Perturbation Theorems for the Joint Spectrum of Commuting Matrices: A Conservative Approach*

Ludwig Elsner
Fakultät für Mathematik
Universität Bielefeld
Postfach 10 01 31
33501 Bielefeld, Germany

Dedicated to Chandler Davis

Submitted by Rajendra Bhatia

ABSTRACT

It is shown that recent perturbation theorems for the joint spectrum of commuting matrices, which have been proved using Clifford-algebra tools, can be obtained and improved by classical means, as used in the case of the standard eigenvalue problem.

1. INTRODUCTION

The effect of a perturbation of a matrix on its spectrum has been investigated for a long time and is now well understood. See e.g. the monographs [2, 8]. In contrast, study of the behavior of joint eigenvalues of m-tuples of commuting matrices has started only recently.

We consider an m-tuple $A = (A_1, \ldots, A_m)$ of complex n-by-n matrices A_i. A joint eigenvalue of A is a vector $\lambda \in \mathbb{C}^m$, $\lambda = (\lambda_1, \ldots, \lambda_m)^T$, such that there exists a nonzero vector $x \in \mathbb{C}^n$ with $A_j x = \lambda_j x$, $j = 1, \ldots, m$. If the

*Parts of this paper were written during a stay at the University of Calgary. The author thanks the members of the Department of Mathematics and Statistics, in particular P. Lancaster and P. Binding, for their hospitality and support.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

A_i are commuting, then there is at least one joint eigenvalue. x is called a joint eigenvector, and the collection of the joint eigenvalues is called the joint spectrum and is denoted by $\mathrm{Sp}(A)$; see [3, 5–7].

The question studied here is how sensitive $\mathrm{Sp}(A)$ is to perturbations in A. The first results in this direction were obtained by Pryde [7], in the case that the unperturbed m-tuple $B = (B_1, \dots, B_m)$ has a basis of joint eigenvectors, i.e., all B_i can be simultaneously diagonalized. He uses the approach of [5] and [6], representing B by the Clifford operator $\mathrm{Cliff}(B)$ acting on a larger space. The results obtained are formally very similar to the Bauer-Fike theorem, where the norm of the perturbation is replaced by the norm of the representing Clifford operator. In the same vein, Bhatia and Bhattacharyya [3] proved a perturbation result generalizing a bound given by Henrici [4].

After introducing the relevant definitions of the Clifford operator, the underlying spaces, and the connection between the joint spectrum of A and the spectrum of $\mathrm{Cliff}(A)$ in Section 2, we outline the above-mentioned results in Sections 3 and 4.

After establishing bounds on $\|\mathrm{Cliff}(A)\|$ in Section 5, we derive our main results in Sections 6 and 7. We show that one can obtain stronger results with elementary tools. Some assumptions of [3] and [7] can be weakened considerably.

2. CLIFFORD ALGEBRAS AND THE CLIFFORD OPERATOR OF AN m-TUPLE

In [5, 6] Clifford algebras were used as a tool to study joint spectra. We denote by $R_{(m)}$ the Clifford algebra generated by R^m. This is an algebra in which R^m can be imbedded. It is generated by elements $e_1, \dots, e_m$ with relations $e_i e_j = -e_j e_i$, $i \neq j$, $e_i^2 = -1$, $i, j = 1, \dots, m$. Define $e_\phi = 1$, and for $S \subset \{1, \dots, m\}$, $S = \{s_1, \dots, s_k\}$, $s_1 < s_2 < \cdots < s_k$ define $e_S = e_{s_1} e_{s_2} \cdots e_{s_k}$. Observe that for $i \in \{1, \dots, m\}$ one has $e_{\{i\}} = e_i$. Then $R_{(m)} = \{\sum_S \lambda_S e_S : \lambda_S \in \mathbb{R}\}$ forms an associative algebra if we define the product of the 2^m basis elements in the canonical way using the generating relations, i.e.

$$e_S e_T = \Bigg(\prod_{\substack{s \in S,\, t \in T \\ s \leqslant t}} (-1) \Bigg) e_{S+T},$$

where $S + T$ is the symmetric sum

$$S + T := (S \cup T)/(S \cap T)$$

The tensor product

$$\mathbb{C}^n \otimes \mathbb{R}_{(m)} = \left\{ \sum_S x_S \otimes e_S,\ x_S \in \mathbb{C}^n \right\}$$

is a Hilbert space if one defines an inner product $\langle x, y\rangle = \langle \Sigma_S x_S \otimes e_S, \Sigma_S y_S \otimes e_S\rangle = \Sigma_S \langle x_S, y_S\rangle$, where $x_S, y_S \in \mathbb{C}^n$ and $\langle x_S, y_S\rangle$ is the usual inner product in $\mathbb{C}^n$, and the norm by $\|\Sigma_S x_S \otimes e_S\| = (\Sigma_S \|x_S\|^2)^{1/2}$.

Let M_n denote the space of n-by-n complex matrices. Then $M_n \otimes R_{(m)} = \{\Sigma_S A_S \otimes e_S,\ A_S \in M_n\}$ is also an algebra, a subalgebra of the space $L(\mathbb{C}^n \otimes R_{(m)})$, of linear mappings of $\mathbb{C}^n \otimes \mathbb{R}_{(m)}$ into itself, if one defines

$$\left(\sum_S A_S \otimes e_S \right)\left(\sum_T x_T \otimes e_T \right) = \sum_{S,T} A_S x_T \otimes e_S e_T .$$

We denote by $\|\Sigma_S A_S \otimes e_S\|$ the operator norm of $\Sigma A_S \otimes e_S$ considered as an endomorphism of the Hilbert space $\mathbb{C}^n \otimes \mathbb{R}_{(m)}$. It is easy to see that for $A = \Sigma A_S \otimes e_S$ the adjoint operator A^*, defined by the relation $\langle Ax, y\rangle = \langle x, A^* y\rangle$ for all $x, y \in \mathbb{C}^n \otimes \mathbb{R}_{(m)}$, is given by

$$A^* = \sum_S A_S^* \otimes \bar{e}_S ,$$

where $\bar{e}_S = \pm e_S$; the sign is chosen to get $e_S \bar{e}_S = 1$. In particular $\bar{e}_i = -e_i$.

For an m-tuple $A = (A_1, \dots, A_m)$ of not necessarily commuting matrices $A_i \in M_n$, define its Clifford operator Cliff$(A) \in M_n \otimes R_{(m)}$ by

$$\text{Cliff}(A) = i \sum_{j=1}^{m} A_j \otimes e_j .$$

The connection between the concept of joint eigenvalues of an m-tuple of commuting matrices and the Clifford operator is given by the following result [5, 7].

LEMMA 2.1. *If $A = (A_1, \dots, A_m)$, where the A_j's are commuting with real spectra, then a vector $\lambda \in \mathbb{R}^m$ is a joint eigenvalue of A iff 0 is an eigenvalue of*

$$\text{Cliff}(A - \lambda I) = i \sum_{j=1}^{m} (A_j - \lambda_j I) \otimes e_j .$$

3. BAUER-FIKE THEOREMS

A classical result due to Bauer and Fike [1] is the following:

THEOREM 3.1. *Let* $A, B \in M_n$, *with spectra* $\sigma(A), \sigma(B) = \{\mu_1, \dots, \mu_n\}$. *Let B be diagonalizable, i.e.,*

$$B = S\,\mathrm{diag}(\mu_i)\,S^{-1}$$

for some nonsingular matrix S. Then for any $\lambda \in \sigma(A)$

$$\min_i |\lambda - \mu_i| \leqslant \|A - B\|\,\|S\|\,\|S^{-1}\|.$$

Here and in the following $\|\ \|$ is the spectral norm (or operator norm).

Using Lemma 2.1, Pryde obtained two analogous results for commuting m-tuples (see [7]). In contrast to the classical Bauer-Fike theorem, it is necessary to distinguish the cases of real and complex spectrum.

THEOREM 3.2. *Let* $A = (A_1, \dots, A_m)$, $B = (B_1, \dots, B_m)$ *be m-tuples of commuting matrices in* M_n *with real spectra. Let the* B_j*'s be simultaneously diagonalized by S. Then for any joint eigenvalue* λ *of A there exists a joint eigenvalue* μ *of B such that*

$$\|\lambda - \mu\| \leqslant \|S\|\,\|S^{-1}\|\,\|\mathrm{Cliff}(A - B)\|.$$

Here $\|\lambda - \mu\|$ is the Euclidean norm in R^m.

For describing the result in the case of complex spectra we have to introduce the concept of a partition of A. One can decompose $A_j = A_{1j} + iA_{2j}$ so that $\pi(A) = (A_{11}, \dots, A_{1m}, A_{21}, \dots, A_{2m})$ is a $2m$-tuple of commuting matrices with real spectra. $\pi(A)$ is called a partition of A. If the B_j's can be simultaneously diagonalized, there is exactly one partition $\pi(B)$ of B such that the matrices B_{ij}, $i = 1, 2$, $j = 1, \dots, m$, can be simultaneously diagonalized. This $\pi(B)$ is called the semisimple partition of B. By applying Theorem 3.2 to $\pi(A)$, $\pi(B)$, Pryde [7] obtained the following result:

THEOREM 3.3. *Let* $A = (A_1, \dots, A_m)$, $B = (B_1, \dots, B_m)$ *be two m-tuples of commuting matrices in* M_n. *Let the* B_j*'s be diagonalizable. Let* $\pi(A)$ *be a partition of A,* $\pi(B)$ *be the semisimple partition of B, and S a matrix simultaneously diagonalizing the components of* $\pi(B)$. *Then for any joint* *eigenvalue* λ *of* A *there exists a joint eigenvalue* μ *of* B *such that*

$$\|\lambda - \mu\| \leqslant \|S\|\,\|S^{-1}\|\,\|\mathrm{Cliff}(\pi(A) - \pi(B))\|.$$

Here $\|\lambda - \mu\|$ *is the Euclidean norm in* C^m, *and* $\|\mathrm{Cliff}(\pi(A) - \pi(B))\|$ *the operator norm in* $M_n \otimes R_{(2m)}$.

4. HENRICI-TYPE THEOREMS

One version of the classical Henrici theorem on the perturbation of spectra is the following (see [4]):

THEOREM 4.1. *Let* $A, B \in M_n$, *with spectra* $\sigma(A)$, $\sigma(B)$ *respectively. Then for any* $\lambda \in \sigma(A)$ *we have*

$$\min_{\mu \in \sigma(B)} |\lambda - \mu| \leqslant S_n(\Delta(B), \|A - B\|) \tag{4.1}$$

Here, $\Delta(B)$ is the $\|\ \|$-departure from normality, as defined below, and $S_n(\Delta, r)$ is the spectral radius $\rho(C)$ of the nonnegative n-by-n matrix $C(\Delta, r)$:

$$S_n(\Delta, r) = \rho(C(\Delta, r)), \qquad C(\Delta, r) = \begin{pmatrix} 0 & \Delta & & \cdots & 0 \\ & & \Delta & & \\ & 0 & & \ddots & \\ & & & & \Delta \\ r & & \cdots & & r \end{pmatrix}, \tag{4.2}$$

$\Delta, r \geqslant 0$. Another way to express $S_n(\Delta, r)$ for $r > 0$, $\Delta > 0$ is the following. If $g = g_n(\Delta/r)$ is the unique positive solution of

$$g + g^2 + \cdots + g^n = \frac{\Delta}{r}, \tag{4.3}$$

then

$$S_n(\Delta, r) = \frac{\Delta}{g_n(\Delta/r)}.$$

Any $B \in M_n$ can be transformed to upper triangular form by a unitary similarity, i.e.

$$U^*BU = \operatorname{diag}(\beta_i) + N, \tag{4.4}$$

N strictly upper triangular and U unitary. Then

$$\Delta(B) = \min\{\|N\|\},$$

where the min is taken over all N appearing in a decomposition like (4.4).

Now consider m-tuples of commuting matrices $A = (A_1, \dots, A_m)$. It is well known that they can be simultaneously triangularized by a unitary transformation

$$U^*A_jU = \Lambda_j + N_j, \qquad j = 1, \dots, m, \tag{4.5}$$

where N_j is strictly upper triangular. Define $N = (N_1, \dots, N_m)$, and

$$\Delta(A) = \min\{\|\operatorname{Cliff}(N)\|\}, \tag{4.6}$$

where again the minimum is taken over all such N's. Bhatia and Bhattacharyya obtained the following results:

THEOREM 4.2. *Let $A = (A_1, \dots, A_m)$ and $B = (B_1, \dots, B_m)$ be two commuting m-tuples of matrices in M_n with real spectra. For any joint eigenvalue λ of A there exists a joint eigenvalue μ of B such that*

$$\|\lambda - \mu\| \leqslant S_n(\Delta(B), \|\operatorname{Cliff}(A - B)\|). \tag{4.7}$$

THEOREM 4.3. *Let $A = (A_1, \dots, A_m)$, $B = (B_1, \dots, B_m)$ be two commuting m-tuples of matrices in M_n with partitions $\pi(A)$, $\pi(B)$. For any joint eigenvalue λ of A there exists a joint eigenvalue of B such that*

$$\|\lambda - \mu\| \leqslant S_n(\Delta(\pi(B)), \|\operatorname{Cliff}(\pi(A) - \pi(B))\|). \tag{4.8}$$

5. BOUNDS FOR THE CLIFFORD OPERATOR

One disadvantage of the previous theorems is the occurrence of the quantity $\|\operatorname{Cliff}(C_1, C_2, \dots, C_m)\|$, which has been determined as a simple function of the C_i's only for $m = 2$ [7]. Pryde showed

$$\|\mathrm{Cliff}(C_1, C_2)\| = \max(\|C_1 + iC_2\|, \|C_1 - iC_2\|).$$

We will give here some computable bounds.

Let $A = (A_1, \dots, A_m)$ be an m-tuple of n-by-n matrices. We may also identify A with an operator

$$\tilde{A} : \mathbb{C}^n \to \mathbb{C}^{nm},$$

$$x \to \begin{pmatrix} A_1 x \\ A_2 x \\ \vdots \\ A_m x \end{pmatrix}, \qquad \tilde{A} = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}.$$

Then the operator norm of $\tilde{A}$ is given by

$$\|\tilde{A}\|^2 = \|\tilde{A}^* \tilde{A}\| = \left\| \sum_{i=1}^{m} A_i^* A_i \right\|.$$

We have

LEMMA 5.1. *Let $A = (A_1, \dots, A_m)$ be an m-tuple of $n \times n$ matrices. Then*

$$\|\tilde{A}\| \leqslant \|\mathrm{Cliff}(A)\| \leqslant \sqrt{1 + \frac{m(m-1)}{2}}\, \|\tilde{A}\|. \tag{5.1}$$

Proof. The lower bound can be found in [5, Proposition 3.5], where also an upper bound

$$\|\mathrm{Cliff}(A)\|^2 \leqslant \left\| \sum_{j=1}^{m} A_j^* A_j \right\| + \sum_{j<k} \|A_j^* A_k - A_k^* A_j\|$$

is provided. But

$$\|A_j^* A_k - A_k^* A_j\| \leqslant \|A_j^* A_j + A_k^* A_k\|,$$

implying the upper bound in (5.1). ■

Another upper bound is given by

$$\|\mathrm{Cliff}(A)\| \leqslant \sum_{j=1}^{m} \|A_j\|, \tag{5.2}$$

but (5.1) and (5.2) seem not to be compatible.

6. IMPROVING THE BAUER-FIKE THEOREMS

In this section we generalize and improve Theorems 4.2 and 4.3. The l_p-norm of $x \in \mathbb{C}^n$ is denoted and defined by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \qquad 1 \leqslant p \leqslant \infty,$$

where we suppress the dependence on the dimension n. The associated operator norm for an r-by-s matrix A is given by

$$\|A\|_p = \max\{\|Ax\|_p, \|x\|_p \leqslant 1\}.$$

Note that $\|A\|_2 = \|A\| =$ spectral norm. We will show

THEOREM 6.1. *Let $A = (A_1, \ldots, A_m)$, $B = (B_1, \ldots, B_m)$ be two m-tuples of complex n-by-n matrices. Assume that the B_i can be simultaneously diagonalized by a nonsingular matrix S.*

Then for any joint eigenvalue $\lambda \in \mathrm{Sp}(A)$ there exists $\mu \in \mathrm{Sp}(B)$ such that

$$\|\lambda - \mu\|_p \leqslant \|S\|_p \|S^{-1}\|_p \|\tilde{A} - \tilde{B}\|_p. \tag{6.1}$$

Note that the B_i's are commuting, but the A_i's need not be. In general, however, $\mathrm{Sp}(A)$ can be void, in which case the statement of Theorem 6.1 is trivially true. Recall that $\tilde{A}$, $\tilde{B}$ are defined in Section 5.

Proof. As $\lambda \in \mathbb{C}^m$ is a joint eigenvalue of A, and $B_i S = S\,\mathrm{diag}(\mu_1^{(i)}, \ldots, \mu_n^{(i)})$, $i = 1, \ldots, m$, we have

$$(A_i - B_i)x = (\lambda_i I - B_i)x$$

for some $x \neq 0$, and with $z = S^{-1}x$

$$S^{-1}(A_i - B_i)Sz = \left[\lambda_i I - \operatorname{diag}(\mu_k^{(i)})\right]z = w_i \qquad \text{(say)},$$

or in matrix terms

$$\begin{pmatrix} S^{-1} & & & 0 \\ & S^{-1} & & \\ & & \ddots & \\ 0 & & & S^{-1} \end{pmatrix} \begin{pmatrix} A_1 - B_1 \\ \vdots \\ A_m - B_m \end{pmatrix} Sz$$

$$= \begin{pmatrix} \operatorname{diag}(\lambda_1 - \mu_k^{(1)}) \\ \vdots \\ \operatorname{diag}(\lambda_m - \mu_k^{(m)}) \end{pmatrix} z = w \qquad \text{(say)}. \tag{6.2}$$

Then on the one side

$$\|w\|_p^p = \sum_{i=1}^m \left(\sum_{j=1}^n |z_j|^p |\lambda_i - \mu_j^{(i)}|^p \right) = \sum_{j=1}^n |z_j|^p \left(\sum_{i=1}^m |\lambda_i - \mu_j^{(i)}|^p \right)$$

$$\geqslant \|z\|_p^p \operatorname*{Min}_j \sum_{i=1}^m |\lambda_i - \mu_j^{(i)}|^p = \|z\|_p^p \operatorname*{Min}_j \|\lambda - \mu^{(j)}\|_p^p.$$

[Observe here that $\mu^{(j)} = (\mu_j^{(1)}, \ldots, \mu_j^{(m)})^T$ is a joint eigenvalue of B.]

On the other side

$$\|w\|_p \leqslant \|S\|_p \|S^{-1}\|_p \|\tilde{A} - \tilde{B}\|_p \|z\|_p.$$

These two inequalities imply (6.1). ∎

REMARKS. In particular, for $p = 2$, one has

$$\min_{\mu \in \mathrm{Sp}(B)} \|\lambda - \mu\| \leqslant \|S\| \, \|S^{-1}\| \, \|\tilde{A} - \tilde{B}\|, \tag{6.3}$$

which, for real spectra, by (5.1) improves Theorem 3.2. If $\lambda = (\lambda_1, \ldots, \lambda_m) \in C^m$ is a joint eigenvalue of A, we can also find

$$\pi(A) = (A_{11}, \ldots, A_{1m}, A_{21}, \ldots, A_{2m}), \qquad A_j = A_{1j} + iA_{2j},$$

$$A_j x = \lambda_j x, \quad A_{1j} x = (\operatorname{Re} \lambda_j) x, \quad A_{21} x = (\operatorname{Im} \lambda_j) x, \qquad j = 1, \ldots, m,$$

i.e., (Re λ, Im λ) is a joint eigenvalue of $\pi(A)$. Applying Theorem 6.1, now, to $(\pi(A), \pi(B))$, we get

$$\min_{\mu \in \mathrm{Sp}(B)} \|\lambda - \mu\| \leqslant \|S\|\,\|S^{-1}\|\left\|\widetilde{\pi(A)} - \widetilde{\pi(B)}\right\|, \tag{6.4}$$

which again improves Theorem 3.3.

It is however not clear whether the bound (6.3) is better than that provided by Theorem 3.3. For $m = 1$ this is true, as

$$\|\mathrm{Cliff}(A + iB)\| = \|A + iB\| \leqslant \max(\|A + iB\|, \|A - iB\|)$$
$$= \|\mathrm{Cliff}(A, B)\|.$$

For the last equality see Section 5. In general we have only

$$\|\tilde{A} - \tilde{B}\| \leqslant \sqrt{2}\left\|\widetilde{\pi(A)} - \widetilde{\pi(B)}\right\| \tag{6.5}$$

7. IMPROVING THE HENRICI-TYPE THEOREMS

Before we formulate the result, we have to define the concept of "departure from normality" used here. As already mentioned in Section 4, for a given m-tuple $B = (B_1, \dots, B_m)$ of commuting n-by-n matrices, there exists a (not uniquely defined) unitary U such that

$$U^*B_iU = M_i + N_i, \qquad i = 1, \dots, m, \tag{7.1}$$

with strictly upper triangular N_i and $M_i = \mathrm{diag}(\mu_1^{(i)}, \dots, \mu_n^{(i)})$. For the m-tuple $N = (N_1, \dots, N_m)$ we define $\tilde{N}$ as in Section 5 and

$$\tilde{\Delta}(B) = \min\|\tilde{N}\|, \tag{7.2}$$

where the minimum is taken over all such N's.

Observe that by Lemma 5.1

$$\tilde{\Delta}(B) \leqslant \Delta(B). \tag{7.3}$$

THEOREM 7.1. *Let $A = (A_1, \dots, A_m)$, $B = (B_1, \dots, B_m)$ be two m-tuples of complex n-by-n matrices. Assume that the B_i's are commuting.*

Then for any joint eigenvalue of A there exists a joint eigenvalue μ of B such that

$$\|\lambda - \mu\| \leqslant S_n\big(\tilde{\Delta}(B), \|\tilde{A} - \tilde{B}\|\big). \tag{7.4}$$

Note that the A_i's need not be commuting; however, Sp(A) may be void. S_n has been defined in Section 4.

Proof. Let $(\lambda_1, \dots, \lambda_m)$ be a joint eigenvalue of A, $A_i x = \lambda_i x$, $i = 1, \dots, m$, $x \neq 0$. Then from (7.1), and $z = U^* x$

$$U^*(B_j - A_j)Uz = (M_j - \lambda_j I + N_j)z$$

or

$$\begin{pmatrix} U^* & & & 0 \\ & U^* & & \\ & & \ddots & \\ 0 & & & U^* \end{pmatrix} \begin{pmatrix} B_1 - A_1 \\ \vdots \\ B_m - A_m \end{pmatrix} U_z = \left[\begin{pmatrix} M_1 - \lambda_1 I \\ \vdots \\ M_m - \lambda_m I \end{pmatrix} + \begin{pmatrix} N_1 \\ \vdots \\ N_m \end{pmatrix} \right] z$$

$$= (\tilde{D} + \tilde{N})z. \tag{7.5}$$

Observe that the columns of $\tilde{D}$ are orthogonal and the ith column has the Euclidean length

$$\alpha_i = \left(\sum_{j=1}^{m} |\lambda_j - \mu_i^{(j)}|^2 \right)^{1/2} = \|\lambda - \mu^{(i)}\|, \tag{7.6}$$

where the vector $\mu^{(i)} = ((\mu_i^{(1)}, \dots, \mu_i^{(m)})^T$ is a joint eigenvalue of B. Hence we may write $\tilde{D} = VD$, where $D = \operatorname{diag}(\alpha_i)$ and V is orthonormal, $V^*V = I_n$. Premultiplying (7.5) by V^*, we get

$$V^* \operatorname{diag}(U^*, \dots, U^*)\,(\tilde{B} - \tilde{A})Uz = (D + V^*\tilde{N})z = (D + N)z. \tag{7.7}$$

Here $N = V^*\tilde{N}$ is strictly upper triangular. If $\alpha_i = 0$ for some i, then by (7.6) there is nothing to prove; otherwise, we have $C = D + N$ invertible and from (7.7)

$$\left\| V^* \operatorname{diag}(U^*, \ldots, U^*)(\tilde{B} - \tilde{A})U \right\| = \|\tilde{B} - \tilde{A}\| \geq \left\| (D + N)^{-1} \right\|^{-1} \quad (7.8)$$

We now proceed as in the original proof of Henrici. As $N^n = 0$, we have $(D+N)^{-1} = D^{-1}\sum_{\nu=0}^{n-1}(ND^{-1})^{\nu}$ and hence with $\delta = \|D^{-1}\| = 1/\min \alpha_i$,

$$\left\| (D+N)^{-1} \right\| \leq \frac{1}{\|N\|} \sum_{\nu=1}^{n} (\|N\|\delta)^{\nu}.$$

This, together with (7.8), gives

$$\gamma = \frac{\|N\|}{\|\tilde{B} - \tilde{A}\|} \leq \sum_{\nu=1}^{n} (\delta\|N\|)^{\nu},$$

from which we get $\delta\|N\| \geq g_n(\gamma)$, or

$$\delta^{-1} \leq \frac{\|N\|}{g_n(\gamma)} = S_n(\|N\|, \|\tilde{A} - \tilde{B}\|).$$

Taking into account that $\|N\| = \|V^*\tilde{N}\| = \|\tilde{N}\|$ can be chosen as $\tilde{\Delta}(B)$ and using $\min\|\lambda - \mu^{(i)}\| = \delta^{-1}$, this last inequality yields (7.4). ∎

REMARK. As noted by H. Schneider, we need only that the matrices B_i are simultaneously upper triangularizable. Then also (7.1) holds, and defines m-tuples $\mu^{(i)} = (\mu_i^{(1)}, \ldots, \mu_i^{(m)})$, $i = 1, \ldots, n$. This coupling of the eigenvalues is unique, i.e. does not depend on the triangularizing similarity transformation, as can be readily seen. It coincides with the concept of "joint eigenvalue," where applicable.

The remarks following Theorem 6.1 apply here also. As $S_n(\Delta, r)$ is obviously strictly monotonic in Δ and r and $\|\tilde{A} - \tilde{B}\| \leq \|\text{Cliff}(A - B)\|$, $\tilde{\Delta}(B) \leq \Delta(B)$ [see (7.3)], we find that for real spectra the bound (7.4) is never worse than the bound (4.7).

In the case of complex eigenvalues we find, by applying Theorem 7.1 to $\pi(A)$ and $\pi(B)$, that both

$$\|\lambda - \mu\| \leq S_n(\tilde{\Delta}(B), \|\tilde{A} - \tilde{B}\|)$$

and

$$\|\lambda - \mu\| \leqslant S_n\left(\bar{\Delta}(\pi(B)), \left\|\widetilde{\pi(A)} - \widetilde{\pi(B)}\right\|\right)$$

hold.

REFERENCES

1 F. L. Bauer and C. T. Fike, Norms and exclusion theorems, *Numer. Math.* 2:137–141 (1960).

2 R. Bhatia, *Perturbation Bounds for Matrix Eigenvalues*, Longman Scientific & Technical, Essex, England, 1987.

3 R. Bhatia and T. Bhattacharyya, A Henrici theorem for joint spectra of commuting matrices, *Proc. Amer. Math. Soc.*, to appear.

4 P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices, *Numer. Math.* 4:24–39 (1962).

5 A. McIntosh and A. Pryde, A functional calculus for several commuting operators, *Indiana Univ. Math. J.* 36:421–439 (1987).

6 A. McIntosh, A. Pryde, and W. Ricker, Comparison of joint spectra for certain classes of commuting operators, *Studia Math.* 88:23–36 (1988).

7 A. Pryde, A Bauer -Fike theorem for the joint spectrum of commuting matrices, *Linear Algebra Appl.* 173:221–230 (1992).

8 G. W. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic, New York, 1990.

Received 20 March 1993; final manuscript accepted 26 April 1993