

## A HENRICI THEOREM FOR JOINT SPECTRA OF COMMUTING MATRICES

RAJENDRA BHATIA AND TIRTHANKAR BHATTACHARYYA

(Communicated by Palle E. T. Jorgensen)

**ABSTRACT.** A version of Henrici's classical perturbation theorem for eigenvalues of matrices is obtained for joint spectra of commuting tuples of matrices. The approach involves Clifford algebra techniques introduced by McIntosh and Pryde.

### 1. INTRODUCTION

Perturbation bounds for eigenvalues of matrices have a long history and several significant results concerning them are known [3].

For commuting tuples of operators the concept of joint spectrum has been developed in several important papers over the last twenty years (see [10] for a recent discussion). However, not many perturbation inequalities seem to be known in this case. Davis [6] drew special attention to this problem and its importance; after that McIntosh and Pryde [10] introduced a novel idea, the use of Clifford algebras, to develop a functional calculus for commuting tuples of operators and used this to extend earlier perturbation results from [4]. This approach was developed further by them and Ricker [11].

In two recent papers [13, 14] Pryde has initiated an interesting program: using the ideas of Clifford analysis to generalize some classical perturbation inequalities for single matrices to the case of joint spectra of commuting tuples of matrices. In [13] he generalizes the classical Bauer-Fike Theorem from single matrices to commuting tuples. In this note we obtain a similar extension of a well-known theorem of Henrici [8]. We follow the ideas of Pryde [13]. We must emphasize that attempts to obtain similar generalizations of other inequalities [14] run into difficulties and stringent conditions need to be imposed. Thus it would be of interest to find out which of the classical 'one variable' theorems can be generalized to the 'several variable' case, which fail to have generalizations and which are true in modified forms. The present note is of interest in this context.

---

Received by the editors August 12, 1991.

1991 *Mathematics Subject Classification*. Primary 15A42, 47A10, 47A55; Secondary 15A66.

*Key words and phrases*. Clifford algebras, commuting tuples of matrices, joint spectrum, spectral variation, Henrici's theorem, measure of nonnormality.

The first author thanks the DAE for a grant.

Other authors with different motivation have also obtained extensions of some classical spectral inequalities from the case of one operator to that of commuting tuples (see, e.g., [12]).

To state the classical Henrici Theorem we need to define a *measure of non-normality* of an  $n$  by  $n$  complex matrix. Any such matrix  $A$  can be reduced to an upper triangular form  $T$  by a unitary conjugation, i.e., there exists a unitary matrix  $U$  and an upper triangular matrix  $T$  such that  $U^*AU = T$ . Further, writing  $T = \Lambda + N$ , where  $\Lambda$  is a diagonal matrix and  $N$  is a strictly upper triangular matrix, we have

$$(1.1) \quad U^*AU = T = \Lambda + N.$$

Of course, neither  $U$  nor  $T$  are uniquely determined. The matrix  $\Lambda$  has as its diagonal entries the eigenvalues of  $A$ . The matrix  $A$  is normal iff the part  $N$  in any decomposition (1.1) of  $A$  is zero. Given a norm  $\nu$  on matrices the  $\nu$  *measure of nonnormality* can be defined as

$$(1.2) \quad \Delta_\nu(A) = \inf \nu(N),$$

where the infimum is taken over all  $N$  occurring in decomposition (1.1) of  $A$ .  $A$  is normal iff  $\Delta_\nu(A) = 0$ .

Identifying  $A$  as usual with an operator on the Euclidean space  $\mathcal{E}^n$  with the Euclidean vector norm  $\|\cdot\|$ , we define the *operator norm* of  $A$  as

$$(1.3) \quad \|A\| = \sup_{\|x\|=1} \|Ax\|.$$

This norm will be of special interest to us.

We then have

**Theorem 1.1** (Henrici). *Let  $A$  be a nonnormal matrix, and let  $B$  be any other matrix,  $B \neq A$ . Let  $\nu$  be any norm majorizing the operator norm. Let*

$$(1.4) \quad y = \frac{\Delta_\nu(A)}{\nu(B - A)}$$

and let  $g_n(y)$  be the unique positive solution of

$$(1.5) \quad g + g^2 + \cdots + g^n = y.$$

Then for each eigenvalue  $\beta$  of  $B$  there exists an eigenvalue  $\alpha$  of  $A$  such that

$$(1.6) \quad |\alpha - \beta| \leq \frac{y}{g_n(y)} \nu(B - A).$$

This theorem can be stated equivalently in the following way. For a fixed eigenvalue  $\beta$  of  $B$  let

$$(1.7) \quad \delta = \min |\alpha - \beta|,$$

where  $\alpha$  varies over all eigenvalues of  $A$ . Then

$$(1.8) \quad \frac{\delta}{1 + \Delta_\nu(A)/\delta + \cdots + \Delta_\nu^{n-1}(A)/\delta^{n-1}} \leq \nu(B - A)$$

(see [15, p. 172] for this formulation and its proof).

We will restrict ourselves to the operator norm and prove a version of the above inequality for the joint spectra of commuting  $m$ -tuples of matrices. The formulation of our result requires some facts from Clifford algebras which are explained in §2. Perturbation theorems are stated and proved in subsequent sections.

## 2. PRELIMINARIES

For the convenience of the reader, we briefly state some facts about joint spectra and Clifford algebras. Further details may be found in [1, 10, 11].

Let  $A = (A_1, \dots, A_m)$  be an  $m$ -tuple of commuting matrices. A *joint eigenvalue* of  $A$  is an element  $\lambda = (\lambda_1, \dots, \lambda_m)$  of  $\mathcal{E}^m$  corresponding to which there exists a nonzero vector  $x \in \mathcal{E}^n$  such that  $A_j x = \lambda_j x$  for all  $j = 1, \dots, m$ . Such an  $x$  is called a *joint eigenvector* of  $A$ . The collection of all joint eigenvalues of  $A$  is called the *joint spectrum* of  $A$  and is denoted by  $\sigma_{pt}(A)$ . The set  $\sigma_{pt}(A)$  is nonempty and has at most  $n$  distinct points. For a single matrix  $T$  we will denote its usual spectrum, i.e., the set of eigenvalues, by  $\sigma(T)$ . If  $A_1, \dots, A_m$  commute and each of them has real spectrum then it is shown in [10] that

$$(2.1) \quad \sigma_{pt}(A) = \left\{ \lambda \in \mathcal{R}^m : 0 \in \sigma \left( \sum (A_j - \lambda_j I)^2 \right) \right\}.$$

Let  $\mathcal{R}_{(m)}$  be the Clifford algebra generated by  $\mathcal{R}^m$ ;  $\mathcal{R}_{(m)}$  is a  $2^m$ -dimensional linear space. Its basis elements can be indexed by subsets of the set  $M = \{1, \dots, m\}$ . Let  $\mathcal{E} = \{e_S : S \subset M\}$  be a basis for  $\mathcal{R}_{(m)}$ .  $\mathcal{R}_{(m)}$  is made into an algebra by defining a multiplication for these basis elements by

$$(2.2) \quad e_S e_T = \prod_{s \in S, t \in T} (s, t) e_{S+T},$$

where  $(s, t) = +1$  if  $s < t$  and  $-1$  if  $s \geq t$ , and  $S + T$  is the subset of  $M$  consisting of integers which are either in  $S$  or in  $T$  but not in both. In particular,  $e_\phi$  is the identity for this algebra and

$$(2.3) \quad e_{\{j\}}^2 = -e_\phi, \quad e_{\{i\}} e_{\{j\}} = -e_{\{j\}} e_{\{i\}}, \quad i, j = 1, \dots, m,$$

and if

$$(2.4) \quad S = \{i_1, \dots, i_k\}, \quad i_1 < i_2 < \dots < i_k, \quad \text{then } e_S = e_{\{i_1\}} \cdots e_{\{i_k\}}.$$

Let  $\{e_1, \dots, e_m\}$  be the standard basis of  $\mathcal{R}^m$ . Identifying  $e_j$  with  $e_{\{j\}}$ ,  $\mathcal{R}^m$  can be thought of as a subspace of  $\mathcal{R}_{(m)}$ . So  $(x_1, \dots, x_m)$  is identified with  $\sum x_j e_{\{j\}}$ . Since

$$(2.5) \quad \left( \sum x_j e_{\{j\}} \right) \left( \sum x_j e_{\{j\}} \right) = \left( - \sum x_j^2 \right) e_\phi$$

and  $e_\phi$  is the identity of  $\mathcal{R}_{(m)}$ , any nonzero element of  $\mathcal{R}^m$  is invertible in  $\mathcal{R}_{(m)}$ .

For any two elements  $\lambda = \sum \lambda_S e_S$  and  $\mu = \sum \mu_S e_S$  of  $\mathcal{R}_{(m)}$ , let  $\langle \lambda, \mu \rangle = \sum \lambda_S \mu_S$ . This defines an inner product in which the basis  $\{e_S : S \subset M\}$  is orthonormal.

Let  $X$  be a finite-dimensional vector space over  $\mathcal{E}$ . The tensor product  $X \otimes \mathcal{R}_{(m)}$  is a finite-dimensional vector space whose elements can be represented as  $\sum x_S \otimes e_S$ . If  $X$  has an inner product  $\langle \cdot, \cdot \rangle$ , the space  $X \otimes \mathcal{R}_{(m)}$  naturally inherits it:

$$(2.6) \quad \left\langle \sum x_S \otimes e_S, \sum y_S \otimes e_S \right\rangle = \sum \langle x_S, y_S \rangle.$$

Let  $\mathcal{M}_n$  be the algebra of all  $n$  by  $n$  matrices. Let  $\{A_S : S \subset M\}$  be any collection of  $2^m$  matrices from  $\mathcal{M}_n$ . Then  $\sum A_S \otimes e_S$  is an element of  $\mathcal{M}_n \otimes \mathcal{R}_{(m)}$ . It can be thought of as a linear operator on  $\mathcal{E}^n \otimes \mathcal{R}_{(m)}$  if we define its action on  $\sum \lambda_T \otimes e_T$  by

$$(2.7) \quad \left( \sum_S A_S \otimes e_S \right) \left( \sum_T \lambda_T \otimes e_T \right) = \sum_{S,T} A_S \lambda_T \otimes e_S e_T.$$

So  $\mathcal{M}_n \otimes \mathcal{R}_{(m)}$  is a subalgebra of the algebra of all linear operators on  $\mathcal{E}^n \otimes \mathcal{R}_{(m)}$  and  $\mathcal{M}_n$  is a subalgebra of  $\mathcal{M}_n \otimes \mathcal{R}_{(m)}$  by the identification of the matrix  $A$  with the operator  $A \otimes e_\phi$  in  $\mathcal{M}_n \otimes \mathcal{R}_{(m)}$ .

Given an  $m$ -tuple  $A = (A_1, \dots, A_m)$  of  $n$  by  $n$  matrices  $A_j$  the Clifford operator of  $A$  is an element of  $\mathcal{M}_n \otimes \mathcal{R}_{(m)}$  defined as

$$(2.8) \quad \text{Cliff}(A) = i \sum A_j \otimes e_{\{j\}}.$$

We list below a few results which we need later and which are proved in Pryde [13].

**Lemma 2.1** (Pryde). (i) If  $A_j$  and  $A_k$  commute for every  $j$  and  $k$  then  $\text{Cliff}(A)^2 = \sum A_j^2$ .

(ii) For any  $\lambda \in \mathcal{R}^m$  define  $A - \lambda I = (A_1 - \lambda_1, \dots, A_m - \lambda_m)$ . If  $A_j$  and  $A_k$  commute for every  $j$  and  $k$  then  $\text{Cliff}(A - \lambda I)$  is invertible iff  $\sum (A_j - \lambda_j I)^2$  is invertible.

(iii) If  $A = (A_1, \dots, A_m)$  is an  $m$ -tuple of commuting selfadjoint matrices then  $\text{Cliff}(A)$  is selfadjoint and

$$(2.9) \quad \|\text{Cliff}(A)\| = \max\{\|\lambda\| : \lambda \in \sigma_{pt}(A)\}.$$

This last quantity  $\max\{\|\lambda\| : \lambda \in \sigma_{pt}(A)\}$  will occur in our discussion often. We will denote it by  $r(A)$  and call it the joint spectral radius of  $A$ .

We close this section with a small observation.

**Proposition 2.2.** If the  $D_j$ 's are diagonal matrices for  $j = 1, \dots, m$  then  $\text{Cliff}(D)$  is invertible iff  $\sum D_j^2$  is invertible, and in this case  $\text{Cliff}(D)^{-1} = (\sum D_j^2)^{-1} \text{Cliff}(D)$ .

*Proof.* The proof uses the previous lemma twice. Taking  $\lambda = (0, \dots, 0)$  in the second part of the lemma we see that  $\text{Cliff}(D)$  is invertible iff  $\sum D_j^2$  is invertible. Since the  $D_j$ 's are diagonal they commute. So by the first part of the lemma  $(\text{Cliff}(D))^2 = \sum D_j^2$ . Hence  $\text{Cliff}(D)^{-1} = (\sum D_j^2)^{-1} \text{Cliff}(D)$ .  $\square$

3. PERTURBATION OF REAL SPECTRA

In this section we consider  $m$ -tuples of matrices  $A = (A_1, \dots, A_m)$  with real eigenvalues only. Our first result in this section concerns the Bauer-Fike theorem [2]. Recall that for single matrices this says that if  $A$  is similar to a diagonal matrix, i.e., if there exists an invertible matrix  $T$  such that  $TAT^{-1} = \Lambda = \text{diag}(\alpha_1, \dots, \alpha_n)$  and  $B$  is any arbitrary matrix then  $\sigma(B)$  is contained in the union of the balls  $B(\alpha_i, \varepsilon)$ , where  $\varepsilon = \|A - B\| \cdot \|T\| \cdot \|T^{-1}\|$ . See [3, p. 114]. This was generalized to the case of  $m$ -tuples of commuting matrices by Pryde in [13]. However, as Stewart and Sun have pointed out in their recent book [15, p. 177] Bauer and Fike proved a stronger result:

Let  $A, B \in \mathcal{M}_n$  and  $T \in \text{GL}(n)$ . Then for any  $\beta \in \sigma(B) \setminus \sigma(A)$

$$(3.1) \quad \|T^{-1}(A - \beta I)^{-1}T\|^{-1} \leq \|T^{-1}(A - B)T\|.$$

A generalization of this to the case of  $m$ -tuples of commuting matrices with real spectra is the following theorem.

**Theorem 3.1.** *Let  $A = (A_1, \dots, A_m)$  and  $B = (B_1, \dots, B_m)$  be two  $m$ -tuples of commuting matrices with real spectra. Let  $\beta \in \sigma_{pt}(B) \setminus \sigma_{pt}(A)$ . Let  $T \in \text{GL}(n)$ . Then*

$$(3.2) \quad \|T^{-1}(\text{Cliff}(A - \beta I))^{-1}T\|^{-1} \leq \|T^{-1}(\text{Cliff}(A - B))T\|.$$

*Proof.* Since  $\beta \notin \sigma_{pt}(A)$  and all  $A_j$  have real spectra by (2.1),  $\sum(A_j - \beta_j I)^2$  is invertible. So by Lemma 2.1(ii)  $\text{Cliff}(A - \beta I)$  is invertible.

On the other hand, since  $\beta \in \sigma_{pt}(B)$ , there exists  $x \in \mathcal{E}^n$  such that  $B_j x = \beta_j x$  for all  $j = 1, \dots, n$ .

Hence,

$$\begin{aligned} \text{Cliff}(A - B)(x \otimes e_\phi) &= i \sum (A_j - B_j)(x \otimes e_{\{j\}}) \\ &= i \sum (A_j - \beta_j)(x \otimes e_{\{j\}}) = \text{Cliff}(A - \beta I)(x \otimes e_\phi). \end{aligned}$$

So,

$$\begin{aligned} x \otimes e_\phi &= (\text{Cliff}(A - \beta I))^{-1}(\text{Cliff}(A - B))(x \otimes e_\phi) \\ &= TT^{-1}(\text{Cliff}(A - \beta I))^{-1}TT^{-1}(\text{Cliff}(A - B))TT^{-1}(x \otimes e_\phi). \end{aligned}$$

Hence,

$$T^{-1}(x \otimes e_\phi) = (T^{-1}(\text{Cliff}(A - \beta I))^{-1}T)(T^{-1}(\text{Cliff}(A - B))T)(T^{-1}(x \otimes e_\phi)).$$

After taking norms and cancelling  $\|T^{-1}(x \otimes e_\phi)\|$  from both sides we have

$$\|(T^{-1}(\text{Cliff}(A - \beta I))^{-1}T)\|^{-1} \leq \|(T^{-1}(\text{Cliff}(A - B))T)\|. \quad \square$$

Now we will define the *measure of nonnormality* of an  $m$ -tuple  $A = (A_1, \dots, A_m)$  of commuting matrices. In this case there exists a unitary matrix  $U$  such that  $U^*A_jU = T_j$  for all  $j$  where the  $T_j$  are upper-triangular. (See [9, p. 81].) Write  $T_j = \Lambda_j + N_j$  where the  $\Lambda_j$  are diagonal and the  $N_j$  are strictly

upper-triangular. Let  $N = (N_1, \dots, N_m)$ . We can define the *measure of non-normality* of  $A$  as

$$(3.3) \quad \Delta(A) = \inf \| \text{Cliff}(N) \|,$$

where the infimum is taken over all choices of unitary  $U$  for which each  $U^* A_j U$  is upper-triangular. We obtain below a Henrici theorem in the case of  $m$ -tuples:

**Theorem 3.2.** *Let  $A = (A_1, \dots, A_m)$  and  $B = (B_1, \dots, B_m)$  be two commuting  $m$ -tuples of matrices with real spectra. Let  $\beta \in \sigma_{pt}(B) \setminus \sigma_{pt}(A)$ . Let  $\delta = \min\{\|\alpha - \beta\| : \alpha \in \sigma_{pt}(A)\}$ . Then*

$$(3.4) \quad \frac{\delta}{1 + \Delta(A)/\delta + \dots + \Delta^{n-1}(A)/\delta^{n-1}} \leq \| \text{Cliff}(A - B) \|.$$

*Proof.* Let  $U$  be a unitary such that the infimum in definition (3.3) of  $\Delta(A)$  is attained. Then  $U^* A_j U = \Lambda_j + N_j$  where the  $\Lambda_j$  are diagonal and the  $N_j$  are strictly upper-triangular. Let  $\Lambda = (\Lambda_1, \dots, \Lambda_m)$  and  $N = (N_1, \dots, N_m)$ . Then

$$(3.5) \quad \begin{aligned} U^* \text{Cliff}(A - \beta I) U &= i U^* \left( \sum (A_j - \beta_j I) \otimes e_{\{j\}} \right) U \\ &= i \sum (\Lambda_j + N_j - \beta_j I) \otimes e_{\{j\}} = \text{Cliff}(\Lambda + N - \beta I). \end{aligned}$$

But

$$(3.6) \quad \begin{aligned} \text{Cliff}(\Lambda + N - \beta I) &= \text{Cliff}(\Lambda - \beta I) + \text{Cliff}(N) \\ &= \text{Cliff}(\Lambda - \beta I) (1 + (\text{Cliff}(\Lambda - \beta I))^{-1} \text{Cliff}(N)). \end{aligned}$$

Note that  $\beta \notin \sigma_{pt}(\Lambda)$  since  $\beta \notin \sigma_{pt}(A)$ . Let  $\Lambda_j - \beta_j I = D_j$  and let  $D = (D_1, \dots, D_m)$ . By (2.1)  $\sum D_j^2$  is invertible. So by Proposition 2.2  $\text{Cliff}(D)$  is invertible and  $(\text{Cliff}(D))^{-1} = \sum_k T_k \otimes e_{\{k\}}$  where  $T_k = (\sum_i D_i^2)^{-1} D_k$ . The  $T_k$  are all diagonal. So,

$$(3.7) \quad \begin{aligned} (\text{Cliff}(\Lambda - \beta I))^{-1} \text{Cliff}(N) &= \left( \sum_k T_k \otimes e_{\{k\}} \right) \left( \sum_j N_j \otimes e_{\{j\}} \right) \\ &= \sum_{k,j} T_k N_j \otimes e_{\{k\}} e_{\{j\}} = X, \quad \text{say.} \end{aligned}$$

Now  $T_k N_j$  being the product of a diagonal and a strictly upper-triangular matrix is again strictly upper-triangular. In different powers of  $X$  various products of  $T_k N_j$  appear. But any product of  $n$  strictly upper-triangular matrices is zero. So  $X^n$  is zero. Therefore,

$$(3.8) \quad \begin{aligned} (\text{Cliff}(\Lambda + N - \beta I))^{-1} \\ = (1 - X + \dots + (-1)^{n-1} X^{n-1}) (\text{Cliff}(\Lambda - \beta I))^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \| (\text{Cliff}(\Lambda + N - \beta I))^{-1} \| &\leq \| (\text{Cliff}(\Lambda - \beta I))^{-1} \| \\ &\cdot (1 + \| (\text{Cliff}(\Lambda - \beta I))^{-1} \| \cdot \| \text{Cliff}(N) \| \\ &+ \dots + \| (\text{Cliff}(\Lambda - \beta I))^{-1} \|^{n-1} \cdot \| \text{Cliff}(N) \|^{n-1}). \end{aligned}$$

Let  $\eta = \|(\text{Cliff}(\Lambda - \beta I))^{-1}\|^{-1}$ . Then

$$(3.9) \quad \|(\text{Cliff}(\Lambda + N - \beta I))^{-1}\| \leq \eta^{-1} \left( 1 + \frac{\Delta(A)}{\eta} + \cdots + \frac{\Delta^{n-1}(A)}{\eta^{n-1}} \right).$$

Taking  $T = U$  in Theorem 3.1 we have

$$(3.10) \quad \begin{aligned} \|(\text{Cliff}(\Lambda + N - \beta I))^{-1}\|^{-1} &= \|U^*(\text{Cliff}(A - \beta I))^{-1}U\|^{-1} \\ &\leq \|U^* \text{Cliff}(A - B)U\| \leq \|\text{Cliff}(A - B)\|. \end{aligned}$$

By (3.9) and (3.10)

$$(3.11) \quad \|\text{Cliff}(A - B)\| \geq \frac{\eta}{(1 + \Delta(A)/\eta + \cdots + \Delta^{n-1}(A)/\eta^{n-1})}.$$

So the proof will be complete if we show that  $\eta = \delta$ . Recall that the  $D_j$  are diagonal matrices with real entries. Let  $d_i^{(j)}$  be the  $(i, i)$ th entry of  $D_j$ ; i.e.,  $D_j = \text{diag}(d_1^{(j)}, \dots, d_n^{(j)})$ . Then  $\sigma_{pt}(D) = \{(d_i^{(1)}, \dots, d_i^{(m)}) : i = 1, \dots, n\}$ . Put  $(d_i^{(1)}, \dots, d_i^{(m)}) = d_i$ . Then

$$\begin{aligned} \eta &= \|(\text{Cliff}(\Lambda - \beta I))^{-1}\|^{-1} \\ &= \left\| \left( \sum D_j^2 \right)^{-1} \text{Cliff}(D) \right\|^{-1} = \left\| \text{Cliff} \left( \sum D_j^2 \right)^{-1} D \right\|^{-1} \\ &= \max \{ \|d_i\| / \|d_i\|^2, i = 1, \dots, n \}^{-1} = \max \{ \|d_i\|, i = 1, \dots, n \} \\ &= \max \{ \|z\|, z \in \sigma_{pt}(\Lambda - \beta I) \} = \max \{ \|\alpha - \beta\| : \alpha \in \sigma_{pt}(A) \} \\ &= \delta \text{ by definition.} \end{aligned}$$

This completes the proof.  $\square$

#### 4. PERTURBATION OF COMPLEX SPECTRA

In this section we consider  $m$ -tuples  $A = (A_1, \dots, A_m)$  of commuting matrices with no restriction on  $A_j$ . In [10] McIntosh and Pryde have shown that any  $A_j$  can be written as  $A_{1j} + iA_{2j}$  where  $A_{1j}$  and  $A_{2j}$  are polynomials in  $A_j$  and each of them has real spectrum. Define  $\pi(A) = (A_{11}, \dots, A_{1m}, A_{21}, \dots, A_{2m})$ . This is called a *partition* of  $A$  in [13, 14]. Since the  $A_j$  commute, the  $A_{qj}$  also commute for all  $q = 1, 2$  and for all  $j = 1, \dots, m$ . So there exists a unitary  $U$  such that  $U^*A_{qj}U = \Lambda_{qj} + N_{qj}$  where  $\Lambda_{qj}$  is diagonal and  $N_{qj}$  upper-triangular. Let

$$(4.1) \quad N = (N_{11}, \dots, N_{1m}, N_{21}, \dots, N_{2m}),$$

and in this case define the measure of nonnormality of  $A$  by

$$(4.2) \quad \Delta(A) = \Delta(\pi(A)) = \inf \|\text{Cliff}(N)\|,$$

where as before the infimum is taken over all  $N$  associated with  $A$  in the above construction. For  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathcal{E}^m$  let  $\alpha_{1j} = \text{Re}(\alpha_j)$ ,  $\alpha_{2j} = \text{Im}(\alpha_j)$ . McIntosh and Pryde have also shown that  $\alpha \in \sigma_{pt}(A)$  iff  $(\alpha_{11}, \dots, \alpha_{1m}, \alpha_{21}, \dots, \alpha_{2m}) \in \sigma_{pt}(\pi(A))$ . Now let  $A$  and  $B$  be any two  $m$ -tuples of commuting

matrices. Let  $\beta \in \sigma_{pt}(B)$ . Then  $(\beta_{11}, \dots, \beta_{1m}, \beta_{21}, \dots, \beta_{2m}) \in \sigma_{pt}(\pi(B))$ . By the main result of § 4 there exists  $(\alpha_{11}, \dots, \alpha_{1m}, \alpha_{21}, \dots, \alpha_{2m}) \in \sigma_{pt}(\pi(A))$  such that if

$$\delta = \|(\beta_{11}, \dots, \beta_{1m}, \beta_{21}, \dots, \beta_{2m}) - (\alpha_{11}, \dots, \alpha_{1m}, \alpha_{21}, \dots, \alpha_{2m})\|$$

then

$$(4.3) \quad \frac{\delta}{1 + \Delta(A)/\delta + \dots + \Delta^{n-1}(A)/\delta^{n-1}} \leq \| \text{Cliff}(\pi(A) - \pi(B)) \|.$$

Note that  $\delta$  is also equal to  $\|\beta - \alpha\|$ , the distance between  $\beta$  and  $\alpha$  in  $\mathcal{E}^n$ . So we have proved

**Theorem 4.1.** *Let  $A = (A_1, \dots, A_m)$  and  $B = (B_1, \dots, B_m)$  be two  $m$ -tuples of commuting  $n$  by  $n$  matrices. Let  $\beta \in \sigma_{pt}(B) \setminus \sigma_{pt}(A)$ . Define  $\delta = \min\{\|\alpha - \beta\| : \alpha \in \sigma_{pt}(A)\}$ . Then*

$$\frac{\delta}{1 + \Delta(A)/\delta + \dots + \Delta^{n-1}(A)/\delta^{n-1}} \leq \| \text{Cliff}(\pi(A) - \pi(B)) \|.$$

## 5. A SPECTRAL VARIATION BOUND

Consider two  $n$  by  $n$  matrices  $A$  and  $B$ . Let  $\lambda_1, \dots, \lambda_n$  be the eigenvalues of  $A$  and  $\mu_1, \dots, \mu_n$  the eigenvalues of  $B$ . Let  $S_A(B) = \max_j \min_i |\lambda_i - \mu_j|$ . In [5] Bhatia and Friedland proved that

$$(5.1) \quad S_A(B) \leq n^{1/n} (2M)^{1-1/n} \|A - B\|^{1/n},$$

where  $M = \max(\|A\|, \|B\|)$ .

The approach in [5] was through characteristic polynomials. In [7] Elsner obtained the same result from Henrici's theorem. Using Elsner's approach we can obtain an analogue of (5.1) for commuting  $m$ -tuples. Let

$$(5.2) \quad S_A(B) = \max_{\mu \in \sigma_{pt}(B)} \min_{\lambda \in \sigma_{pt}(A)} \|\lambda - \mu\|.$$

Define

$$(5.3) \quad S_n(\Delta, r) = \begin{cases} yr/g_n(y) & \text{where } y = \Delta/r \text{ for } r > 0, \\ 0 & \text{for } r = 0. \end{cases}$$

Then  $S_n(\Delta, r)$  is strictly monotone in both its arguments.

The following lemma can be found in Elsner [7] and is crucial to the proof.

**Lemma 5.1.** *Given  $\tau \geq 0$ ,  $\delta > 0$ , and a positive integer  $n$  define*

$$\gamma = (\delta^{n-1} + \delta^{n-2}\tau + \dots + \tau^{n-1})^{1/n}.$$

*Then  $\gamma$  is the minimal number such that*

$$\min\{S_n(\tau M, r), \delta M\} \leq \gamma M^{1-1/n} r^{1/n} \quad \text{for all } M \geq 0, r \geq 0.$$

Theorem 4.1 can be equivalently stated as

$$(5.4) \quad S_A(B) \leq S_n(\Delta(A), \| \text{Cliff}(\pi(A) - \pi(B)) \|).$$

Let  $M = \max(\| \text{Cliff}(\pi(A)) \|, \| \text{Cliff}(\pi(B)) \|)$ . Let  $U$  be a unitary such that  $U^* A_{qj} U = N_{qj} + \Lambda_{qj}$ , and suppose the infimum in definition (4.2) is attained

for this choice, i.e.,  $\Delta(\pi(A)) = \|\text{Cliff}(N)\|$  where  $N$  is as in (4.1). Then by definition

$$(5.5) \quad \Delta(A) = \|\text{Cliff}(N)\| \leq \|\text{Cliff}(N + \Lambda)\| + \|\text{Cliff}(\Lambda)\|.$$

The first term on the right-hand side is  $\|\text{Cliff}(\pi(A))\|$ , and  $\|\text{Cliff}(\Lambda)\| = r(\Lambda) = r(\pi(A)) \leq \|\text{Cliff}(\pi(A))\|$ . So  $\Delta(A) \leq 2M$ . Hence by monotonicity of  $S_n$  in the first component

$$(5.6) \quad S_A(B) \leq S_n(2M, \|\text{Cliff}(\pi(A) - \pi(B))\|).$$

Also for all  $\lambda \in \sigma_{pt}(A)$  and  $\mu \in \sigma_{pt}(B)$

$$(5.7) \quad \|\lambda - \mu\| \leq \|\lambda\| + \|\mu\| \leq \|\text{Cliff}(A)\| + \|\text{Cliff}(B)\|,$$

so that  $S_A(B) \leq 2M$ . So,

$$(5.8) \quad S_A(B) \leq \min \{S_n(2M, \|\text{Cliff}(\pi(A) - \pi(B))\|), 2M\}$$

$$(5.9) \quad \leq n^{1/n}(2M)^{1-1/n} \|\text{Cliff}(\pi(A) - \pi(B))\|^{1/n}$$

by the lemma. This is nothing but the Bhatia-Friedland inequality in the present context. We state this as a theorem:

**Theorem 5.2.** *Let  $A = (A_1, \dots, A_m)$  and  $B = (B_1, \dots, B_m)$  be two  $m$ -tuples of commuting  $n$  by  $n$  matrices. Let  $\pi(A)$  and  $\pi(B)$  be partitions of  $A$  and  $B$  respectively. Let  $M = \max(\|\text{Cliff}(\pi(A))\|, \|\text{Cliff}(\pi(B))\|)$ . Let  $S_A(B)$  be as defined in (5.2). Then we have the following bound on  $S_A(B)$ :*

$$(5.10) \quad S_A(B) \leq n^{1/n}(2M)^{1-1/n} \|\text{Cliff}(\pi(A) - \pi(B))\|^{1/n}.$$

Finally we should point out that for the case of pairs of matrices  $A = (A_1, A_2)$  the norms  $\|\text{Cliff}(A)\|$  and  $\|\text{Cliff}(\pi(A))\|$  have been evaluated by Pryde [13]. For general  $m$ -tuples only bounds for them are known so far [10]. These can be combined with our results to give computable upper bounds for estimation of joint spectra. It remains an interesting open problem to find exact expressions for these norms.

#### REFERENCES

1. E. Artin, *Geometric algebra*, Wiley Interscience, New York, 1957.
2. F. L. Bauer and C. T. Fike, *Norms and exclusion theorems*, *Numer. Math.* **2** (1960), 137–141.
3. R. Bhatia, *Perturbation bounds for matrix eigenvalues*, Longman Scientific and Technical, Essex, England, 1987.
4. R. Bhatia, Ch. Davis, and A. McIntosh, *Perturbations of spectral subspaces and solutions of linear operator equations*, *Linear Algebra Appl.* **52** (1983), 45–67.
5. R. Bhatia and S. Friedland, *Variation of Grassman powers and spectra*, *Linear Algebra Appl.* **40** (1981), 1–18.
6. Ch. Davis, *Perturbation of spectrum of normal operators and of commuting tuples*, *Linear and Complex Analysis Problem Book* (V. P. Havin, S. V. Hruschev, and N. K. Nikol'ski, eds.), *Lecture Notes in Math.*, vol. 1043, Springer, Berlin and New York, 1984.
7. L. Elsner, *On the variation of spectra of matrices*, *Linear Algebra Appl.* **47** (1982), 127–138.
8. P. Henrici, *Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*, *Numer. Math.* **4** (1962), 24–39.
9. R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge Univ. Press, Cambridge, 1985.

10. A. McIntosh and A. Pryde, *A functional calculus for several commuting operators*, Indiana Univ. Math. J. **36** (1987), 421–439.
11. A. McIntosh, A. Pryde, and W. Ricker, *Comparison of joint spectra for certain classes of commuting operators*, Studia Math. **88** (1988), 23–36.
12. F. Ming, *Garske's inequality for an  $n$ -tuple of operators*, Integral Equations Operator Theory **14** (1991), 787–793.
13. A. Pryde, *A Bauer-Fike theorem for the joint spectrum of commuting matrices*, Linear Algebra Appl. **173** (1992), 221–230.
14. ———, *Optimal matching of joint eigenvalues for normal matrices*, Monash Univ. Analysis Paper 74, March 1991.
15. G. W. Stewart and J. Sun, *Matrix perturbation theory*, Academic Press, New York, 1990.

INDIAN STATISTICAL INSTITUTE, NEW DELHI 110016, INDIA

*E-mail address*, R. Bhatia: rbh@isid.ernet.in

*E-mail address*, T. Bhattacharyya: tirtha@isid.ernet.in