Exclusion Theorems and The Perturbation Analysis of The Generalized Eigenvalue Problem.

by

Dr. King-wah Eric Chu

NUMERICAL ANALYSIS REPORT No. NA/11/85

Abbreviated Title:-

Perturbation Analysis of The GEVP.

Abstract:- A Bauer-Fike type exclusion theorem is proved for the eigenvalue problem $Ax = \lambda Bx$. A generalization is then applied to obtain perturbation bounds for multiple eigenvalues. The conditioning of an eigenvalue (finite or infinite) is proved to be dependent on a Jordan type condition number, its eigenvector deficiency and how regular is the matrix pencil $(A-\lambda B)$.

Acknowledgement.

The work in this paper has been carried out under sponsorship by the Science and Engineering Research Council of the United Kingdom.

1. Introduction.

The aim of this paper is to prove a Bauer-Fike type exclusion theorem [1] for the generalized eigenvalue problem (GEVP) $Ax = \lambda Bx$. Using techniques similar to that in [2], the most general multiple eigenvalue cases are considered and a perturbation theorem for the eigenvalues is proved. The implications on the sensitivity of the eigenvalues is then discussed. The result complements those by Stewart [8]-[11]. (See also [4] and [12].)

The paper is essentially a generalization of [2] by the author on the ordinary eigenvalue problem $Ax = \lambda x$. Other interesting references concerning the GEVP are in [5] [6] [7] [14] [15] and the references therein.

Consider the GEVP in the more convenient and sensible form $\det(\lambda_\beta A-\lambda_\alpha B) = 0 \iff$

$$\begin{cases} \lambda_\beta Ax = \lambda_\alpha Bx , & \text{(1a)} \\ \lambda_\beta y^H A = \lambda_\alpha y^H B , & \text{(1b)} \end{cases}$$

where x and y are the right and left eigenvectors respectively, with $(\cdot)^H$ denoting the hermitian. The classical eigenvalue λ is thus the ratio $\lambda_\alpha/\lambda_\beta$. Assume some sort of normalization for the eigenvectors x and y, the generalized eigenvalue pair $(\lambda_\alpha,\lambda_\beta)$ is then unique up to a scaling factor and one may scale $(\lambda_\alpha,\lambda_\beta)$ by the factor $(|\lambda_\alpha|^2 + |\lambda_\beta|^2)^{-\frac{1}{2}}$ so that the resulting eigenvalue pair in $\mathbb{C}^2$ has 2-norm equals to unity.

Note that equation (1) treats both finite and infinite eigenvalues similarly and a symmetry in the roles of the matrices A and B evolves.

In addition, equation (1) represents precisely how generalized eigenvalues are calculated numerically by the QZ algorithm [7] [14], in which ordered pairs $(\lambda_\alpha, \lambda_\beta)$ are produced instead of λ .

Equivalently, one is trying to look for eigenvectors x and y such that, for simple eigenvalues,

$$\begin{cases} y^H A x = \lambda_\alpha , & \text{(2a)} \\ y^H B x = \lambda_\beta , & \text{(2b)} \end{cases}$$

with some scaling for the vectors x, y, e.g.

$$\begin{cases} \|x\|_2^2 \triangleq x^H x = 1 , & \text{(3a)} \\ \|y\|_2^2 = 1 . & \text{(3b)} \end{cases}$$

Assuming from now on that the matrix pencil $(A-\lambda B)$ is regular so that $\det(\lambda_\beta A - \lambda_\alpha B)$ cannot be identically zero unless $\lambda_\alpha = \lambda_\beta = 0$.

For the most general multiple eigenvalue structure, one has, similar to equation (2) and the Kronecker canonical form [5],

$$Y^H A X = \Lambda_\alpha = \begin{bmatrix} \Lambda_{\alpha_1} & 0 \\ 0 & J_{\alpha_2} \end{bmatrix} , \qquad \text{(4a)}$$

$$Y^H B X = \Lambda_\beta = \begin{bmatrix} N_{\beta_1} & 0 \\ 0 & \Lambda_{\beta_2} \end{bmatrix} , \qquad \text{(4b)}$$

with some scaling for the columns of the non-singular eigenvector matrices X and Y . (One can keep to the scaling defined by equation (3) but the scaling only changes the values of the non-zero elements of the matrices Λ_{α_1} , J_{α_2} , N_{β_1} and Λ_{β_2} , and does not affect the development of the theory.)

Here Λ_{α_1} and Λ_{β_2} are diagonal matrices,

$$\Lambda_{\alpha_1} \triangleq \mathrm{diag}(\lambda_{\alpha_{1i}}) \tag{5}$$

and

$$\Lambda_{\beta_2} \triangleq \mathrm{diag}(\lambda_{\beta_{2i}}) . \tag{6}$$

The matrices J_{α_2} and N_{β_1} are in Jordan canonical forms

$$J_{\alpha_2} \triangleq \mathrm{diag}(J_{\alpha_{2i}}) \tag{7}$$

and

$$N_{\beta_1} \triangleq \mathrm{diag}(N_{\beta_{1i}}) , \tag{8}$$

with diagonal elements $\lambda_{\alpha_{2i}}$ and $\lambda_{\beta_{1i}}$ $(= 0)$ respectively, and the dimensions of the Jordan blocks $J_{\alpha_{2i}}$ and $N_{\beta_{1i}}$ equal to p_i and q_i respectively. The generalized eigenvalues are then the ordered pairs $(\lambda_{\alpha_{ki}}, \lambda_{\beta_{ki}})$, $k = 1,2$.

The columns x_j and y_j of the matrices X and Y now contain the eigenvectors as well as the generalized eigenvectors or principal vectors.

Let the matrices A and B be perturbed to $\tilde{A} = A + \delta A$ and $\tilde{B} = B + \delta B$ respectively, and let $\tilde{\Lambda}_\alpha$ and $\tilde{\Lambda}_\beta$ be defined similarly as Λ_α and Λ_β in equations (4) to (8), for the perturbed regular pencil. We can now present the Bauer-Fike type exclusion theorem.

2. The Exclusion Theorem.

Theorem 1. Let $(\tilde{\lambda}_\alpha, \tilde{\lambda}_\beta)$ be a generalized eigenvalue pair of the regular matrix pencil $(\tilde{A}-\lambda\tilde{B})$. In addition assume $\Delta \triangleq \|\tilde{\lambda}_\beta \cdot \delta A - \tilde{\lambda}_\alpha \cdot \delta B\|$ and $\tilde{\kappa} \triangleq \|X\| \cdot \|Y\|$. One has

$$\min_{i,j} \left| \tilde{\lambda}_\beta \cdot \lambda_{\alpha_{ij}} - \tilde{\lambda}_\alpha \cdot \lambda_{\beta_{ij}} \right| \leq \phi \tag{9}$$

with $p \triangleq \max_i(p_i)$, $q \triangleq \max_i(q_i)$ the dimensions of the biggest Jordan blocks in J_{α_2} and N_{β_1} respectively, and

(i) $\phi = |\tilde{\lambda}_\beta| \cdot \max\{\theta, \theta^{1/p}\}$, $\theta = \mathscr{C} \cdot \Delta \cdot \tilde{\kappa} \cdot |\tilde{\lambda}_\beta|^{-1}$,

when $i = 2$ and $\tilde{\lambda}_\beta \neq 0$;

$\mathscr{C} = p$ for the 1- or ∞- norm in Δ and $\tilde{\kappa}$ and

$\mathscr{C} = [p(p+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm.

(ii) $\phi = \mathscr{C} \cdot \Delta \cdot \tilde{\kappa}$, when $i = 2$ and $\tilde{\lambda}_\beta = 0$; $\mathscr{C} = 1$ for the 1- or ∞-norm and $\mathscr{C} = p^{\frac{1}{2}}$ for the 2- or F-norm.

(iii) $\phi = |\tilde{\lambda}_\alpha| \cdot \max\{\theta, \theta^{1/q}\}$, $\theta = \mathscr{C} \cdot \Delta \cdot \tilde{\kappa} \cdot |\tilde{\lambda}_\alpha|^{-1}$, when $i = 1$ and $\tilde{\lambda}_\alpha \neq 0$;

$\mathscr{C} = q$ for the 1- or ∞-norm and

$\mathscr{C} = [q(q+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm.

and (iv) $\phi = \mathscr{C} \cdot \Delta \cdot \tilde{\kappa}$, when $i = 1$ and $\tilde{\lambda}_\alpha = 0$;

$\mathscr{C} = 1$ for the 1- or ∞-norm and

$\mathscr{C} = q^{\frac{1}{2}}$ for the 2- or F-norm.

(Proof) Following the techniques in [1] and [2], consider the matrix

$$Y^H\{\tilde{\lambda}_\beta \tilde{A} - \tilde{\lambda}_\alpha \tilde{B}\}X = (\tilde{\lambda}_\beta \cdot \Lambda_\alpha - \tilde{\lambda}_\alpha \cdot \Lambda_\beta) \cdot (I + M_1) \tag{10}$$

with

$$M_1 = (\tilde{\lambda}_\beta \cdot \Lambda_\alpha - \tilde{\lambda}_\alpha \cdot \Lambda_\beta)^{-1} \cdot Y^H (\tilde{\lambda}_\beta \cdot \delta A - \tilde{\lambda}_\alpha \cdot \delta B) X$$

$$= M_2^{-1} \cdot M_3 \; . \tag{11}$$

One can assume the invertibility of the matrix M_2, otherwise the LHS of the inequality (9) vanishes.

From the definition of the eigenvalue $(\tilde{\lambda}_\alpha, \tilde{\lambda}_\beta)$, the matrix in equation (10) is singular and the matrix M_1 in equations (10) and (11) satisfies

$$\| M_1 \| \geq 1 \implies \| M_2^{-1} \| \cdot \| M_3 \| \geq 1 \; . \tag{12}$$

One has to estimate $\| M_2^{-1} \|$ and for the Holder norms we are considering, one has

$$\| M_2^{-1} \| = \max_j \| (\tilde{\lambda}_\beta \Lambda_{\alpha_{1j}} - \tilde{\lambda}_\alpha N_{\beta_{1j}})^{-1} \| \tag{13a}$$

or

$$\| M_2^{-1} \| = \max_j \| (\tilde{\lambda}_\beta J_{\alpha_{2j}} - \tilde{\lambda}_\alpha \Lambda_{\beta_{2j}})^{-1} \| \; , \tag{13b}$$

depending on which block in equation (4) the maximum in equation (13) occurs at.

Consider equation (13a), when the maximum occurs at j and denote $(\tilde{\lambda}_\beta \lambda_{\alpha_{1k}} - \tilde{\lambda}_\alpha \lambda_{\beta_{1k}})$ by z_k, one has

$$(\tilde{\lambda}_\beta \Lambda_{\alpha_{1j}} - \tilde{\lambda}_\alpha N_{\beta_{1j}})^{-1} = \begin{bmatrix} z_1 & -\tilde{\lambda}_\alpha & & \bigcirc \\ & z_2 & -\tilde{\lambda}_\alpha & \\ & & \ddots \; z_{p_j-1} & \ddots -\tilde{\lambda}_\alpha \\ \bigcirc & & & z_{p_j} \end{bmatrix}^{-1}$$

$$= \left[\begin{array}{c|c|c|c|c} z_1^{-1} & \tilde{\lambda}_\alpha z_1^{-1} z_2^{-1} & \tilde{\lambda}_\alpha^2 z_1^{-1} z_2^{-1} z_3^{-1} & \cdots & \\ \hline 0 & z_2^{-1} & \tilde{\lambda}_\alpha z_2^{-1} z_3^{-1} & \cdots & \\ \hline 0 & 0 & z_3^{-1} & \cdots & \\ \hline 0 & 0 & 0 & \ddots & \tilde{\lambda}_\alpha z_{p_j-1}^{-1} z_{p_j}^{-1} \\ \hline 0 & 0 & 0 & 0 & z_{p_j}^{-1} \end{array}\right] . \tag{14}$$

Consider the cases $\tilde{\lambda}_\alpha \neq 0$ and $\tilde{\lambda}_\alpha = 0$, $\| M_2^{-1} \|$ is then majorized using equation (14), by considering $\min_k |z_k|$, which in turn majorizes the LHS of the inequality (9). The constant $\mathcal{C}$ is obtained through using different norms and thus cases (iii) and (iv) have been proved. Note that the 2-norm result is proved by considering the F-norm of M_2^{-1} .

Cases (i) and (ii), mirror images of cases (iii) and (iv), are proved similarly by using equation (13b).

■

Some interesting observations can be made through Theorem 1:-

(1) the quantity on the LHS of the inequality (9), with the scaling factor $(|\tilde{\lambda}_\alpha|^2 + |\tilde{\lambda}_\beta|^2)^{-\frac{1}{2}} \cdot (|\lambda_{\alpha_{ij}}|^2 + |\lambda_{\beta_{ij}}|^2)^{-\frac{1}{2}}$

incorporated, has been used by various authors in [9],[10],[11],[12],[13]. Apart from evolving naturally from the Bauer-Fike type theorem in Theorem 1, the quantity can be interpreted in several ways. It is a chordal metric on the Neumann sphere [9], and the sine of the angle between the vectors $(\tilde{\lambda}_\alpha, \tilde{\lambda}_\beta)$ and $(\lambda_{\alpha_{ij}}, \lambda_{\beta_{ij}})$. In addition, one can define a partition of equivalence classes of generalized eigenvalue pairs $(\lambda_\alpha, \lambda_\beta)$, using the usual equivalence relation $\sim$ for quotients

$$(\lambda_{\alpha_1},\lambda_{\beta_1}) \sim (\lambda_{\alpha_2},\lambda_{\beta_2}) \quad \text{iff} \quad (\lambda_{\alpha_1}\lambda_{\beta_2} - \lambda_{\beta_1}\lambda_{\alpha_2}) = 0 \ . \tag{15}$$

The LHS of the inequality (9) is thus a measure of the distances between equivalence classes.

(2) One does not need to assume the smallness of the perturbations.

(3) Theorem 1 degenerates into the Bauer-Fike Theorem [1] and its generalizations [2], if $B = I$ and $\delta B = 0$.

(4) Zero and infinite eigenvalues are treated similarly and do not have to be ill-conditioned.

(5) For definite matrix pairs [11][12], or when $p = q = 1$, Theorem 1 indicates well-conditioning for the GEVP. Consider

$$\phi = \Delta\tilde{\kappa} = \|\tilde{\lambda}_\beta \cdot \delta A - \tilde{\lambda}_\alpha \cdot \delta B\| \cdot \|X\| \cdot \|Y\|$$

$$\leq \max\{|\tilde{\lambda}_\alpha|, |\tilde{\lambda}_\beta|\} \cdot \|X\| \cdot \|Y\| \cdot (\|\delta A\| + \|\delta B\|)$$

$$\Longrightarrow \quad |\tilde{\lambda}_\beta\lambda_{\alpha_{ij}} - \tilde{\lambda}_\alpha\lambda_{\beta_{ij}}|/(\tilde{\nu}\nu_{ij}) \leq \tilde{\kappa}\nu_{ij}^{-1} \cdot (\|\delta A\| + \|\delta B\|) \ , \tag{16}$$

where $\tilde{\nu} \triangleq (|\tilde{\lambda}_\alpha|^2 + |\tilde{\lambda}_\beta|^2)^{\frac{1}{2}}$ and $\nu_{ij} \triangleq (|\lambda_{\alpha_{ij}}|^2 + |\lambda_{\beta_{ij}}|^2)^{\frac{1}{2}}$.

The result in inequality (16) is similar to that by Stewart in [9] [10] [11]. Similarly, the quantity $\tilde{\kappa}\nu_{ij}^{-1}$ can be considered to be a condition number of the GEVP, with $\tilde{\kappa}$ measuring the conditioning of the eigenvectors, and ν_{ij}^{-1} the regularity of the generalized eigenvalue pair $(\lambda_{\alpha_{ij}},\lambda_{\beta_{ij}})$ of the matrix pencil $(A-\lambda B)$. From the QZ decomposition [7] of the pencil, it is obvious that $\nu \triangleq \min_{i,j}(\nu_{ij})$ is the size of the smallest possible perturbation which produces singularity and can be considered to be a measure of the distance between the matrix pencil $(A-\lambda B)$ and its nearest singular neighbour. Note that ν_{ij} in inequality (9) is replacable by ν .

(6) For cases (i) and (ii) and $p > 1$, the condition number will be $\mathscr{E}\tilde{\kappa}\nu_{ij}^{-1}$ or $(\mathscr{E}\tilde{\kappa}\nu_{ij}^{-p})^{1/p}$, depending on which cases one is considering, and where the maximum in ϕ occurs at. The conditioning of the GEVP is thus dependent on the sizes of p , $\tilde{\kappa}$ and ν_{ij}. The quantity p reflects the deficiency of the eigenvectors, $\tilde{\kappa}$ the conditioning of the invariant subspaces, and ν_{ij} the regularity of the generalized eigenvalue $(\lambda_{\alpha_{ij}},\lambda_{\beta_{ij}})$. Again, note that ν_{ij} can be replaced by ν and the results still hold. Note also that $\tilde{\kappa}$ is a generalization of the usual condition number and the Jordan condition number. One can choose a minimum $\tilde{\kappa}$, as it is not unique.

(7) For case (ii) with $\tilde{\lambda}_\beta = 0$, there is no fractional power of the perturbation on the RHS of inequality (9), and the geometric structure of the canonical form Λ_α does not come into play. It is because a finite eigenvalue has been perturbed to an infinite one and one is interested in the null space of the matrix $\tilde{B}$. Thus the matrix Λ_α does not feature in the analysis. A similar observation holds for case (iv).

(8) Return to inequality (9) and the definition of Δ , perturbation in the matrix A does not affect the singularity of the matrix B , and vise versa. In other words, perturbations in the matrix A does not change the behaviour of the generalized eigenvalues $(\lambda_\alpha,0)$, a traditional infinite eigenvalue. Similarly observations

hold for the zero eigenvalues as well.

(9) One can use the usual continuity argument in [1] [9] to show that the Gershgorin regions, defined by the inequality (9) for different values of i and j, satisfying the exclusion type theorems as in Theorems 2.1 and 2.2 by Stewart [9].

3. The Matrix Equation $AXB - CXD = E$.

Following the techniques in [2], we then try to generalize Theorem 1 to cope with the situation when one is interested only in part of the spectrum. As a bonus, one obtains "condition numbers" for different clusters of generalized eigenvalues.

One needs the following theorem from [3, Theorem 1] for the generalization:-

Theorem 2. The matrix equation

$$AXB - CXD = E \tag{17}$$

has a unique solution if and only if

(i) The matrix pencils $(A-\lambda C)$ and $(D-\lambda B)$ are regular,

and (ii) $\rho(A,C) \cap \rho(D,B) = \emptyset$,

with $\rho(A,C)$ denoting the equivalence classes of generalized eigenvalue pairs of the matrix pencil $(A-\lambda C)$, defined by the equivalence relation (15).

(Proof): It is trivial, using the QZ decompositions [7] of the matrix pencils. ∎

The importance of the equation (17) to the GEVP is the same as that of the Sylvester equation (of which equation (17) is a generalization) to the ordinary eigenvalue problem $Ax = \lambda x$. In addition, for regular matrix pencils $(A-\lambda C)$ and $(D-\lambda B)$, equation (17) can be proved [3] to be equivalent to the equation

$$(YA-DZ, YC-BZ) = (E,F) \tag{18}$$

introduced by Stewart [8].

Numerical algorithms for the solution of equations (17) and (18) are discussed in [3].

The following corollary of Theorem 1 can be proved, for eigenvalues satisfying an equation with a small residual.

Consider only part of the spectrum of the matrix pencil $(A-\lambda B)$. Let $\hat{X}_1$ be an $n \times n_1$ matrix such that

$$A(\hat{X}_1,\hat{X}_2)\begin{bmatrix}\hat{\Lambda}_{\beta_1} & O \\ O & \hat{\Lambda}_{\beta_2}\end{bmatrix} = B(\hat{X}_1,\hat{X}_2)\begin{bmatrix}\hat{\Lambda}_{\alpha_1} & O \\ O & \hat{\Lambda}_{\alpha_2}\end{bmatrix}, \quad (19)$$

and

$$A\tilde{X}_1\tilde{\Lambda}_{\beta_1} = B\tilde{X}_1\tilde{\Lambda}_{\alpha_1} + D, \quad (20)$$

with D is small residual matrix.
As the matrix pencils are regular, $\begin{pmatrix}\tilde{\Lambda}_{\alpha_1} \\ \tilde{\Lambda}_{\beta_1}\end{pmatrix}$ is full-ranked. One can then construct the decomposition of $\hat{Y}_1^H D$, with $\hat{Y}_i^H$ the left eigenvector matrix corresponding to $\hat{X}_i$:-

$$\hat{Y}_1^H D = Z_\alpha\tilde{\Lambda}_{\alpha_1} - Z_\beta\tilde{\Lambda}_{\beta_1}. \quad (21)$$

Denote $\hat{X}_1^H\tilde{X}_1$ by P and define the operator T by

$$T(M) \triangleq \hat{\Lambda}_{\alpha_2}M\tilde{\Lambda}_{\beta_1} - \hat{\Lambda}_{\beta_2}M\tilde{\Lambda}_{\alpha_1}, \quad (22)$$

the matrix P will be non-singular if T is invertible and $T^{-1}(D)$ is small enough. (See Corollary 3 below and c.f. Thoerem 5.1 in [8].) Note that the operator T is invertible if the conditions similar to that in Theorem 2 are satisfied.

Denote the generalized eigenvalues of the matrix pencils $(\hat{\Lambda}_{\alpha_1} - \lambda\hat{\Lambda}_{\beta_1})$ and $(\tilde{\Lambda}_{\alpha_1} - \lambda\tilde{\Lambda}_{\beta_1})$ by $(\hat{\alpha},\hat{\beta})$ and $(\tilde{\alpha},\tilde{\beta})$ respectively. Let $\hat{p}$ and $\hat{q}$ be the dimensions of the biggest Jordan blocks in $\hat{\Lambda}_{\alpha_1}$ and $\hat{\Lambda}_{\beta_1}$ respectively, and denote $\|\hat{X}_1\|\cdot\|\hat{Y}_1\|$ by $\hat{\kappa}$. One has the following corollary:

<u>Corollary 3.</u> With the assumptions and notations in this section, one has

$$\min_{\hat{\alpha},\hat{\beta}} |\tilde{\alpha}\hat{\beta} - \tilde{\beta}\hat{\alpha}| \leq \phi \tag{23}$$

where (i) $\phi = |\tilde{\beta}|\cdot(\mathcal{C}\cdot\mathcal{D}\cdot\hat{\kappa}\cdot|\tilde{\beta}|^{-1})^{1/\hat{p}}$, when $\hat{\beta} \neq 0$, $\tilde{\beta} \neq 0$, and $\mathcal{C} = \hat{p}$ for the 1- or ∞-norm in

$$\mathcal{D} \triangleq \| \tilde{\beta}\cdot Z_\alpha P^{-1} - \tilde{\alpha}\cdot Z_\beta P^{-1} \| ; \tag{24}$$

$\mathcal{C} = [\hat{p}(\hat{p}+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm ;

(ii) $\phi = \mathcal{C}\cdot\mathcal{D}\cdot\hat{\kappa}$, when $\hat{\beta} \neq 0$, $\tilde{\beta} = 0$, and $\mathcal{C} = 1$ for the 1- or ∞-norm, and $\mathcal{C} = \hat{p}^{\frac{1}{2}}$ for the 2- or F-norm;

(iii) $\phi = |\tilde{\alpha}|\cdot(\mathcal{C}\cdot\mathcal{D}\cdot\hat{\kappa}\cdot|\tilde{\alpha}|^{-1})^{1/\hat{q}}$ when $\hat{\beta} = 0$, $\tilde{\alpha} \neq 0$, and $\mathcal{C} = \hat{q}$ for the 1- or ∞-norm, and $\mathcal{C} = [\hat{q}(\hat{q}+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm;

and (iv) $\phi = \mathcal{C}\cdot\mathcal{D}\cdot\hat{\kappa}$ when $\hat{\beta} = 0$, $\tilde{\alpha} = 0$, and $\mathcal{C} = 1$ for the 1- or ∞-norm, and $\mathcal{C} = \hat{q}^{\frac{1}{2}}$ for the 2- or F-norm.

(Proof) Let $Q \triangleq \hat{X}_2^H \tilde{X}_1$. Equation (20) implies

$$\begin{cases} \hat{\Lambda}_{\alpha_1} P \tilde{\Lambda}_{\beta_1} - \hat{\Lambda}_{\beta_1} P \tilde{\Lambda}_{\alpha_1} = \hat{Y}_1^H D \triangleq \hat{D}_1 & \text{(25a)} \\ \hat{\Lambda}_{\alpha_2} Q \tilde{\Lambda}_{\beta_1} - \hat{\Lambda}_{\beta_2} Q \tilde{\Lambda}_{\alpha_1} = \hat{Y}_2^H D \triangleq \hat{D}_2 . & \text{(25b)} \end{cases}$$

From equation (25), $Q = T^{-1}(\hat{D}_2)$ and $P \cong I$ if Q is small enough. From equations (21) and (25a), one has

$$(\hat{\Lambda}_{\alpha_1} + Z_\alpha P^{-1}) P \tilde{\Lambda}_{\alpha_1} = (\hat{\Lambda}_{\beta_1} + Z_\beta P^{-1}) P \tilde{\Lambda}_{\beta_1} \tag{26}$$

and the matrices $Z_\alpha P^{-1}$ and $Z_\beta P^{-1}$ are now perturbations in the above GEVP to matrices $\hat{\Lambda}_{\alpha_1}$ and $\hat{\Lambda}_{\beta_1}$. Apply Theorem 1 to equation (26) and the corollary has been proved. ■

Corollary 3 will be required to prove Theorem 5 in section 4.

4. A Perturbation Theorem.

Consider a partitioning of the GEVP in equation (4):

$$A(X_1,X_2)\begin{bmatrix}\Lambda_{\beta_1} & O\\ O & \Lambda_{\beta_2}\end{bmatrix} = B(X_1,X_2)\begin{bmatrix}\Lambda_{\alpha_1} & O\\ O & \Lambda_{\alpha_2}\end{bmatrix}, \tag{27a}$$

and

$$\begin{bmatrix}\Lambda_{\beta_1} & O\\ O & \Lambda_{\beta_2}\end{bmatrix}\begin{bmatrix}Y_1^H\\ Y_2^H\end{bmatrix}A = \begin{bmatrix}\Lambda_{\alpha_1} & O\\ O & \Lambda_{\alpha_2}\end{bmatrix}\begin{bmatrix}Y_1^H\\ Y_2^H\end{bmatrix}B\ ; \tag{27b}$$

or

$$\begin{bmatrix}Y_1^H\\ Y_2^H\end{bmatrix}A(X_1,X_2) = \begin{bmatrix}\Lambda_{\alpha_1} & O\\ O & \Lambda_{\alpha_2}\end{bmatrix}, \tag{28a}$$

and

$$\begin{bmatrix}Y_1^H\\ Y_2^H\end{bmatrix}B(X_1,X_2) = \begin{bmatrix}\Lambda_{\beta_1} & O\\ O & \Lambda_{\beta_2}\end{bmatrix}. \tag{28b}$$

The columns of the matrices (X_1,X_2) and (Y_1,Y_2) are permutations of that of X and Y in equation (4) respectively, and thus the matrices on the RHS of equation (28) are block diagonal, with sub-blocks being permutations of those in Λ_α and Λ_β respectively in equation (4).

Applying Theorem 1 to the $n_1 \times n_1$ matrix pencil

$$Y_1^H(\tilde{A}-\lambda\tilde{B})X_1 \triangleq (\hat{A}-\lambda\hat{B}) = [(\Lambda_{\alpha_1}+Y_1^H\delta AX_1) - \lambda(\Lambda_{\beta_1}+Y_1^H\delta BX_1)],$$

one has

$$\min_i \left|\hat{\lambda}_{\alpha 1}\lambda_{\beta_{1i}} - \hat{\lambda}_{\beta 1}\lambda_{\alpha_{1i}}\right| \leq \hat{\phi}, \tag{29}$$

where $(\hat{\lambda}_{\alpha 1},\hat{\lambda}_{\beta 1}) \in \rho(\hat{A},\hat{B})$, $(\lambda_{\alpha_{ki}},\lambda_{\beta_{ki}}) \in \rho(\Lambda_{\alpha_k},\Lambda_{\beta_k})$ for $k = 1,2$; $\hat{\phi}$ is now defined similarly as ϕ in Theorem 1, with δA and δB replaced by $Y_1^H \delta AX_1$, and $Y_1^H \delta BX_1$ respectively, $\tilde{\kappa}$ by 1 (as the generalized eigenvectors of $(\Lambda_{\alpha_1} - \lambda\Lambda_{\beta_1})$ are now columns of I_{n_1}), and p and q by p_1 and q_1, the dimensions of the biggest Jordan blocks in Λ_{α_1} and Λ_{β_1} respectively.

However, one is interested in the perturbation of the generalized eigenvalues of the matrix pencil $(A-\lambda B)$, rather than that of $(\hat{A}-\lambda\hat{B})$. We can prove that $(\hat{\lambda}_{\alpha 1},\hat{\lambda}_{\beta 1})$ is approximately equal to some $(\tilde{\lambda}_{\alpha 1},\tilde{\lambda}_{\beta 1}) \in \rho(\tilde{A},\tilde{B})$, after giving up the freedom of the size of the perturbations δA and δB, under some mild restrictions on the partitioning in equations (27) and (28):-

<u>Lemma 4</u>. Let $\varepsilon = \max\{\|\delta A\|,\|\delta B\|\}$ be a small perturbation parameter and let $(\lambda_{\alpha 1},\lambda_{\beta 1}) \in \rho(\Lambda_{\alpha_1},\Lambda_{\beta_1})$ be perturbed to $(\tilde{\lambda}_{\alpha 1},\tilde{\lambda}_{\beta 1})$ and $(\hat{\lambda}_{\alpha 1},\hat{\lambda}_{\beta 1})$ respectively. Let $(\tilde{\Lambda}_{\alpha_1},\tilde{\Lambda}_{\beta_1})$ and $(\hat{\Lambda}_{\alpha_1},\hat{\Lambda}_{\beta_1})$ be the corresponding Kronecker canonical forms. Define the operator $\hat{T}$ by

$$\hat{T}(M) \triangleq (\Lambda_{\alpha_2}+Y_2^H\delta AX_2)M\tilde{\Lambda}_{\beta_1} - (\Lambda_{\beta_2}+Y_2^H\delta BX_2)M\tilde{\Lambda}_{\alpha_1} \ . \tag{30}$$

Let $\hat{T}$ be invertible and ε be small enough.

One has

$$\min_{(\hat{\lambda}_{\alpha},\hat{\lambda}_{\beta 1})} \left| \tilde{\lambda}_{\alpha 1}\hat{\lambda}_{\beta 1} - \tilde{\lambda}_{\beta 1}\hat{\lambda}_{\alpha 1} \right| \leqq \phi \ , \tag{31}$$

with $\phi = O(\varepsilon^{2/\hat{p}_1})$ or $O(\varepsilon^{2/\hat{q}_1})$, with $\hat{p}_1$ and $\hat{q}_1$ the dimensions of the biggest Jordan blocks in $\hat{\Lambda}_{\alpha_1}$ and $\hat{\Lambda}_{\beta_1}$ respectively.

(One can write out the complicated ϕ in equation (31) in detail, though it will not be worthwhile as it only appears in Theorem 5 as a higher order term.)

(Proof) Equations (27) and (28) imply that

$$\begin{cases} \begin{bmatrix} Y_1^H \\ Y_2^H \end{bmatrix} (A+\delta A)(X_1,X_2) = \begin{bmatrix} \Lambda_{\alpha_1} & O \\ O & \Lambda_{\alpha_2} \end{bmatrix} + \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix} , \\ \begin{bmatrix} Y_1^H \\ Y_2^H \end{bmatrix} (B+\delta B)(X_1,X_2) = \begin{bmatrix} \Lambda_{\beta_1} & O \\ O & \Lambda_{\beta_2} \end{bmatrix} + \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix} ; \end{cases} \qquad (32)$$

or

$$\left[\begin{bmatrix} \Lambda_{\alpha_1} & O \\ O & \Lambda_{\alpha_2} \end{bmatrix} + \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix} \right] \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \tilde{\Lambda}_{\beta_1} = \left[\begin{bmatrix} \Lambda_{\beta_1} & O \\ O & \Lambda_{\beta_2} \end{bmatrix} + \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix} \right] \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \tilde{\Lambda}_{\alpha_1} , \qquad (33)$$

for some matrices Z_1 and Z_2 .

Equation (33) implies

$$\begin{cases} (\Lambda_{\alpha_1}+E_1)Z_1\tilde{\Lambda}_{\beta_1} + E_2Z_2\tilde{\Lambda}_{\beta_1} = (\Lambda_{\beta_1}+F_1)Z_1\tilde{\Lambda}_{\alpha_1} + F_2Z_2\tilde{\Lambda}_{\alpha_1} , & (34a) \\ \hat{T}(Z_2) = F_3Z_1\tilde{\Lambda}_{\alpha_1} - E_3Z_1\tilde{\Lambda}_{\beta_1} . & (34b) \end{cases}$$

Equation (34b) implies

$$Z_2 = \hat{T}^{-1}(F_3Z_1\tilde{\Lambda}_{\alpha_1} - E_3Z_1\Lambda_{\beta_1}) = O(\varepsilon)$$

and substituting Z_2 back into equation (34a) to obtain an equation similar to that in (20), with the residual $= O(\varepsilon^2)$. One then proves the results in equation (31) by applying Corollary 3.

■

Note that the invertibility of $\hat{T}$ will be ensured if ε is small enough and the operator $\{\Lambda_{\alpha_2}(\cdot)\Lambda_{\beta_1} - \Lambda_{\beta_2}(\cdot)\Lambda_{\alpha_1}\}$ is invertible, i.e. $\rho(\Lambda_{\alpha_1},\Lambda_{\beta_1}) \cap \rho(\Lambda_{\alpha_2},\Lambda_{\beta_2}) = \emptyset$. Thus a reasonable "distance" between the two partitions in equations (27) and (28) should do the trick. (c.f.[4].) Note that the operator $\hat{T}^{-1}$ behaves like $(\lambda_{\alpha 1}\lambda_{\beta 2}-\lambda_{\beta 1}\lambda_{\alpha 2})^{-1}$, and can be analysed using the "diff" concept in [8].

One can then prove a perturbation theorem for the GEVP. Let the partitioning in equations (27) and (28) be chosen such that a group of multiple eigenvalues are grouped together in the pencil $(\Lambda_{\alpha_1} - \lambda \Lambda_{\beta_1})$. If one has a cluster of close eigenvalues instead of a set of pure multiple eigenvalues, the cluster should be averaged and considered multiple, with some small perturbations added to δA and δB. The conditions in Lemma 4 is then likely to be satisfied, unless the additional perturbations force the operator $\hat{T}$ in equation (30) singular. As in the case of ordinary eigenvalue problem, the perturbation of a set of pathologically close eigenvalues are most difficult to analyse [15].

One can prove the following theorem from Corollary 3 and Lemma 4:-

Theorem 5. With the assumptions in this section satisfied, one has

$$\min_{(\lambda_{\alpha 1}, \lambda_{\beta 1})} \left| \tilde{\lambda}_{\alpha 1} \lambda_{\beta 1} - \tilde{\lambda}_{\beta 1} \lambda_{\alpha 1} \right| \leq \hat{\phi} + O(\varepsilon^{2/r}) , \tag{35}$$

where $\Delta \triangleq \| \hat{\lambda}_{\beta 1} \cdot \delta A - \hat{\lambda}_{\alpha 1} \cdot \delta B \|$, $\hat{\kappa} = \|X_1\| \cdot \|Y_1\|$, (36)

$$\hat{\phi} = (\tilde{\lambda}_{\alpha 1} / \hat{\lambda}_{\alpha 1}) \cdot \phi, \text{ or } (\tilde{\lambda}_{\beta 1} / \hat{\lambda}_{\beta 1}) \cdot \phi, \tag{37}$$

(whichever is finite)

p_1 and q_1 are the dimensions of the biggest Jordan blocks in Λ_{α_1} and Λ_{β_1} respectively, and anologous to the four cases in Theorem 1:-

(i) $\phi = |\hat{\lambda}_{\beta 1}| \cdot (\mathscr{C} \cdot \Delta \cdot \hat{\kappa} \cdot |\hat{\lambda}_{\beta 1}|^{-1})^{1/p_1}$, when $\lambda_{\beta 1} \neq 0$, $\hat{\lambda}_{\beta 1} \neq 0$;

$\mathscr{C} = p_1$ for the 1- or ∞-norm, and

$\mathscr{C} = [p_1(p_1+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm;

$r = \hat{p}_1$, defined in Lemma 4.

(ii) $\phi = \mathscr{C} \cdot \Delta \cdot \hat{\kappa}$, when $\lambda_{\beta 1} \neq 0$, $\hat{\lambda}_{\beta 1} = 0$;

$\mathscr{C} = 1$ for the 1- or ∞-norm, and

$\mathscr{C} = p_1^{\frac{1}{2}}$ for the 2- or F-norm;

$r = \hat{q}_1$, defined in Lemma 4.

(iii) $\phi = |\hat{\lambda}_{\alpha 1}| \cdot (\mathcal{C} \cdot \Delta \cdot \hat{\kappa} \cdot |\hat{\lambda}_{\alpha 1}|^{-1})^{1/q_1}$, when $\lambda_{\beta 1} = 0, \hat{\lambda}_{\alpha 1} \neq 0$;

$\mathcal{C} = q_1$ for the 1- or ∞-norm, and

$\mathcal{C} = [q_1(q_1+1)/2]^{\frac{1}{2}}$ for the 2- or F-norm;

$r = \hat{p}_1$.

and (iv) $\phi = \mathcal{C} \cdot \Delta \cdot \hat{\kappa}$; when $\lambda_{\beta 1} = 0$, $\hat{\lambda}_{\alpha 1} = 0$;

$\mathcal{C} = 1$ for the 1- or ∞-norm, and

$\mathcal{C} = q_1^{\frac{1}{2}}$ for the 2- or F-norm;

$r = \hat{q}_1$.

(The above four cases represent the perturbation of the finite or infinite eigenvalue $(\lambda_{\alpha 1}, \lambda_{\beta 1})$ to the finite or infinite eigenvalue $(\hat{\lambda}_{\alpha 1}, \hat{\lambda}_{\beta 1})$.)

(Proof) From the argument in this section leading to equation (29) and Lemma 4, one has

$$|\tilde{\lambda}_{\alpha 1}\hat{\lambda}_{\beta 1} - \tilde{\lambda}_{\beta 1}\hat{\lambda}_{\alpha 1}| = O(\varepsilon^{\tau_1}) \tag{38a}$$

and

$$|\hat{\lambda}_{\alpha 1}\lambda_{\beta 1} - \hat{\lambda}_{\beta 1}\lambda_{\alpha 1}| = \phi = O(\varepsilon^{\tau_2}) \tag{38b}$$

for some $\tilde{\lambda}$, $\hat{\lambda}$ and λ's , and hopefully with $\tau_1 > \tau_2$.

In other words, equations (38) is equivalent to

$$\tilde{\lambda}_{\alpha 1}\hat{\lambda}_{\beta 1} = \tilde{\lambda}_{\beta 1}\hat{\lambda}_{\alpha 1} + O(\varepsilon^{\tau_1}) \tag{39a}$$

and

$$\hat{\lambda}_{\alpha 1}\lambda_{\beta 1} = \hat{\lambda}_{\beta 1}\lambda_{\alpha 1} \pm \phi \quad . \tag{39b}$$

One has to prove from equations (39) that

$$\tilde{\lambda}_{\alpha 1}\lambda_{\beta 1} = \tilde{\lambda}_{\beta 1}\lambda_{\alpha 1} \pm \hat{\phi} + O(\varepsilon^{\tau_1}) \quad . \tag{40}$$

Multiplying equation (39b) by $\tilde{\lambda}_{\alpha 1}$ and using equation (39a), one obtains

$$\hat{\lambda}_{\alpha 1}\tilde{\lambda}_{\alpha 1}\lambda_{\beta 1} - (\tilde{\lambda}_{\alpha 1}\hat{\lambda}_{\beta 1})\lambda_{\alpha 1} = \tilde{\lambda}_{\alpha 1}\cdot\phi$$

$$\Longrightarrow \hat{\lambda}_{\alpha 1}\tilde{\lambda}_{\alpha 1}\lambda_{\beta 1} - (\hat{\lambda}_{\alpha 1}\tilde{\lambda}_{\beta 1})\lambda_{\alpha 1} = 0(\tilde{\lambda}_{\alpha 1}\varepsilon^{\tau_1}) \pm \tilde{\lambda}_{\alpha 1}\cdot\phi \quad . \tag{41}$$

If $\hat{\lambda}_{\alpha 1} \neq 0$, equation (41) implies

$$\tilde{\lambda}_{\alpha 1}\lambda_{\beta 1} = \tilde{\lambda}_{\beta 1}\lambda_{\alpha 1} + 0[(\tilde{\lambda}_{\alpha 1}/\hat{\lambda}_{\alpha 1})\cdot\varepsilon^{\tau_1}] \pm (\tilde{\lambda}_{\alpha 1}/\hat{\lambda}_{\alpha 1})\cdot\phi$$

which implies equation (40).

If $\hat{\lambda}_{\alpha 1} = 0$, then equations (39) now read

$$\tilde{\lambda}_{\alpha 1}\hat{\lambda}_{\beta 1} = 0(\varepsilon^{\tau_1}) \tag{42a}$$

and

$$\hat{\lambda}_{\beta 1}\lambda_{\alpha 1} = \pm\phi = 0(\varepsilon^{\tau_2}) \quad . \tag{42b}$$

As $\hat{\lambda}_{\alpha 1} = 0$, $\hat{\lambda}_{\beta 1} \neq 0$ because of the regularity of pencils, and equations (42) imply

$$\tilde{\lambda}_{\alpha 1}\lambda_{\beta 1} = 0[(\lambda_{\beta 1}/\hat{\lambda}_{\beta 1})\cdot\varepsilon^{\tau_1}] \tag{43a}$$

and

$$\tilde{\lambda}_{\beta 1}\lambda_{\alpha 1} = \pm(\tilde{\lambda}_{\beta 1}/\hat{\lambda}_{\beta 1})\cdot\phi \quad . \tag{43b}$$

Subtracting the equation (43b) from (43a), one proves equation (40).

∎

Similar comments to Theorem 1 in section 2 apply for Theorem 5. Again, one has to partition the spectra in equations (27) and (28) with care, in accordance to the discussion prior to Theorem 5, to make sure that (i) the conditions of Corollary 3 and Lemma 4 are satisfied, and (ii) the RHS of inequality (35) is dominated by $\hat{\phi}$. Condition (ii) can be ensured by grouping multiple eigenvalues together carefully and adding perturbations to δA and δB. Note that only very special perturbations δA and δB will enable the RHS of inequality (35) be dominated by the 2nd term, when $(\lambda_{\alpha 1}, \lambda_{\beta 1})$ is perturbed from being a non-zero finite eigenvalue to a zero or infinite $(\hat{\lambda}_{\alpha 1}, \hat{\lambda}_{\beta 1})$, when $\hat{\phi} = 0$ because $\hat{\lambda}_{\alpha 1} = \tilde{\lambda}_{\beta 1} = 0$ or $\tilde{\lambda}_{\alpha 1} = \hat{\lambda}_{\beta 1} = 0$ (result of nearly singular matrix-pencils), or when $p_1, q_1 \leq \hat{p}_1/2, \hat{q}_1/2$ and the deficiency in eigenvectors becomes more severe after perturbations. Adding small perturbations to $\hat{A}$ and $\hat{B}$ should pull the GEVP of the matrix pencil $(\hat{A}-\lambda\hat{B})$ out of these pathological situations. The same perturbations can then be subtracted from $(\tilde{A}, \tilde{B})$ when perturbed from $(\hat{A}, \hat{B})$ (in Lemma 4). As a result, one can assume that the RHS in inequality (35) is always dominated by $\hat{\phi}$. In other words, $(\hat{A}, \hat{B})$ is only a bridge between (A, B) and $(\tilde{A}, \tilde{B})$, and perturbations can be added and subtracted so that the matrix pencil $(\hat{A}-\lambda\hat{B})$ is in desirable form.

For a non-defective p-multiple eigenvalue, similar to equation (16), one has

$$|\tilde{\lambda}_\beta \lambda_{\alpha_{ij}} - \tilde{\lambda}_\alpha \lambda_{\beta_{ij}}| / (\tilde{\nu} \nu_{ij}) \leq \hat{\kappa} \nu_{ij}^{-1} \varepsilon + O(\varepsilon^2) . \tag{44}$$

Inequality (44) is similar to what Stewart has derived in [9] [10]. (Stewart has a factor of p on the RHS of inequality (44).)

Similarly, results as in (44) can be obtained for special cases where the matrix pair (A, B) is definite [11] [12] or symmetric, from Theorem 5.

Finally, the conditioning of an individual eigenvalue will then be reflected by $\mathscr{C}\hat{\kappa}\nu_{ij}^{-1}$ or $(\mathscr{C}\hat{\kappa}\nu_{ij}^{-p_1})^{1/p_1}$. Again p_1 reflects the deficiency of the eigenvectors, $\hat{\kappa}$ the conditioning of the invariant subspaces, and ν_{ij} the distances between the matrix pencil $(A-\lambda B)$ and its nearest singular neighbour. Note that $\hat{\kappa} = \|x_1\| \cdot \|y_1\|$ for the case $n_1 = p_1 = 1$ and is a generalization of the sensitivity s_1 introduced by Wilkinson [15].

6. Conclusions.

A Bauer-Fike type exclusion theorem is proved for the GEVP. A further generalization is then proved to provide a perturbation analysis for the generalized eigenvalues of the GEVP. Perturbation results are obtained for all generalized eigenvalues, finite and infinite, defective or otherwise, using a chordal metric.

A generalized eigenvalue $(\lambda_\alpha, \lambda_\beta)$ will be better conditioned, if it is less defective (in terms of its generalized eigenvectors), its generalized invariant subspaces better conditioned, and $(|\lambda_\alpha|^2 + |\lambda_\beta|^2)$ further away from zero.

Although the exclusion theorem in Theorem 1 holds for any size of perturbations, the perturbation analysis in Theorem 5 only holds for small perturbations.

Note that the perturbation analysis for the invariant subspaces has been thoroughly investigated by Stewart [8], and recently by Sun [12] [13].

Finally, as in [13], the results in this paper can be applied to analyse the perturbation of generalized singular values.

References:-

[1] F. L. Bauer and C. T. Fike. "Norms and Exclusion Theorems." Numer. Math. 2 (1960), pp. 137-144.

[2] K.-w. E. Chu. "Generalizations of The Bauer-Fike Theorem." Numer. Anal. Rpt. NA/5/85, Univ. of Reading, Dept. of Maths., Reading, U.K., 1985.

[3] K.-w. E. Chu. "The Solution of the Matrix Equations AXB-CXD = E and (YA-DZ,YC-BZ) = (E,F)." Numer. Anal. Rpt. NA/10/85, Univ. of Reading, Dept. of Maths., Reading, U.K., 1985.

[4] L. Elsen and J. G. Sun. "Perturbation Theorems for The Generalized Eigenvalue Problem." Lin. Alg. Applic. 48 (1982), pp. 341-357.

[5] F. R. Gantmacher. "The Theory of Matrices, II." Chelsea, New York, 1960.

[6] G. H. Golub and C. Van Loan. "Matrix Computations." The Johns Hopkins Univ. Press, Baltimore, 1983.

[7] C. B. Moler and G. W. Stewart. "An Algorithm for The Generalized Matrix Eigenvalue Problem $Ax = \lambda Bx$." SIAM J. Numer. Anal. 10 (1973), pp. 241-256.

[8] G. W. Stewart. "On The Sensitivity of The Eigenvalue Problem $Ax = \lambda Bx$." SIAM J. Numer. Anal. 9(1972), pp. 669-686.

[9] G. W. Stewart. "Gershgorin Theory for The Generalized Eigenvalue Problem $Ax = \lambda Bx$." Math. Comp. 29(1975), pp. 600-606.

[10] G. W. Stewart. "Perturbation Theory for The Generalized Eigenvalue Problem." In "Recent Advances in Numerical Analysis," ed. C. de Boor and G. H. Golub, Academic Press, New York, 1978.

[11] G. W. Stewart. "Perturbation Bounds for The Definite Generalized Eigenvalue Problem." Lin. Alg. Applic. 23(1979), pp. 69-85.

[12] J. G. Sun. "The Perturbation Bounds for Eigenspaces of A Definite Matrix Pair." Numer. Math. 41 (1983), pp. 321-343.

[13] J. G. Sun. "Perturbation Analysis for The Generalized Singular Value Problem." SIAM J. Numer. Anal. 20 (1983), pp. 611-625.

[14] R. C. Ward. "The Combination Shift QZ Algorithm." SIAM J. Numer. Anal. 12 (1975), pp. 835-852.

[15] J. H. Wilkinson. "The Algebraic Eigenvalue Problem." Clarendon, Oxford, 1965.