2 are chosen to satisfy (4.3). Replacing X_jS_j by X_j yields (4.4). ∎

Corollary 4.4. *The triple described in* (4.4) *has the properties*

$$X^*AX = P_0 \oplus P_\infty J_\infty, \quad X^*BX = P_0 J_0 \oplus P_\infty. \tag{4.5}$$

Proof. In (2.15) we may now put

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} P_0 & 0 \\ 0 & P_\infty \end{bmatrix} \begin{bmatrix} X_1^* \\ X_2^* \end{bmatrix} = [P_0 \oplus P_\infty] X^*$$

and the result follows on noting that $P_0^2 = I$, $P_\infty^2 = I$. ∎

Note also that, using the triple (4.4), the resolvent form of equation (2.8) becomes

$$(\lambda A - B)^{-1} = X_1(\lambda I - J_0)^{-1} P_0 X_1^* + X_2(\lambda J_\infty - I)^{-1} P_\infty X_2^*. \tag{4.6}$$

The hermitian pair (A, B) is said to be definite if

$$c(A, B) = \inf_{\|x\|=1} \{|x^*(A + iB)x|\} > 0. \tag{4.7}$$

Clearly, this is equivalent to stating that $x^*Ax = x^*Bx = 0$ if and only if $x = 0$. This concept was first introduced by Crawford[1]. It is well known that a definite pair has only real eigenvalues and is diagonable. These properties are usually established by showing that there exists a positive definite linear combination of A and B (see the review of Uhlig[8], for example). The structure developed here allows a purely algebraic proof.

Proposition 4.5. If the pair (A, B) is definite then

(i) all eigenvalues of $\lambda A - B$ are real;

(ii) all elementary divisors are linear, i.e. all partial multiplicities are equal to one;

(iii) there is a nonsingular X such that

$$X^*AX = P_0 \oplus P_\infty J_\infty, \quad X^*BX = P_0 J_0 \oplus P_\infty, \tag{4.8}$$

where P_0, P_∞, J_0, J_∞ are diagonal.

Proof. Let (X, J, Y) be a decomposable triple with the properties (4.4) and (4.5). If there is a nonreal eigenvalue $\lambda_0 \in \sigma(J_0)$ then as the pair is hermitian, $\bar{\lambda}_0$ is also an eigenvalue of $\lambda A - B$ and, due to the strict decomposability, is in $\sigma(J_0)$. Hence there is a principal submatrix of J_0 of the form $\begin{bmatrix} \lambda_0 & 0 \\ 0 & \bar{\lambda}_0 \end{bmatrix}$, the corresponding submatrix of P_0 is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and hence of $P_0 J_0$ is $\begin{bmatrix} 0 & \bar{\lambda}_0 \\ \lambda_0 & 0 \end{bmatrix}$. Thus for a suitable unit vector e_j we have $e_j^* P_0 e_j = e_j^* P_0 J_0 e_j = 0$. For $x_j = X e_j$ we get from (4.5) $x_j^* A x_j = x_j^* B x_j = 0$, $x_j \neq 0$. A similar argument holds if $\lambda_0 \in \sigma(J_\infty)$ and is nonreal. This establishes (i). Also, a similar argument applies if λ_0 is real with a partial multiplicity larger than one.

Thus, J_0 and J_∞ are real diagonal matrices and it follows that P_0, P_∞ are also diagonal. Equation (4.8) is just (4.5). ∎

It is apparent from the definition 4.7 that the property of being definite is stable under small hermitian perturbations of A and B. We remark that this implies a further property of the canonical matrices P_0 and P_∞. In the terminology of [5] or [4], a definite pair may be said to be "stably real-diagonable" and it follows from Theorem III.1.3 of [4], for example, that when a multiple eigenvalue appears in J_0 or J_∞ the corresponding submatrix of P_0 or P_∞, respectively, is just I or $-I$.

To see that the real spectrum of more general hermitian pencils is not stable under hermitian perturbations consider the example

$$A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix},\quad B=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

and the perturbed pencil

$$\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\lambda-\begin{bmatrix}0 & \varepsilon\\ \varepsilon & 0\end{bmatrix},$$

where ε is real.

Earlier analyses of spectral variations for hermitian pairs have been confined to definite pairs, but with the preparations made so far we can apply our analysis to a hermitian pair (A, B) which is merely diagonable and has only real eigenvalues. Such a pair will be called (hermitian) r-diagonable. It should be emphasized that the perturbed system need not be hermitian.

§ 5. Spectral Variation for Hermitian Pairs

Note first of all that if $\lambda A-B$ is a hermitian pencil and (A, B) is r-diagonable then Theorem 3.3 applies with X and J chosen as in Theorem 4.3. We first note that, in this case, X^{-1} can be estimated in a potentially useful way.

Proposition 5.1. Let (A, B) be an r-diagonable pair. Then the matrix X of Theorem 4.3 (and Corollary 4.4) has the property

$$\|X^{-1}\|\leqslant\|X\|\,\|A^2+B^2\|^{1/2}\leqslant\|X\|(\|A\|+\|B\|). \tag{5.1}$$

Proof. Equations (4.5) imply that

$$(X^*AX)^2+(X^*BX)^2=I+D,$$

where $D=J_0^2\oplus J_\infty^2\geqslant 0$. Hence, for any $x\neq 0$

$$\begin{aligned}x^*X^{*-1}X^{-1}x\leqslant x^*X^{*-1}(I+D)X^{-1}x&=x^*(AXX^*A+BXX^*B)x\\ &=\|X^*Ax\|^2+\|X^*Bx\|^2\leqslant\|X\|^2(\|Ax\|^2+\|Bx\|^2)\\ &=\|X\|^2x^*(A^2+B^2)x\leqslant\|X\|^2\|A^2+B^2\|\,\|x\|^2\end{aligned}$$

and the first inequality of (5.1) follows. The second is elementary.∎

Before stating a perturbation theorem for r-diagonable pairs it is convenient to introduce the condition number of $(A^2+B^2)^{1/2}$. Thus, let

$$\varkappa=\|(A^2+B^2)^{1/2}\|\,\|(A^2+B^2)^{-1/2}\|.$$

Combining Proposition 5.1 with Theorem 3.3 yields:

Theorem 5.2. *Let (A, B) be a hermitian r-diagonable pair, and let a strictly decomposable triple for $\lambda A-B$ be formed as in Eq.* (4.4). *Assume that*

$$\eta_0\triangleq 2\varkappa\|X\|^2\|[E,F]\|<1.$$

Then the inequalities (3.8) *hold with η replaced by η_0.*

The spectral variation for definite pairs is usually estimated in terms of the constant $c(A, B)$ of Equation (4.7) and we conclude with the derivation of estimates of this kind. Combining the notations of Proposition 4.5 and Lemma 3.1 we have for a definite pair (A, B):

$$X^*AX = P_0 \oplus P_\infty J_\infty = \Lambda = \operatorname{diag}[\alpha_1, \cdots, \alpha_n],$$
$$X^*BX = P_0 J_0 \oplus P_\infty = \Omega = \operatorname{diag}[\beta_1, \cdots, \beta_n].$$

Lemma 3.1 implies that, for any eigenvalue $\tilde{\lambda}$ of the perturbed pair $\tilde{A} = A+E$, $\tilde{B} = B+F$,

$$\min_j |\tilde{\lambda}\alpha_j - \beta_j| \leqslant \|X\|^2 \|[E, F]\| (1+|\tilde{\lambda}|^2)^{1/2}. \tag{5.2}$$

We estimate $\|X\|^2$ following an argument of [2]. Thus, for any $x \neq 0$,

$$c(A, B) \leqslant \frac{|x^*X^*(A+iB)Xx|}{x^*X^*Xx} = \left[\frac{|x^*(\Lambda+i\Omega)x|}{x^*x}\right] \Big/ \left[\frac{x^*X^*Xx}{x^*x}\right].$$

Hence

$$\frac{\|Xx\|^2}{\|x\|^2} \leqslant \frac{1}{c(A, B)} \frac{|x^*(\Lambda+i\Omega)x|}{x^*x}.$$

With the decomposable pair (X, J) chosen to satisfy (3.3) and using the fact that the eigenvalues are real we have

$$\|X\|^2 \leqslant \frac{\sqrt{2}}{c(A, B)}. \tag{5.3}$$

If $|\tilde{\lambda}| \leqslant 1$ then (5.2) and (5.3) yield

$$\min_j |\tilde{\lambda}\alpha_j - \beta_j| \leqslant \frac{2}{c(A, B)} \|[E, F]\|. \tag{5.4}$$

If $|\tilde{\lambda}| > 1$ we write $\tilde{\mu} = \tilde{\lambda}^{-1}$ so that $|\tilde{\mu}| < 1$. As in (3.2),

$$\min_j |\alpha_j - \tilde{\mu}\beta_j| \leqslant \|X\|^2 \|[E, F]\| (1+|\tilde{\mu}|^2)^{1/2}$$

and so

$$\min_j |\alpha_j - \tilde{\mu}\beta_j| \leqslant \frac{2}{c(A, B)} \|[E, F]\|. \tag{5.5}$$

Proceeding as in Theorem 3.3 we find from (5.4) and (5.5) that the first inequality of (5.6) below is obtained.

Theorem 5.3. *Let (A, B) be a hermitian definite pair and $\tilde{A} = A+E$, $\tilde{B} = B+F$ where*

$$\eta_1 \triangleq \frac{2}{c(A, B)} \|[E, F]\| < 1.$$

Then

$$S_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \frac{\eta_1}{1-\eta_1}, \quad S^{(1)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \frac{\eta_1}{(1-\eta_1)^{1/2}}, \quad S^{(2)}_{(A,B)}(\tilde{A}, \tilde{B}) \leqslant \eta_1/2. \tag{5.6}$$

For the proof of the remaining inequalities we change the reasoning slightly. We get

$$\|X\|^2 \leqslant \frac{1}{c(A, B)},$$

where X satisfies $X^*AX = \operatorname{diag}[\hat{\alpha}_1, \cdots, \hat{\alpha}_n]$, $X^*BX = \operatorname{diag}[\hat{\beta}_1, \cdots, \hat{\beta}_n]$ and $\hat{\alpha}_i^2 + \hat{\beta}_i^2 = 1$, $i = 1, \cdots, n$. (5.2) now gives

$$\min_i \rho(\tilde{\lambda}, \hat{\alpha}_i | \hat{\beta}_i) \leqslant \eta_1/2$$

or the last inequality of (5.6). The second inequality is obtained as in the proof of Theorem 3.3.∎

We remark that the last inequality in (5.6) improves considerably a bound which is derived from Theorem 2.3 and (1.13) in [2]. This result should be compared with a bound in [7]. There it is shown that if $(\tilde{A}, \tilde{B})$ is also a definite pair, then its eigenvalues $\tilde{\lambda}_i$ can be ordered so that $\rho(\lambda_i, \tilde{\lambda}_i) \leqslant \eta_1/2$ holds for $i=1, \cdots, n$.

References

[1] C. R. Crawford, A stable generalized eigenvalue problem, *SIAM J. Numer. Anal.*, **6** (1976), 854—860.

[2] L. Elsner, Sun Ji-guang, Perturbation theorems for the generalized eigenvalue problem, *Linear Algebra Appl.*, **48** (1982), 341—357.

[3] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press, New York, 1982.

[4] I. Gohberg, P. Lancaster, L. Rodman, Matrices and Indefinite Scalar Products, Birkhäuser Verlag, Basel, 1983.

[5] I. Gohberg, P. Lancaster, L. Rodman, Perturbations of H-selfadjoint matrices, *Integral Eq. and Operator Theory*, **5** (1982), 718—757.

[6] P. Lancaster, M. Tismenetsky, Theory of Matrices (2nd Edition), Academic Press, New York, 1985.

[7] G. W. Stewart, Perturbation bounds for the definite generalized eigenvalue problem, *Linear Algebra Appl.*, **23** (1979), 69—85.

[8] F. Uhlig, A recurring theorem about pairs of quadratic forms and extensions: a survey, *Linear Algebra Appl.*, **25** (1979), 219—237.