ON PERTURBATION THEOREMS FOR THE GENERALIZED EIGENVALUES OF REGULAR MATRIX PAIRS†

Ren-cang Li*

(Computing Center, Academia Sinica, Beijing, P.R.China)

Abstract

In this paper, we study the Bauer-Fike type theorems for regular matrix pairs and its general form. Besides these, we also obtain some bounds for measuring approximate generalized eigenvalues and for the perturbation of generalized eigenvalues concerning an approximate generalized invariant subspace. To estimate all these bounds, p-Hölder norm and an ad hoc pseudo-metric called *p-Chordal matric* which isn't a metric on the Riemann sphere when $p \neq 2$ are used.

§1. Introduction

As to perturbations of eigenvalues of matrices, Bauer-Fike Theorem[3] is well-known, and its generalized version has also been deduced in Kahan et al.[5] and Li[6] so as to it can be applied to non-diagonalizable case. Note that the results in Li[6], which also generalize one of the results in Kahan et al.[5], improve slightly the results in [5]. On the other hand, for perturbations of generalized eigenvalue problems, we have similar results obtained by Elsner and Sun[1]. One of the aims of this paper is to improve slightly some of results in [1] and to generalize them to general regular matrix pair case. We remark that the norms used later are not confined to spectral norm, i.e., 2-Hölder norm.

We use the following notation: capital letters for matrices, lowercase Latin letters for column vectors and lowercase Greek letters for scalars; Also we use $\mathbb{C}^{m\times n}$ for the set of m by n complex matrices, $\mathbb{C}^m = \mathbb{C}^{m\times 1}$, $\mathbb{C} = \mathbb{C}^1$. $I^{(n)}$ stands for the n by n unit matrix (also we just write I for convenience when no confusion arises). A^T and A^H denote the transpose, conjugate transpose of A, respectively. For $p \geq 1$, $x \in \mathbb{C}^n$ and $A \in \mathbb{C}^{m\times n}$, we use

$$\|x\|_p \equiv \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}},$$
$$\|A\|_p \equiv \max_{\|x\|_p=1} \|Ax\|_p$$

to denote the p-Hölder norm of vector x and matrix A, respectively, where $x = (\xi_1, \cdots, \xi_n)^T$.

1.1 Regular Matrix Pairs Of Order n.

Definition 1.1. *Let A, $B \in \mathbb{C}^{n\times n}$. Matrix pair $\{A, B\}$ is called a regular matrix pair of order n, if*

$$\det(A + \lambda B)\text{is not zero, identically, for } \lambda \in \mathbb{C}. \tag{1.1}$$

†Published in *Math. Numer. Sinica*, **11**:1(1989), 10-19; Engl. transl. in *Chinese J. Numer. Math. Appl.*, **11**:2(1989), 24-35.

*Current address: Institute of Applied Physics and Computational Mathematics, P.O.Box 8009, Beijing, P.R.China

Definition 1.2. *The generalized eigenvalue set of a regular matrix pair $\{A,B\}$ of order n is defined by*

$$\lambda(A,B) \equiv \{(\alpha,\beta) \neq 0 \,|\, \det(\beta A - \alpha B) = 0,\ \alpha,\beta \in \mathbb{C}\}, \tag{1.2}$$

and each element $(\alpha,\beta) \in \lambda(A,B)$ is a generalized eigenvalue of the regular matrix pair $\{A,B\}$.

The canonical form of a regular matrix pair has been known a century ago, and here we just cite a theorem of Weierstrass(1867) (see e.g., Gantmacher[2]).

Theorem 1.1(Weierstrass). *Let $\{A,B\}$ be a regular matrix pair of order n, then there exist invertible matrices P, $Q \in \mathbb{C}^{n\times n}$ such that*

$$A = PJ_AQ, \quad B = PJ_BQ\,, \tag{1.3}$$

where

$$J_A = \begin{pmatrix} J & \\ & I^{(n_2)} \end{pmatrix}, \quad J_B = \begin{pmatrix} I^{(n_1)} & \\ & N \end{pmatrix}, \tag{1.4}$$

$$n_1 + n_2 = n, \tag{1.5}$$

$$J = \mathrm{diag}(\, J_1(\lambda_1), \cdots J_r(\lambda_r)\,) \in \mathbb{C}^{n_1\times n_1}, \tag{1.6}$$

$$J_i(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{n(\lambda_i)\times n(\lambda_i)},\ 1 \le i \le r, \tag{1.7a}$$

$$\sum_{i=1}^{r} n(\lambda_i) = n_1, \tag{1.7b}$$

$$N = \mathrm{diag}(\, N^{(l_1)}, \cdots, N^{(l_s)}\,) \in \mathbb{C}^{n_2\times n_2}, \tag{1.8}$$

$$N^{(l_j)} = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{C}^{l_j\times l_j},\ 1 \le j \le s, \tag{1.9a}$$

$$\sum_{j=1}^{s} l_j = n_2, \tag{1.9b.}$$

Here λ_i and λ_j $(i \neq j)$ may be equal.

As to above decomposition, we call $\{J_A, J_B\}$ is *the Weierstrass-Jordan canonical form* (WJCF) of regular matrix pair $\{A,B\}$. It is clear that $(\lambda_i, 1)$ (if $n_1 \geq 1$),$1 \le i \le r$ and/or $(1,0)$ (if $n_2 \geq 1$) are generalized eigenvalues of $\{A,B\}$. Now partitioning $I^{(n_1)}$ and $I^{(n_2)}$ in (1.4) as

$$I^{(n_1)} = \mathrm{diag}(\, I^{(n(\lambda_1))}, \cdots, I^{(n(\lambda_r))}\,)\,, I^{(n_2)} = \mathrm{diag}(\, I^{(l_1)}, \cdots, I^{(l_s)}\,),$$

then we get submatrx pairs $\{J_i(\lambda_i), I^{(n(\lambda_i)}\}$ $(1 \le i \le r)$ and $\{I^{(l_j)}, N^{(l_j)}\}$ $(1 \le j \le s)$. We call these pairs *Weierstrass-Jordan canonical submatrix pairs* (WJCSP) of order $n(\lambda_i)$ $(1 \le i \le r)$, l_j $(1 \le j \le s)$, respectively.

A very important regular matrix pair is the diagonalizable regular matrix pair.

Definition 1.3. *A regular matrix pair $\{A,B\}$ of order n is diagonalizable, or normalizable, if there exist invertible matrices P, $Q \in \mathbb{C}^{n\times n}$ such that*

$$\left.\begin{aligned} A = P\Lambda Q \equiv P\mathrm{diag}(\, \alpha_1, \cdots, \alpha_n\,)Q \\ B = P\Omega Q \equiv P\mathrm{diag}(\, \beta_1, \cdots, \beta_n\,)Q \end{aligned}\right\}. \tag{1.10}$$

1.2 Metric.

A peculiarity is the use of a pair of positive number (p, q) for a dual number pair, i.e.,

$$\frac{1}{p} + \frac{1}{q} = 1 \text{for } q > 1;\ p = 1 \text{ for } q = +\infty \text{ and } p = +\infty \text{ for } q = 1. \tag{1.11}$$

From Definition 1.2, we see that every generalized eigenvalue is represented by a nonzero complex pair, therefore the two generalized eigenvalues represent the same one if the two components of one generalized eigenvalue are the product of the corresponding components of the other with a nonzero complex number. Hence we introduce the following pseudo-metric on the Riemann sphere: Let (α, β), (γ, δ) be two generalized eigenvalues, define

$$\rho_p(\,(\alpha, \beta), (\gamma, \delta)\,) = \frac{|\delta\alpha - \beta\gamma|}{\sqrt[p]{|\alpha|^p + |\beta|^p}\,\sqrt[q]{|\gamma|^q + |\delta|^q}},\ 1 \le p \le +\infty. \tag{1.12}$$

Clearly, $p = 2$, $\rho_p(\,(\alpha, \beta), (\gamma, \delta)\,)$ is just the well-known *Chordal Metric* and is denoted simply by $\rho((\alpha, \beta), (\gamma, \delta))$ in literature. Unfortunately, (1.12) isn't a metric on the Riemann sphere when $p \neq 2$ (this can be seen by noting that $\rho_p((0, 1), (1, 1)) \neq \rho_p((1, 1), (0, 1))$). Nevertheless we will simply call (1.12) *p-Chordal metric*

Let $\{A, B\}$ and $\{C, D\}$ be two regular matrix pairs of order n, and denote $\lambda(A, B) = \{\,(\alpha_i, \beta_i),\ 1 \le i \le n\}$, $\lambda(C, D) = \{\,(\gamma_i, \delta_i),\ 1 \le i \le n\}$. Corresponding to (1.12) we define *p-spectral variation* of $\{C, D\}$ with respect to $\{A, B\}$ by

$$S^{(p)}_{\{A,B\}}\{C, D\} \equiv \max_{1 \le j \le n} \min_{1 \le i \le n} \rho_p((\alpha_i, \beta_i), (\gamma_j, \delta_j)),\ 1 \le p \le +\infty. \tag{1.13}$$

1.3 Moore-Penrose Inverse and Orthogonal Projection.

Readers are referred to Sun[4, pp.28-43] for definitions and properties of Moore-Penrose inverse and orthogonal projection. Here we just cite some of them for later use. Using $A^\dagger$ to denote the Moore-Penrose inverse of $A \in \mathbb{C}^{m\times n}$, then we have

(1) $A^\dagger = (A^H A)^{-1} A^H$, if $\mathrm{rank}A = n$, therefore $A^\dagger A = I^{(n)}$;

(2) $A^\dagger = A^H (A A^H)^{-1}$, if $\mathrm{rank}A = m$, therefore $A A^\dagger = I^{(m)}$.

Here $\mathrm{rank}A$ stands for the rank of A.

Let $X \in \mathbb{C}^{m\times n}$, $\mathcal{R}(X)$ be the column space, the subspace spanned by the column vectors of X. It can be proved that

$$P_{X^H} \equiv X^\dagger X$$

is the orthogonal projection onto the column space $\mathcal{R}(X^H)$, a $\mathrm{rank}X$-dimensional subspace.

§2. Lemmas

Lemma 2.1. *Suppose* $D_i = \mathrm{diag}(\,d_1^{(j)}, \cdots, d_n^{(j)}\,) \in \mathbb{C}^{n\times n}$, $j = 1, 2$. *Then for any* $1 \le p \le +\infty$, *we have*

$$\|\,(\,D_1, D_2\,)\,\|_p = \max_{1 \le i \le n}(\,|d_i^{(1)}|^q + |d_i^{(2)}|^q\,)^{\frac{1}{q}} \tag{2.1a}$$

$$\left\|\begin{pmatrix} D_1 \\ D_2 \end{pmatrix}\right\|_p = \max_{1 \le i \le n}(\,|d_i^{(1)}|^p + |d_i^{(2)}|^p\,)^{\frac{1}{p}}. \tag{2.1b}$$

Moreover, if every column of $X = (x_1, \cdots, x_m)^T \in \mathbb{C}^{m\times n}$ $(x_i \in \mathbb{C}^n,\ 1 \le i \le m)$ *has only one nonzero element, and every row of* $Y = (y_1, \cdots, y_n) \in \mathbb{C}^{m\times n}$ $(y_i \in \mathbb{C}^m,\ 1 \le i \le n)$ *has only one nonzero element. Then for any* $1 \le p \le +\infty$, *we have*

$$\|X\|_p = \max_{1 \le i \le m}\{\,\|x_i\|_q\,\}, \tag{2.2a}$$

$$\|Y\|_p = \max_{1 \le i \le n}\{\,\|y_i\|_p\,\}, \tag{2.2b}$$

A very simple and easy proof, omitted here, of this lemma may be given by just recalling the definition of p-norm.

Lemma 2.2. *Suppose $\{J_A, J_B\}$ is of form (1.4)-(1.9), and $(\delta J_A - \gamma J_B)$invertible, where γ, $\delta \in \mathbb{C}$ and $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$. Then there exist (α, β) such that*

$$\|(\delta J_A - \gamma J_B)^{-1}(J_A, J_B)\|_p \leq \sqrt[q]{2} \sum_{j=1}^{m} \frac{1}{\rho_q^j}, \quad 1 \leq p \leq +\infty, \tag{2.3}$$

where (α, β) is one of $(\lambda_i, 1)$ $(1 \leq i \leq r)$ or $(1, 0)$, m is the maximum order of all WJCSPs corresponding to (α, β), and

$$\rho_q \equiv \rho_q(\,(\alpha, \beta), (\gamma, \delta)\,) = \frac{|\delta\alpha - \beta\gamma|}{\sqrt[q]{|\alpha|^q + |\beta|^q}\,\sqrt[p]{|\gamma|^p + |\delta|^p}}.$$

Proof: Let's introduce the following notation:

$$\tilde{J}_{AB} \equiv (\delta J_A - \gamma J_B) = \begin{pmatrix} \tilde{J} & \\ & \tilde{N} \end{pmatrix}, \tag{2.4}$$

where

$$\tilde{J} = \text{diag}(\,\tilde{J}_1(\lambda_1), \cdots \tilde{J}_r(\lambda_r)\,), \tag{2.5a}$$

$$\tilde{J}_i(\lambda_i) \equiv \delta J_i(\lambda_i) - \gamma I^{(n(\lambda_i))} = \begin{pmatrix} \delta\lambda_i - \gamma & \delta & & \\ & \ddots & \ddots & \\ & & \ddots & \delta \\ & & & \delta\lambda_i - \gamma \end{pmatrix}, \; 1 \leq i \leq r, \tag{2.5b}$$

$$\tilde{N} = \text{diag}(\,\tilde{N}^{(l_1)}, \cdots, \tilde{N}^{(l_s)}\,), \tag{2.6a}$$

$$\tilde{N}^{(l_j)} \equiv \delta I^{(l_j)} - \gamma N^{(l_j)} = \begin{pmatrix} \delta & -\gamma & & \\ & \ddots & \ddots & \\ & & \ddots & -\gamma \\ & & & \delta \end{pmatrix}, \; 1 \leq j \leq s, \tag{2.6b}$$

Simple verification shows

$$
\begin{aligned}
&(\delta J_A-\gamma J_B)^{-1}(J_A,J_B)\\
&=\begin{pmatrix}\tilde J & \\ & \tilde N\end{pmatrix}\left(\begin{pmatrix}J & \\ & I^{(n_2)}\end{pmatrix},\begin{pmatrix}I^{(n_1)} & \\ & N\end{pmatrix}\right)\\
&=\begin{pmatrix}\ddots & & & & \\ & \tilde J_i(\lambda_i)^{-1} & & & \\ & & \ddots & & \\ & & & \tilde N^{(l_j)\,-1} & \\ & & & & \ddots\end{pmatrix}\times\\
&\quad\left(\begin{pmatrix}\ddots & & & & \\ & J_i(\lambda_i) & & & \\ & & \ddots & & \\ & & & I^{(l_j)} & \\ & & & & \ddots\end{pmatrix},\begin{pmatrix}\ddots & & & & \\ & I^{(n(\lambda_i))} & & & \\ & & \ddots & & \\ & & & N^{(l_j)} & \\ & & & & \ddots\end{pmatrix}\right)\\
&=\left(\begin{pmatrix}\ddots & & & & \\ & \tilde J_i(\lambda_i)^{-1}J_i(\lambda_i) & & & \\ & & \ddots & & \\ & & & \tilde N^{(l_j)\,-1} & \\ & & & & \ddots\end{pmatrix},\right.\\
&\quad\left.\begin{pmatrix}\ddots & & & & \\ & \tilde J_i(\lambda_i)^{-1} & & & \\ & & \ddots & & \\ & & & \tilde N^{(l_j)\,-1}N^{(l_j)} & \\ & & & & \ddots\end{pmatrix}\right).
\end{aligned}
\tag{2.7}
$$

We remark here that if $d_j^{(i)}$ $(i=1,2;\ 1\le j\le n)$ in Lemma 2.1 is replaced by matrices with appropriate orders, then more general conclusions may be obtained. In fact for (2.7) we have

$$\|(\delta J_A-\gamma J_B)^{-1}(J_A,J_B)\|_p=\max_{i,j}\begin{pmatrix}\|(\tilde J_i(\lambda_i)^{-1}J_i(\lambda_i),\tilde J_i(\lambda_i)^{-1})\|_p\\ \|(\tilde N^{(l_j)\,-1},\tilde N^{(l_j)\,-1}N^{(l_j)})\|_p\end{pmatrix}. \tag{2.8}$$

Now we are in the position of estimating the upper bounds for the terms in the right hand of (2.8). First, we take $\|(\tilde J_i(\lambda_i)^{-1}J_i(\lambda_i),\tilde J_i(\lambda_i)^{-1})\|_p=\|\tilde J_i(\lambda_i)^{-1}(J_i(\lambda_i),I)\|_p$ into consideration. Denote

$$D=\begin{pmatrix}0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1\\ & & & 0\end{pmatrix}\text{ with appropriate size and } d=\frac{\delta}{\delta\lambda_i-\gamma}. \tag{2.9}$$

Since

$$\tilde{J}_i(\lambda_i)^{-1} = \begin{pmatrix} \delta\lambda_i - \gamma & \delta & & \\ & \ddots & \ddots & \\ & & \ddots & \delta \\ & & & \delta\lambda_i - \gamma \end{pmatrix}^{-1} = \frac{1}{\delta\lambda_i - \gamma}\begin{pmatrix} 1 & d & & \\ & \ddots & \ddots & \\ & & \ddots & d \\ & & & 1 \end{pmatrix}^{-1} = \frac{1}{\delta\lambda_i - \gamma}(I - dD + (dD)^2 - \cdots + (-dD)^{n(\lambda_i)-1}),$$

Hence

$$\begin{aligned} &\tilde{J}_i(\lambda_i)^{-1}(J_i(\lambda_i), I) \\ &\quad = \tilde{J}_i(\lambda_i)^{-1}(\lambda_i + D, I) \\ &\quad = \frac{1}{\delta\lambda_i - \gamma}\Big(\lambda_i I - (\lambda_i d - 1)D + (\lambda_i d - 1)d(D)^2 - \cdots + (\lambda_i d - 1)(-d)^{n(\lambda_i)-2}(D)^{n(\lambda_i)-1}, \\ &\qquad I - dD + (dD)^2 - \cdots + (-dD)^{n(\lambda_i)-1}\Big), \end{aligned}$$

together with Lemma 2.1, we have

$$\begin{aligned} \|\tilde{J}_i(\lambda_i)^{-1}(J_i(\lambda_i), I)\|_p &\le \frac{1}{|\delta\lambda_i - \gamma|}\left[(|\lambda_i|^q + 1)^{\frac{1}{q}} + \sum_{j=1}^{n(\lambda_i)-1} \frac{|\delta|^{j-1}}{|\delta\lambda_i - \gamma|^{j-1}} \cdot \frac{(|\gamma|^q + |\delta|^q)^{\frac{1}{q}}}{|\delta\lambda_i - \gamma|}\right] \\ &\le \frac{1}{|\delta\lambda_i - \gamma|}\left[(|\lambda_i|^q + 1)^{\frac{1}{q}} + \sum_{j=1}^{n(\lambda_i)-1} \frac{\sqrt[q]{2}}{|\delta\lambda_i - \gamma|^j}\right] \\ &\le \sqrt[q]{2}\frac{1}{\frac{|\delta\lambda_i - \gamma|}{\sqrt[q]{|\lambda_i|^q + 1}}} \sum_{j=0}^{n(\lambda_i)-1} \frac{1}{\left(\frac{|\delta\lambda_i - \gamma|}{\sqrt[q]{|\lambda_i|^q + 1}}\right)^j}. \end{aligned} \tag{2.10}$$

As to the estimation of $\|(\tilde{N}^{(l_j)\,-1}, \tilde{N}^{(l_j)\,-1}N^{(l_j)})\|_p = \|\tilde{N}^{(l_j)\,-1}(I, N^{(l_j)})\|_p$, it follows from the invertibility of $(\delta J_A - \gamma J_B)$ that if $n_2 \ge 1$ then $\delta \ne 0$, otherwise $n_2 = 0$, (2.10) is enough for the proof of Lemma 2.2. Assume $n_2 \ge 1$, hence $\delta \ne 0$. Denote

$$\tilde{d} = \frac{\gamma}{\delta},$$

then it can be verified that

$$\begin{aligned} \tilde{N}^{(l_j)\,-1}(I, N^{(l_j)}) &= \tilde{N}^{(l_j)\,-1}(I, D) \\ &= \frac{1}{\delta}\left(\sum_{j=0}^{l_j-1} \tilde{d}^j D^j\right)(I, D) \\ &= \frac{1}{\delta}\left(\sum_{j=0}^{l_j-1} \tilde{d}^j (D^j, D^{j+1})\right). \end{aligned}$$

Together with Lemma 2.1, we have

$$\|\tilde{N}^{(l_j)\ -1}(I, N^{(l_j)})\|_p \leq \frac{1}{|\delta|}\left(\sum_{j=0}^{l_j-1}\frac{|\gamma|^j}{|\delta|^j}\cdot\sqrt[q]{2}\right)$$

$$\leq \sqrt[q]{2}\,\frac{1}{\frac{|\delta\cdot 1-\gamma\cdot 0|}{\sqrt[q]{1^q+0^q}}}\sum_{j=0}^{l_j-1}\frac{1}{\left(\frac{|\delta\cdot 1-\gamma\cdot 0|}{\sqrt[q]{1^q+0^q}}\right)^j}. \tag{2.11}$$

The combination of (2.11) with (2.10) leads to (2.3). Q.E.D.

§3. Bauer-Fike Type Theorems and Its General Form

Theorem 3.1. *Suppose that regular matrix pair $\{A, B\}$ of order n is diagonalizable, and admits the decomposition (1.10). Assume also that $\{C, D\}$ is another regular matrix pair of the same order n. Denote*

$$Z \equiv (A, B), \quad W \equiv (C, D).$$

Then

$$S^{(q)}_{\{A,B\}}\{C, D\} \leq \left\|\begin{pmatrix} Q & \\ & Q\end{pmatrix}(P_{Z^H} - P_{W^H})\begin{pmatrix} Q^{-1} & \\ & Q^{-1}\end{pmatrix}\right\|_p, \ 1 \leq p \leq +\infty. \tag{3.1}$$

Proof: Arbitrarily choose $(\gamma, \delta) \in \lambda(C, D)$. It is easy to see that we only need to study the case when $(\gamma, \delta) \notin \lambda(A, B)$ and without loss of generality, we assume that $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$. It follows from the singularity of $\delta C - \gamma D$ that $ZW^\dagger(\delta C - \gamma D) = \delta(ZW^\dagger C) - \gamma(ZW^\dagger D)$ is singular. On the other hand, simple verification shows

$$\begin{aligned}
&-\delta(ZW^\dagger C) + \gamma(ZW^\dagger D)\\
&= \delta(A - ZW^\dagger C) - \gamma(B - ZW^\dagger D) - (\delta A - \gamma B)\\
&= Z[\,\delta(Z^\dagger A - W^\dagger C) - \gamma(Z^\dagger B - W^\dagger D)\,] - (\delta A - \gamma B) \qquad (\text{by } ZZ^\dagger = I)\\
&= P\left\{(\Lambda Q, \Omega Q)\left[(Z^\dagger Z - W^\dagger W)\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix}\right]Q^{-1} - (\delta\Lambda - \gamma\Omega)\right\}Q\\
&= P(\delta\Lambda - \gamma\Omega)\cdot\\
&\quad\left[(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\begin{pmatrix} Q & \\ & Q\end{pmatrix}(P_{Z^H} - P_{W^H})\begin{pmatrix} Q^{-1} & \\ & Q^{-1}\end{pmatrix}\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix} - I\right]Q.
\end{aligned} \tag{3.2}$$

Hence the matrix in the brackets [ ] of (3.2) must be singular, this shows that for any consistent matrix norm $\|\cdot\|$, we have

$$\begin{aligned}
1 &\leq \left\|(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\begin{pmatrix} Q & \\ & Q\end{pmatrix}(P_{Z^H} - P_{W^H})\begin{pmatrix} Q^{-1} & \\ & Q^{-1}\end{pmatrix}\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix}\right\|\\
&\leq \|(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\|\cdot\left\|\begin{pmatrix} Q & \\ & Q\end{pmatrix}(P_{Z^H} - P_{W^H})\begin{pmatrix} Q^{-1} & \\ & Q^{-1}\end{pmatrix}\right\|\cdot\left\|\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix}\right\|,
\end{aligned}$$

thus

$$\|(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\|^{-1} \leq \left\|\begin{pmatrix} Q & \\ & Q\end{pmatrix}(P_{Z^H} - P_{W^H})\begin{pmatrix} Q^{-1} & \\ & Q^{-1}\end{pmatrix}\right\|\cdot\left\|\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix}\right\|. \tag{3.3}$$

We know that $\|\cdot\|_p$ ($1 \leq p \leq +\infty$) is consistent, therefore (3.3) is valid for norm $\|\cdot\|_p$. By Lemma 2.1 and by $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$, we have

$$\|(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\|_p^{-1} = \min_{1\leq i\leq n}\frac{|\delta\alpha_i - \gamma\beta_i|}{\sqrt[q]{|\alpha_i|^q + |\beta_i|^q}\,\sqrt[p]{|\gamma|^p + |\delta|^p}},$$

$$\left\|\begin{pmatrix}\delta I^{(n)}\\ -\gamma I^{(n)}\end{pmatrix}\right\|_p = \sqrt[p]{|\gamma|^p + |\delta|^p} = 1.$$

The above equation, (3.3) and (1.13) lead to (3.1). Q.E.D.

REMARK 3.1: It follows from (3.1) that

$$S^{(q)}_{\{A,B\}}\{C,D\} \le \|Q\|_p\|Q^{-1}\|_p\|P_{Z^H} - P_{W^H}\|_p\,,\ 1 \le p \le +\infty. \tag{3.4}$$

If we let $p = 2$ in (3.4), then (3.4) is just the Bauer-Fike type Theorem proved in Elsner and Sun[1] (see also Sun[4, pp.278-281]).

THEOREM 3.2. *Suppose that $\{A, B\}$ is a regular matrix pair of order n, and admits WJCF as described in Theorem 1.1. Assume also that $\{C, D\}$ is another regular matrix pair of the same order n. Z, W are defined as in Theorem 3.1. Then for any $(\gamma, \delta) \in \lambda(C, D)$, there exists $(\alpha, \beta) \in \lambda(A, B)$ such that*

$$\frac{\rho_q}{1 + \frac{1}{\rho_q} + \cdots + \frac{1}{\rho_q^{m-1}}} \le \sqrt[q]{2}\left\| \begin{pmatrix} Q & \\ & Q \end{pmatrix} (P_{Z^H} - P_{W^H}) \begin{pmatrix} Q^{-1} & \\ & Q^{-1} \end{pmatrix} \right\|_p,\ 1 \le p \le +\infty, \tag{3.5}$$

where

$$\rho_q = \rho_q(\,(\alpha,\beta),(\gamma,\delta)\,) = \frac{|\delta\alpha - \beta\gamma|}{\sqrt[q]{|\alpha|^q + |\beta|^q}\,\sqrt[p]{|\gamma|^p + |\delta|^p}}\,,$$

(α, β) and m are as described in Lemma 2.2. We define that the right hand of (3.5) is equal to zero when $\rho_q = 0$ (the reasonability of such definition can be seen by limiting procedure).

PROOF: If $(\gamma, \delta) \in \lambda(A, B)$, then (3.5) is obviously right. Without loss of generality, we assume that $(\gamma, \delta) \notin \lambda(A, B)$ and that $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$. Thus $\delta(ZW^\dagger C) - \gamma(ZW^\dagger D)$ is singular, while $\delta J_A - \gamma J_B$ is nonsingular. Using the same technique in the deduction of (3.2), we get

$$\begin{aligned} &-\delta(ZW^\dagger C) + \gamma(ZW^\dagger D) \\ &= P(\delta J_A - \gamma J_B)\cdot \\ &\quad \left[(\delta J_A - \gamma J_B)^{-1}(J_A, J_B) \begin{pmatrix} Q & \\ & Q \end{pmatrix} (P_{Z^H} - P_{W^H}) \begin{pmatrix} Q^{-1} & \\ & Q^{-1} \end{pmatrix} \begin{pmatrix} \delta I^{(n)} \\ -\gamma I^{(n)} \end{pmatrix} - I \right] Q. \end{aligned}$$

Similarly to (3.3), the following inequality can be obtained for any consistent matrix norm

$$\|(\delta J_A - \gamma J_B)^{-1}(J_A, J_B)\|^{-1} \le \left\| \begin{pmatrix} Q & \\ & Q \end{pmatrix} (P_{Z^H} - P_{W^H}) \begin{pmatrix} Q^{-1} & \\ & Q^{-1} \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} \delta I^{(n)} \\ -\gamma I^{(n)} \end{pmatrix} \right\|. \tag{3.6}$$

(3.6) is valid, especially for $\|\cdot\|_p$ $(1 \le p \le +\infty)$, combining (3.6) with Lemma 2.2, we finally get (3.5). Q.E.D.

REMARK 3.2: If l is the maximum order of all WJCSPs in WJCF of $\{A, B\}$, then from (3.5), we have

$$\frac{S^{(q)}_{\{A,B\}}\{C,D\}}{\sum_{j=0}^{l-1} \frac{1}{(S^{(q)}_{\{A,B\}}\{C,D\})^j}} \le \sqrt[q]{2}\left\| \begin{pmatrix} Q & \\ & Q \end{pmatrix} (P_{Z^H} - P_{W^H}) \begin{pmatrix} Q^{-1} & \\ & Q^{-1} \end{pmatrix} \right\|_p,\ 1 \le p \le +\infty. \tag{3.7}$$

When $p = 1$, then $\sqrt[q]{2} = 1$, if, in addition, at this time $l = 1$, i.e., $\{A, B\}$ diagonalizable, then (3.7) is just (3.1) with $p = 1$. Hence Theorem 3.2 may be regarded as a generalization of Theorem 3.1.

COROLLARY 3.1. *The conditions are as described in Theorem 3.2, l as in Remark 3.2. Then*

$$S^{(q)}_{\{A,B\}}\{C,D\} \le \max\left\{ l^{\frac{1}{l}}\eta^{\frac{1}{l}}\,, l\eta \right\},\ 1 \le p \le +\infty, \tag{3.8}$$

where

$$\eta = \sqrt[q]{2}\left\| \begin{pmatrix} Q & \\ & Q \end{pmatrix} (P_{Z^H} - P_{W^H}) \begin{pmatrix} Q^{-1} & \\ & Q^{-1} \end{pmatrix} \right\|_p. \tag{3.9}$$

The proof of Corollary 3.1 is easy, and it is only needed to take the features of the function of the left hand of (3.7) into consideration (see Li[6]). The following lemma (see Li[6] or Lin[7]) also enables us to deduce a corollary of Theorem 3.2.

LEMMA 3.1(LI[6] OR LIN[7]). *Suppose that $\tau > 0$. Equation*

$$x^n - \tau(1 + x + \cdots + x^{n-1}) = 0$$

has a unique positive solution x, which satisfies

$$x \le \begin{cases} \tau^{\frac{1}{n}} + \tau^{\frac{2}{n}}, & \text{if } \tau \le 1, \\ \tau^{\frac{1}{n}} + \tau, & \text{if } \tau > 1. \end{cases} \tag{3.10}$$

COROLLARY 3.2. *The conditions are as described in Corollary 3.1, then*

$$S^{(q)}_{\{A,B\}}\{C, D\} \le \begin{cases} \eta^{\frac{1}{l}} + \eta^{\frac{2}{l}}, & \text{if } \eta \le 1, \\ \eta^{\frac{1}{l}} + \eta, & \text{if } \eta > 1, \end{cases} \tag{3.11}$$

where η is defined by (3.9).

We see that (3.8) is better than (3.11) if $l = 1$, while (3.11) is better than (3.8) if $l \ge 2$. It is worth mentioning that all conclusions containing m and/or l remain valid if m and l are replaced by the order n of regular matrix pairs considered.

§4. APPROXIMATE GENERALIZED EIGENVALUES

The following question is frequently considered in matrix computations: Let $\{A, B\}$ be a regular matrix pair of order n, (γ, δ) approximate generalized eigenvalue obtained by some means, x a corresponding approximate generalized vector. How close is (γ, δ) to a generalized eigenvalue of $\{A, B\}$. It is the aim of this section to give an answer to this question.

THEOREM 4.1. *Suppose that regular matrix pair $\{A, B\}$ of order n is diagonalizable, and admits the decomposition (1.10). Assume also that (γ, δ) is its an approximate generalized eigenvalue, $x \in \mathbb{C}^n$ an approximate generalized vector corresponding to (γ, δ) and $\|x\|_p = 1$ $(1 \le p \le +\infty)$. Denote $Z = (A, B)$, then there exists $(\alpha, \beta) \in \lambda(A, B)$ such that*

$$\rho_q(\,(\alpha, \beta), (\gamma, \delta)\,) \le \|Q\|_p \|Q^{-1}\|_p \left\| P_{Z^H} \begin{pmatrix} \delta x \\ -\gamma x \end{pmatrix} \right\|_p, \; 1 \le p \le +\infty. \tag{4.1}$$

The proof, omitted here, of Theorem 4.1 is similar to that of Theorem 4.2 below.

THEOREM 4.2. *Suppose that $\{A, B\}$ is a regular matrix pair of order n, and admits WJCF as described in Theorem 1.1. Assume also that (γ, δ) is its approximate generalized eigenvalue, $x \in \mathbb{C}^n$ an approximate generalized vector corresponding to (γ, δ) and $\|x\|_p = 1$ $(1 \le p \le +\infty)$. Denote $Z = (A, B)$, then there exists $(\alpha, \beta) \in \lambda(A, B)$ such that*

$$\frac{\rho_q}{1 + \frac{1}{\rho_q} + \cdots + \frac{1}{\rho_q^{m-1}}} \le \sqrt[q]{2}\|Q\|_p \|Q^{-1}\|_p \left\| P_{Z^H} \begin{pmatrix} \delta x \\ -\gamma x \end{pmatrix} \right\|_p, \; 1 \le p \le +\infty, \tag{4.2}$$

where (α, β), m and ρ_q are as described in Lemma 2.2.

PROOF: If $(\gamma, \delta) \in \lambda(A, B)$, then (4.2) is obviously right. Without loss of generality, we assume that $(\gamma, \delta) \notin \lambda(A, B)$ and that $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$. Denote

$$e \equiv \delta Ax - \gamma Bx = Z \begin{pmatrix} \delta x \\ -\gamma x \end{pmatrix}.$$

It is easy to verify that

$$\delta Ax - \gamma Bx = P(\delta J_A - \gamma J_B)Qx,$$

and that

$$\begin{aligned}\delta Ax - \gamma Bx &= e \\ &= ZZ^{\dagger}e \qquad \text{(since rank}Z = n) \\ &= P(J_A, J_B)\begin{pmatrix} Q & \\ & Q\end{pmatrix} Z^{\dagger}Z\begin{pmatrix}\delta x \\ -\gamma x\end{pmatrix}.\end{aligned}$$

Therefore

$$x = Q^{-1}(\delta J_A - \gamma J_B)^{-1}(J_A, J_B)\begin{pmatrix} Q & \\ & Q\end{pmatrix} P_{Z^H}\begin{pmatrix}\delta x \\ -\gamma x\end{pmatrix},$$

from which follows that

$$\|x\| = \|Q^{-1}\|\|(\delta J_A - \gamma J_B)^{-1}(J_A, J_B)\|\left\|\begin{pmatrix} Q & \\ & Q\end{pmatrix}\right\|\left\|P_{Z^H}\begin{pmatrix}\delta x \\ -\gamma x\end{pmatrix}\right\|,$$

for any consistent matrix norm. Especially for $\|\cdot\|_p$ $(1 \le p \le +\infty)$ the above inequality is valid. Again utilizing Lemma 2.2, we will get (4.2). Q.E.D.

REMARK 4.1: As in Corollaries 3.1, 3.2, we also have the following inequalities: Suppose that the conditions are as described in Theorem 4.2, then

$$\rho_q \le \max\left\{ m^{\frac{1}{m}}\eta^{\frac{1}{m}}, m\eta\right\}, \tag{4.3}$$

$$\rho_q \le \begin{cases} \eta^{\frac{1}{m}} + \eta^{\frac{2}{m}}, & \text{if } \eta \le 1, \\ \eta^{\frac{1}{m}} + \eta, & \text{if } \eta > 1, \end{cases} \tag{4.4}$$

where ρ_q is as shown in Theorem 4.2, and

$$\eta = \sqrt[q]{2}\|Q\|_p\|Q^{-1}\|_p\left\|P_{Z^H}\begin{pmatrix}\delta x \\ -\gamma x\end{pmatrix}\right\|_p.$$

The inequalies (4.2)-(4.4) are all valid, if m is replaced by l or n, where l is as described in Remark 3.2.

REMARK 4.2: In order to make the right hands of (4.1)-(4.4) contain $\|\delta Ax - \gamma Bx\|$ explicitly, we may use the following estimation

$$\left\|P_{Z^H}\begin{pmatrix}\delta x \\ -\gamma x\end{pmatrix}\right\|_p = \|Z^{\dagger}(\delta Ax - \gamma Bx)\|_p \le \|Z^{\dagger}\|_p\|\delta Ax - \gamma Bx\|_p. \tag{4.5}$$

§5. EIGENVALUE PERTURBATION CONCERNING APPROXIMATE GENERALIZED INVARIANT SUBSPACES

DEFINITION 5.1. *Let $\{A, B\}$ be a regular matrix pair of order n, $\{A_1, B_1\}$ be a regular matrix pair of order t $(1 \le t \le n-1)$, and $X \in \mathbb{C}^{n\times t}$. We call $\mathcal{X} = \mathcal{R}(X)$ a t-dimensional generalized invariant subspace of $\{A, B\}$, if*

$$\text{rank}X = t, \quad AXB_1 = BXA_1. \tag{5.1}$$

A throrough investigation have been made by Sun[8], in which a sufficient and necessary condition, under which (5.1) is satisfied, has been given. We know from the hypothese of Definition 5.1 that

$$\lambda(A_1, B_1) \subset \lambda(A, B). \tag{5.2}$$

However, if (5.1) isn't satisfied, i.e.,

$$0 \ne AXB_1 - BXA_1 = R \in \mathbb{C}^{n\times t}, \tag{5.3}$$

then (5.2), generally, isn't true. Thus the following question arises: Are there any relation between $\lambda(A_1, B_1)$ and $\lambda(A, B)$? we will answer it by Theorem 5.1 below.

THEOREM 5.1. *Let $\{A, B\}$ be a regular matrix pair of order n, $\{A_1, B_1\}$ be a regular matrix pair of order t $(1 \le t \le n-1)$, and $X \in \mathbb{C}^{n\times t}$,*

$$\text{rank}X = t, \quad AXB_1 - BXA_1 = R \in \mathbb{C}^{n\times t}. \tag{5.4}$$

Denote

$$Z = (A, B), \quad Z_1 = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}.$$

(i) if $\{A, B\}$ is diagonalizable, and admits the decomposition (1.10). Then for any $(\gamma, \delta) \in \lambda(A_1, B_1)$, there exist $(\alpha, \beta) \in \lambda(A, B)$ such that

$$\rho_q(\,(\alpha,\beta),(\gamma,\delta)\,) \le \|Z_1^\dagger\|_p \|X^\dagger Q^{-1}\|_p \|Q\|_p \|Z^\dagger(AXB_1 - BXA_1)\|_p,\ 1 \le p \le +\infty. \tag{5.5}$$

(ii) Generally, $\{A, B\}$ admits WJCF as described in Theorem 1.1. Then for any $(\gamma, \delta) \in \lambda(A_1, B_1)$, there exist $(\alpha, \beta) \in \lambda(A, B)$ such that

$$\frac{\rho_q}{1 + \frac{1}{\rho_q} + \cdots + \frac{1}{\rho_q^{m-1}}} \le \sqrt[q]{2}\|Z_1^\dagger\|_p \|X^\dagger Q^{-1}\|_p \|Q\|_p \|Z^\dagger(AXB_1 - BXA_1)\|_p,\ 1 \le p \le +\infty, \tag{5.6}$$

where (α, β) and m are as described in Lemma 2.2.

PROOF:

(i) If $(\gamma, \delta) \in \lambda(A, B)$, then (5.5) is obviously right. Without loss of generality, we assume that $(\gamma, \delta) \notin \lambda(A, B)$ and that $\sqrt[p]{|\gamma|^p + |\delta|^p} = 1$. Suppose that y is a generalized eigenvector of $\{A_1, B_1\}$ corresponding to (γ, δ), and moreover $\|y\|_p = 1$, then

$$\delta A_1 y = \gamma B_1 y,$$

from which and

$$Ry = AXB_1y - BXA_1y$$

it follows that

$$\begin{aligned} \delta Ry &= (\delta A - \gamma B)XB_1 y \\ &= P(\delta\Lambda - \gamma\Omega)QXB_1y. \end{aligned} \tag{5.7}$$

On the other hand, we also have

$$\delta Ry = \delta ZZ^\dagger Ry = \delta P(\Lambda, \Omega) \begin{pmatrix} Q & \\ & Q \end{pmatrix} Z^\dagger Ry, \tag{5.8}$$

the right-hand sides of (5.8) and of (5.7) give

$$\begin{aligned} &\delta Q^{-1}(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega) \begin{pmatrix} Q & \\ & Q \end{pmatrix} Z^\dagger Ry = XB_1y \\ \Rightarrow &\delta(X^HX)^{-1}X^HQ^{-1}(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega) \begin{pmatrix} Q & \\ & Q \end{pmatrix} Z^\dagger Ry = B_1y \\ \Rightarrow &\delta Cy = B_1y, \end{aligned} \tag{5.9}$$

where

$$C = (X^HX)^{-1}X^HQ^{-1}(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega) \begin{pmatrix} Q & \\ & Q \end{pmatrix} Z^\dagger Ry.$$

By similar arguments, we also have

$$\gamma Cy = A_1 y. \tag{5.10}$$

Combining (5.9) and (5.10), we get

$$\begin{aligned} &\begin{pmatrix} \gamma Cy \\ \delta Cy \end{pmatrix} = \begin{pmatrix} \gamma I \\ \delta I \end{pmatrix} Cy = \begin{pmatrix} A_1 y \\ B_1 y \end{pmatrix} = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} y = Z_1 y \\ \Rightarrow &(Z_1^H Z_1)^{-1} Z_1^H \begin{pmatrix} \gamma I \\ \delta I \end{pmatrix} Cy = y, \end{aligned} \tag{5.11}$$

since $\mathrm{rank} Z_1 = t$ for $\{A_1, B_1\}$ being a regular matrix pair of order t. Take p-norm in the two sides of (5.11)

$$\begin{aligned} 1 &\le \|Z_1^\dagger\|_p \left\| \begin{pmatrix} \gamma I \\ \delta I \end{pmatrix} \right\|_p \|C\|_p \\ &\le \|Z_1^\dagger\|_p \|X^\dagger Q^{-1}\|_p \|Q\|_p \|(\delta\Lambda - \gamma\Omega)^{-1}(\Lambda, \Omega)\|_p \|Z^\dagger R\|_p. \end{aligned}$$

Again applying Lemma 2.1 to the above inequality, we will get (5.5).

(ii) Similarly to (i), we will also get (5.11) with C defined by

$$C = X^\dagger Q^{-1}(\delta J_A - \gamma J_B)^{-1}(J_A, J_B) \begin{pmatrix} Q & \\ & Q \end{pmatrix} Z^\dagger R y,$$

and then apply Lemma 2.2 to complete our proof. Q.E.D.

Remark 5.1: As we remark in Remark 4.1, the following inequalities can been deduced: Suppose that the conditions are as described in Theorem 5.1(ii), then

$$\rho_q \le \max\left\{ m^{\frac{1}{m}} \eta^{\frac{1}{m}}, m\eta \right\}, \tag{5.12}$$

$$\rho_q \le \begin{cases} \eta^{\frac{1}{m}} + \eta^{\frac{2}{m}}, & \text{if } \eta \le 1, \\ \eta^{\frac{1}{m}} + \eta, & \text{if } \eta > 1, \end{cases} \tag{5.13}$$

where m, ρ_q are as in Theorem 5.1(ii), and

$$\eta = \sqrt[q]{2} \|Z_1^\dagger\|_p \|X^\dagger Q^{-1}\|_p \|Q\|_p \|Z^\dagger (AXB_1 - BXA_1)\|_p. \tag{5.14}$$

Acknowledgement. *The author would like to thank Professor Sun Ji-guang for his supervision and guidance.*

References

[1] Elsner,L. and Sun,Ji-guang, Perturbation theorems for the generalized eigenvalue problem , *Lin. Alg. Appl.*, 48(1982), 341-51.

[2] Gantmacher, F., *The Theory Of Matrices*, Gosud. Izdat, Tehn, -Teor. Lit., Moscow, 1953; transl. by Hirsch, K., Chelsea, New York, 1959.

[3] Bauer, F. L. and Fike, C. T., Norms and exclusion theorems, *Numer., Math.*, 2(1960), 137-141.

[4] Sun, Ji-guang, *Matrix Perturbation Analysis* (in Chinese), Science Press, Beijing , 1987.

[5] Kahan, W. M., Parlett, B. N. and Jiang, Erxiong, Residual bounds on approximate eigensystems of nonnormal matrices, *SIAM J. Numer. Anal.*, 19(1982), 470-484.

[6] Li, Ren-cang, On the perturbation bounds for eiegnvalues of a matrix, preprint, 1985.

[7] Lin, Ping, On the perturbation bouns for eiegnvalues concerning approximate generalized invariant subspaces of nonnormal matrices, Nanjing Univ., 1985, to appear.

[8] Sun, Ji-guang, Invariant subspaces and generalized invariant subspaces (I), *Math., Numer., Sinica,* 4(1982),23-29.