Some Problems in Joint Spectral Theory

Tirthankar Bhattacharyya
Indian Statistical Institute
New Delhi 110016, India

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements for a Ph.D. degree

THESIS SUPERVISOR: PROF. RAJENDRA BHATIA

April 1996

for

AMAL KIRAN DAS

of course

ACKNOWLEDGEMENTS

It would be impossible to thank individually all of those whose advice and support have gone behind this thesis. My warmest thanks go to my adviser Prof. R. Bhatia for many helpful discussions and encouragements and for suggesting the topic in the first place. It was a pleasure to work with him. I am indebted to Prof. K. B. Sinha for many inspiring lectures on operator theory and to Prof. H. Radjavi for a brief but stimulating interaction. I am also thankful to Prof. K. R. Parthasarathy and Prof. L. Elsner for helpful comments and suggestions.

All the staff members of the Delhi centre of the Indian Statistical Institute were extremely helpful. The facilities provided by them are gratefully acknowledged. My friends at the hostel made the stay an enjoyable experience.

The financial support for the work came from the University Grants Commission, India.

Finally, I wish I could say something for my parents.

Tirthankar Bhattacharyya

New Delhi
March, 1995

Contents

Introduction

In the last few years, joint spectral theory of several commuting operators has been studied by several authors. It is very different from the single variable operator theory. Attempts to extend results for a single operator to tuples of operators quickly result in easy counterexamples. Some one-dimensional facts do not have their analogues in the new situation. Thus it would be interesting to investigate how much of the one-variable theory has multi-variable analogues and where exactly they fail to carry over. The present thesis is of interest in this context. Here we extend some well-known resuls of the single operator theory to commuting tuples of operators.

The first chapter is an introduction to the basic facts of joint spectral theory. Here we briefly describe the joint spectrum and the functional calculus introduced by J. L. Taylor using the notion of Koszul complex in homological algebra. We also introduce the left and right spectra, the Harte spectrum and the approximate point spectrum. A commuting n-tuple of $d \times d$ matrices is simultaneously upper-triangularizable with the eigenvalues occuring along the diagonals. Then the d number of scalar n-tuples formed by picking up the (i,i)th. entries of the matrices for $i = 1, \ldots, d$ are called joint eigenvalues because to each of them there corresponds a vector $x_i \in \mathbb{C}^d$ which is a common eigenvector for all the n matrices. So, there is a natural definition of joint spectrum for tuples of matrices. The Taylor joint spectrum in this case reduces to these joint eigenvalues. In Section 1.3 we prove this fact following a partitioning argument of McIntosh, Pryde and Ricker. This chapter ends with a short introduction to Clifford algebras which are used in later chapters.

It is a well-known fact in operator theory that a compact operator on a Banach space can be upper-triangularised with respect to a maximal chain of invariant subspaces. In Chapter 2, we show that a commuting tuple of compact operators can be simultaneously upper-triangularised and the joint diagonal coefficients are in one-one correspondence with the non-zero joint eigenvalues of the tuple. This generalises the result for commuting matrices mentioned above. Moreover, exploiting a rule of multiplication of tuples we show that every non-zero point in the Taylor spectrum of a compact tuple is a joint eigenvalue with finite algebraic multiplicity.

Chapters 3 and 4 are devoted to finite-dimensional results.

Chō and Huruya proved a joint spectral radius formula for commuting $d \times d$ matrices considering them as operators on $\mathbb{C}^d$ with the Euclidean norm. In a recent paper, Müller and Soltysiak extended this result to operators on Hilbert spaces. In Chapter 3 we consider matrices as operators on the finite-dimensional Banach space $\mathbb{C}^d$ equipped with the p-norm. We then define a joint spectral radius with respct to the p-norm and prove a new spectral radius formula. We introduce a new operator corresponding to any n-tuple of commuting Banach space operators. This operator proves to be helpful in deriving the spectral radius formulae.

Chapter 4 is about perturbation of joint eigenvalues. The first results in this direction came from Pryde who obtained for commuting tuples of matrices an analogue of the Bauer-Fike theorem using ideas from Clifford analysis. In Chapter 4, using the Clifford operator, we obtain, in a similar spirit, extensions to commuting tuples of two well-known perturbation inequalities - the Henrici theorem and the Hoffman-Wielandt theorem.

Chapter 1

Preliminaries

Let $\mathcal{B}$ be a commutative Banach algebra with identity. For $a \in \mathcal{B}$, the spectrum of a in $\mathcal{B}$ is the set $\sigma_{\mathcal{B}}(a) = \{\lambda \in \mathbf{C} : (a-\lambda)\mathcal{B} \neq \mathcal{B}\}$, where $(a-\lambda)\mathcal{B}$ is the set of elements $\{(a-\lambda)b : b \in \mathcal{B}\}$. If $\mathcal{M}_{\mathcal{B}}$ is the maximal ideal space of $\mathcal{B}$ then the Gelfand transform $\hat{}: \mathcal{B} \to C(\mathcal{M}_{\mathcal{B}})$ is defined by $\hat{a}(\varphi) = \varphi(a), a \in \mathcal{B}, \varphi \in \mathcal{M}_{\mathcal{B}}$. It is a well-known fact (see [Co, page 224])that the spectrum of an element a can be described as $\sigma_{\mathcal{B}}(a) = \hat{a}(\mathcal{M}_{\mathcal{B}})$. In other words a is invertible if and only if $\hat{a}$ is never zero.

When $\mathcal{B}$ is not commutative, say $\mathcal{B} = \mathcal{L}(\mathcal{X})$, the space of linear operators on a Banach space $\mathcal{X}$, then $\sigma(a)$ is defined as $\{\lambda \in \mathbf{C} : a - \lambda \text{ is not invertible in } \mathcal{B}\}$. If $\mathcal{A}$ is any commutative subalgebra of $\mathcal{B}$ containing a, then the spectrum $\sigma_{\mathcal{A}}(a)$ depends on the algebra $\mathcal{A}$ in general. However when $\mathcal{A}$ is a maximal abelian subalgebra then it is easy to see that $\sigma_{\mathcal{A}}(a)$ is independent of $\mathcal{A}$ and equals $\sigma(a)$.

1.1 Algebraic Joint Spectra

For several Banach algebra elements, the (joint) spectral theory is more complicated. There is a difference between the study of commutative and non-commutative tuples and in this thesis we shall consider only commutative tuples.

1.1.1. **Definition.** Let $\mathcal{B}$ be a commutative Banach algebra and let $\boldsymbol{a}$ be the n-tuple $(a_1, \ldots, a_n) \in \mathcal{B}^n$. We say that $\boldsymbol{a}$ is *invertible* with respect to $\mathcal{B}$ if there exist $b_1, \ldots, b_n \in \mathcal{B}$ such that $\boldsymbol{a} \circ \boldsymbol{b} \stackrel{\text{def}}{=} \sum_{i=1}^n a_i b_i = 1$. For $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \mathbf{C}^n$, let $\boldsymbol{a} - \boldsymbol{\lambda}$ denote the n-tuple $(a_1 - \lambda_1, \ldots, a_n - \lambda_n)$. The *(algebraic) joint spectrum* of $\boldsymbol{a}$ in $\mathcal{B}$ is defined as

$$\sigma_{\mathcal{B}}(\boldsymbol{a}) = \{\lambda \in \mathbf{C}^n : \boldsymbol{a} - \boldsymbol{\lambda} \text{ is not invertible in } \mathcal{B}\}.$$

Let $\boldsymbol{a} \circ \mathcal{B}^n$ denote the following subalgebra of $\mathcal{B}$:

$$\boldsymbol{a} \circ \mathcal{B}^n \stackrel{\text{def}}{=} \{\boldsymbol{a} \circ \boldsymbol{b} : \boldsymbol{b} = (b_1, \ldots, b_n) \in \mathcal{B}^n\} \tag{1.1.2}$$

1.1.3. **Proposition.**

$$\sigma_{\mathcal{B}}(\boldsymbol{a}) = \{\boldsymbol{\lambda} \in \mathbf{C}^n : (\boldsymbol{a} - \boldsymbol{\lambda}) \circ \mathcal{B}^n \neq \mathcal{B}\} = \hat{\boldsymbol{a}}(\mathcal{M}_{\mathcal{B}}) \stackrel{\text{def}}{=} \{(\varphi(a_1), \ldots, \varphi(a_n)) : \varphi \in \mathcal{M}_{\mathcal{B}}\}.$$

PROOF Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n) \in \hat{\boldsymbol{a}}(\mathcal{M}_{\mathcal{B}})$ i.e., there exists a $\varphi \in \mathcal{M}_{\mathcal{B}}$ such that $\varphi(a_i) = \lambda_i$. Then for any n-tuple $\boldsymbol{b} = (b_1, \ldots, b_n)$ of elements from $\mathcal{B}$, we have $\varphi(\sum_{i=1}^n (a_i - \lambda_i) b_i) = \sum_{i=1}^n (\varphi(a_i) - \lambda_i) b_i = 0$. So $\sum_{i=1}^n (a_i - \lambda_i) b_i = 1$ has no solution in $\mathcal{B}$.

Conversely if $\sum_{i=1}^n (a_i - \lambda_i) b_i = 1$ has no solution in $\mathcal{B}$ then the ideal generated by $a_1 - \lambda_1, \ldots, a_n - \lambda_n$ is a proper one. So it is contained in some maximal ideal, say $\mathcal{I}$. The maximal ideals are in one-one correspondence with the multiplicative linear functionals. Let φ be the multiplicative linear functional with $\mathcal{I}$ as its kernel. Then $\boldsymbol{\lambda} = \varphi(\boldsymbol{a})$. ∎

If $p : \mathbf{C}^n \to \mathbf{C}^k$ is a polynomial mapping then it is immediate from Proposition 1.1.3 that $\sigma_{\mathcal{B}}(p(\boldsymbol{a})) = p(\sigma_{\mathcal{B}}(\boldsymbol{a}))$. In particular, $\sigma_{\mathcal{B}}(a_i) = \{\lambda_i : \boldsymbol{\lambda} \in \sigma_{\mathcal{B}}(\boldsymbol{a})\}$ for $i = 1, \ldots, n$. The following theorem due to Shilov, Arens-Calderón and Waelbroeck says that the algebraic joint spectrum carries an analytic functional calculus. We state it without proof. A detailed proof can be found in [Cu3].

1.1.4. **Theorem.** Let $\mathcal{U}(\sigma_{\mathcal{B}}(a))$ denote the algebra of germs of functions which are analytic in a neighbourhood of $\sigma_{\mathcal{B}}(a)$. There exists a continuous homomorphism $f \mapsto f(a)$ from $\mathcal{U}(\sigma_{\mathcal{B}}(a))$ into $\mathcal{B}$ such that

(i) $1(a) = 1$,

(ii) $z_i(a) = a_i$ for $i = 1, \ldots, n$ and

(iii) $\widehat{f(a)} = f \circ \hat{a}$ for all $f \in A(\sigma_{\mathcal{B}}(a))$.

As a consequence of Theorem 1.1.4 we have $\sigma_{\mathcal{B}}(f(a)) = f(\sigma_{\mathcal{B}}(a))$ for any function $f \in \mathcal{U}(\sigma_{\mathcal{B}}(a))$.

1.2 Spatial Joint Spectra

In this section we present the axiomatic approach to the joint spectrum of a tuple of elements from a *non commutative* Banach algebra $\mathcal{B}$. We discuss various spectral systems with a special emphasis on the Taylor joint spectrum.

1.2.1. **Definition.** For an integer $n \geq 1$ let $\mathcal{B}^n_{\text{com}}$ be the set of all commuting n-tuples of elements from $\mathcal{B}$.

$$\mathcal{B}^n_{\text{com}} \stackrel{\text{def}}{=} \{a = (a_1, \ldots, a_n) \in \mathcal{B}^n : a_i a_j = a_j a_i \text{ for all } i, j\}.$$

$\cup_{n=1}^{\infty} \mathcal{B}^n_{\text{com}}$ will be denoted by $\mathcal{B}_{\text{com}}$.

To define the joint spectrum of $a \in \mathcal{B}_{\text{com}}$ for a general Banach algebra $\mathcal{B}$ one might look for a maximal abelian subalgebra. But unlike the single element case the joint spectrum *does depend* on the maximal abelian subalgebra. For an example see [Al].

If X is any set let $\mathcal{P}(X)$ denote its power set i.e., the set of all subsets of X. Let $\mathbb{C}^\infty$ denote the space of all complex sequences.

1.2.2. **Definition.** A *spectral system* for a Banach algebra $\mathcal{B}$ is a map

$$\tilde{\sigma} : \mathcal{B}_{\text{com}} \longrightarrow \mathcal{P}(\mathbb{C}^\infty)$$

such that

(i) $\boldsymbol{a} \in \mathcal{B}_{\text{com}} \Rightarrow \tilde{\sigma}(\boldsymbol{a}) \neq \emptyset$,

(ii) $\boldsymbol{a} \in \mathcal{B}^n_{\text{com}} \Rightarrow \tilde{\sigma}(\boldsymbol{a}) \subseteq \mathbb{C}^n \subseteq \mathbb{C}^\infty$, and

(iii) $\boldsymbol{a} \in \mathcal{B}_{\text{com}} \Rightarrow \tilde{\sigma}(\boldsymbol{a})$ is compact.

1.2.3. **Definition.** Let $\tilde{\sigma}(\boldsymbol{a})$ be a spectral system on a Banach algebra $\mathcal{B}$. Let $\boldsymbol{a} \in \mathcal{B}^n_{\text{com}}$ and $\boldsymbol{b} \in \mathcal{B}^k_{\text{com}}$. Let $P_1 : \mathbb{C}^{n+k} \to \mathbb{C}^n$ and $P_2 : \mathbb{C}^{n+k} \to \mathbb{C}^k$ be the projections $P_1(z_1, \ldots, z_{n+k}) = (z_1, \ldots, z_n)$ and $P_2(z_1, \ldots, z_{n+k}) = (z_{n+1}, \ldots, z_{n+k})$ respectively. The spectral system $\tilde{\sigma}$ for $\mathcal{B}$ is said to possess the *projection property* if

$$P_1\tilde{\sigma}(\boldsymbol{a}, \boldsymbol{b}) = \tilde{\sigma}(\boldsymbol{a}) \text{ and } P_2\tilde{\sigma}(\boldsymbol{a}, \boldsymbol{b}) = \tilde{\sigma}(\boldsymbol{b}).$$

$\tilde{\sigma}$ possesses the *spectral mapping property for polynomials* if $\tilde{\sigma}(p(\boldsymbol{a})) = p(\tilde{\sigma}(\boldsymbol{a}))$ for every polynomial $p : \mathbb{C}^n \to \mathbb{C}^k$ and for every n-tuple $\boldsymbol{a} \in \mathcal{B}^n_{\text{com}}$.

1.2.4. **Examples.**

(i) Let $\mathcal{B}$ be a commutative Banach algebra. For $\boldsymbol{a} \in \mathcal{B}^n$ the *rational spectrum* of $\boldsymbol{a}$ in $\mathcal{B}$ is $\sigma_R(\boldsymbol{a})$, where R is the smallest inverse-closed closed subalgebra of $\mathcal{B}$ that contains $a_1, \ldots, a_n$. In general $\sigma_{\mathcal{B}}(\boldsymbol{a}) \subset \sigma_R(\boldsymbol{a})$. Both $\sigma_{\mathcal{B}}$ and σ_R are spectral systems. Obviously $\sigma_{\mathcal{B}}$ has the projection property by virtue of Proposition 1.1.3. But σ_R does not have the projection property. For this result and more about the rational spectrum, see [Wa].

(ii) For a subset $\mathcal{S}$ of $\mathcal{B}$, let $\mathcal{S}' \overset{\text{def}}{=} \{b \in \mathcal{B} : bs = sb \text{ for every } s \in \mathcal{S}\}$. It is easy to see that $\mathcal{S}'$ is always an algebra. The algebra $\mathcal{S}'$ is called the *commutant* of $\mathcal{S}$. Similarly $\mathcal{S}'' \overset{\text{def}}{=} (\mathcal{S}')'$ is called the *double commutant* of $\mathcal{S}$. For $\boldsymbol{a} \in \mathcal{B}^n_{\text{com}}$ we let $(\boldsymbol{a}), (\boldsymbol{a})'$ and $(\boldsymbol{a})''$ denote the Banach subalgebras of $\mathcal{B}$ generated by $a_1, \ldots, a_n$ and 1, by its commutant (relative to $\mathcal{B}$), and by its double commutant respectively. We define

$$\hat{\sigma}(\boldsymbol{a}) \overset{\text{def}}{=} \sigma_{(\boldsymbol{a})}(\boldsymbol{a}), \tag{1.2.5}$$

$$\sigma''(\boldsymbol{a}) \overset{\text{def}}{=} \sigma_{(\boldsymbol{a})''}(\boldsymbol{a}), \tag{1.2.6}$$

and

(1.2.7) $$\sigma'(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : (a-\lambda)\, o\, (a)' \neq (a)'\}.$$

More generally, if $\mathcal{A}$ is any closed subalgebra of $\mathcal{B}$ containing a in its center, then we let

(1.2.8) $$\sigma_{\mathcal{A}}(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : (a-\lambda)\, o\, \mathcal{A} \neq \mathcal{A}\}.$$

In general, by an *algebraic spectrum* we mean $\sigma_{\mathcal{A}}$ for some $\mathcal{A}$. The sets $\hat{\sigma}, \sigma'$ and σ'' are spectral systems without the projection property (see [SlZ]). Also $\sigma' \subseteq \sigma'' \subseteq \hat{\sigma}$.

(iii) The *left spectrum* of a is defined by

(1.2.9) $$\sigma_l(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : \mathcal{B}\, o\, (a-\lambda) \neq \mathcal{B}\}.$$

Similarly, the *right spectrum* of a is defined by

(1.2.10) $$\sigma_r(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : (a-\lambda)\, o\, \mathcal{B} \neq \mathcal{B}\}.$$

The *Harte spectrum* is $\sigma_{\mathrm{H}} \stackrel{\text{def}}{=} \sigma_l \cup \sigma_r \subseteq \sigma'$. The spectral systems σ_l, σ_r and σ_{H} all possess the projection property (see [Bu], [Har1] and [Har2]).

(iv) Suppose $\mathcal{B} = \mathcal{L}(\mathcal{X})$, where $\mathcal{X}$ is a Banach space. The tuple $a \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ is said to be *jointly bounded below* if there exists a positive constant $\epsilon > 0$ such that

(1.2.11) $$\sum_{i=1}^{n} \|a_i x\| \geq \epsilon \|x\|,$$

The tuple a is *jointly onto* if

(1.2.12) $$\sum_{i=1}^{n} a_i \mathcal{X} = \mathcal{X}.$$

The range space of a is defined to be the subspace

(1.2.13) $$\mathcal{R}(a) = \{y \in \mathcal{X} : \text{There exist } x_1, \ldots, x_n \in \mathcal{X} \text{ such that } y = a_1 x_1 + \cdots + a_n x_n\}.$$

In other words, $\mathcal{R}(a) = \mathcal{R}(a_1) + \cdots + \mathcal{R}(a_n)$. In this notation, a is jointly onto if $\mathcal{R}(a) = \mathcal{X}$.

For $a \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ the *approximate point spectrum* $\sigma_\pi(a)$ and the *defect spectrum* $\sigma_\delta(a)$ are defined by

$$\sigma_\pi(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : a - \lambda \text{ is not jointly bounded below}\} \tag{1.2.14}$$

and

$$\sigma_\delta(a) \stackrel{\text{def}}{=} \{\lambda \in \mathbb{C}^n : a - \lambda \text{ is not jointly onto}\}. \tag{1.2.15}$$

Then σ_π and σ_δ are spectral systems with the projection property (see [ChDa]).

(v) For $a \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ let $\sigma_\Pi(a, \mathcal{X}) \stackrel{\text{def}}{=} \sigma(a_1, \mathcal{X}) \times \cdots \times \sigma(a_n, \mathcal{X})$, where $\sigma(a_i, \mathcal{X})$ denotes the ordinary spectrum of a_i as an element of $\mathcal{L}(\mathcal{X})$. Obviously, the set σ_Π is a spectral system with the projection property. However, σ_Π does not have the spectral mapping property for polynomial mappings. (Example: Let p be a nontrivial idempotent in $\mathcal{L}(\mathcal{X})$. Then $\sigma_\Pi(p, p) = \{(0,0), (0,1), (1,0), (1,1)\}$, so that $\{\lambda_1 + \lambda_2 : (\lambda_1, \lambda_2) \in \sigma_\Pi(p,p)\} = \{0, 1, 2\}$, while $\sigma(2p) = \{0, 2\}$.)

(vi) *The Taylor spectrum*: In this example we discuss the joint spectrum introduced by J. L. Taylor in [T1]. Let Λ_n be the exterior algebra on n generators with identity $e_0 = 1$. This is the algebra of forms in $e_1, \ldots, e_n$ with complex coefficients, subject to the collapsing property $e_i e_j + e_j e_i = 0 (1 \le i, j \le n)$. The algebra Λ_n is graded: $\Lambda_n = \oplus_{k=1}^n \Lambda_n^k$ with $\{e_{i_1} \cdot \ldots \cdot e_{i_k} : 1 \le i_1 < \ldots < i_k \le n\}$ as the basis for Λ_n^k. Let $E_i : \Lambda_n \to \Lambda_n$ be given by

$$E_i \xi = e_i \xi, \quad i = 1, \ldots, n, \quad \xi \in \Lambda_n. \tag{1.2.16}$$

The operators $E_1, \ldots, E_n$ are called the *creation operators*. Clearly $E_i E_j + E_j E_i = 0 (1 \le i, j \le n)$. We regard Λ_n as a Hilbert space by declaring $\{e_{i_1} \cdot \ldots \cdot e_{i_k} : 1 \le i_1 < \ldots < i_k \le n \text{ and } k = 1, 2, \ldots, n\}$ to be an orthonormal basis. Then each E_i is a partial isometry and $E_i^* E_j + E_j E_i^* = \delta_{ij} (1 \le i, j \le n)$. If $\mathcal{X}$ is a vector space, we define $\Lambda_n(\mathcal{X}) \stackrel{\text{def}}{=} \mathcal{X} \otimes \Lambda_n$. Then for $\mathbf{A} = (A_1, \ldots A_n) \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ the operator $D_{\mathbf{A}} : \Lambda_n(\mathcal{X}) \to \Lambda_n(\mathcal{X})$ is defined by

$$D_{\mathbf{A}} \stackrel{\text{def}}{=} \sum_{i=1}^n A_i \otimes E_i. \tag{1.2.17}$$

Then

$$D^2_{\mathbf{A}} = \sum_{i,j=1}^{n} A_j A_i \otimes E_j E_i = \sum_{i<j} A_i A_j \otimes (E_i E_j + E_j E_i) = 0, \tag{1.2.18}$$

If $\mathcal{R}(D_{\mathbf{A}})$ and $\mathcal{N}(D_{\mathbf{A}})$ denote the range and kernel respectively of the operator $D_{\mathbf{A}}$, then the equation (1.2.18) says that $\mathcal{R}(D_{\mathbf{A}}) \subseteq \mathcal{N}(D_{\mathbf{A}})$. We say that $\mathbf{A}$ is *nonsingular* on $\mathcal{X}$ if $\mathcal{R}(D_{\mathbf{A}}) = \mathcal{N}(D_{\mathbf{A}})$. The *Taylor spectrum* of $\mathbf{A}$ on $\mathcal{X}$ is

$$\sigma_{\mathrm{T}}(\mathbf{A}, \mathcal{X}) \stackrel{\mathrm{def}}{=} \{\boldsymbol{\lambda} \in \mathbf{C}^n : \mathcal{R}(D_{\mathbf{A}-\boldsymbol{\lambda}}) \neq \mathcal{N}(D_{\mathbf{A}-\boldsymbol{\lambda}})\}. \tag{1.2.19}$$

When $n = 1$, D_A has the following 2×2 matrix representation relative to the direct sum decomposition $(\mathcal{X} \otimes e_0) \oplus (\mathcal{X} \otimes e_1)$:

$$D_A = \begin{pmatrix} 0 & 0 \\ A & 0 \end{pmatrix} \tag{1.2.20}$$

So,

$$\mathcal{N}(D_{\mathbf{A}}) = \mathcal{N}(A) \oplus \mathcal{X}, \tag{1.2.21}$$

$$\mathcal{R}(D_{\mathbf{A}}) = 0 \oplus \mathcal{R}(A), \tag{1.2.22}$$

and

$$\mathcal{N}(D_{\mathbf{A}})/\mathcal{R}(D_{\mathbf{A}}) = \mathcal{N}(A) \oplus \mathcal{X}/\mathcal{R}(A), \tag{1.2.23}$$

so that A is nonsingular if and only if A is one-one and onto. For $n = 2$ the matrix of $D_{\mathbf{A}}$ relative to the direct sum decomposition
$\Lambda_2(\mathcal{X}) = (\mathcal{X} \otimes e_0) \oplus (\mathcal{X} \otimes e_1) \oplus (\mathcal{X} \otimes e_2) \oplus (\mathcal{X} \otimes e_1 e_2)$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ A_1 & 0 & 0 & 0 \\ A_2 & 0 & 0 & 0 \\ 0 & -A_2 & A_1 & 0 \end{pmatrix}. \tag{1.2.24}$$

So

$$\mathcal{N}(D_{\mathbf{A}}) = \{\mathcal{N}(A_1) \cap \mathcal{N}(A_2)\} \oplus \{(x_1, x_2) : A_2 x_1 = A_1 x_2\} \oplus \mathcal{X}, \tag{1.2.25}$$

$$\mathcal{R}(D_{\mathbf{A}}) = 0 \oplus \{(A_1 x_0, A_2 x_0) : x_0 \in \mathcal{X}\} \oplus \{\mathcal{R}(A_1) + \mathcal{R}(A_2)\}. \tag{1.2.26}$$

One can also consider a chain complex $K(\mathbf{A},\mathcal{X})$, called the *Koszul complex* as follows:

(1.2.27) $K(\mathbf{A},\mathcal{X}): 0 \to \Lambda_n^0(\mathcal{X}) \xrightarrow{D^0_{\mathbf{A}}} \Lambda_n^1(\mathcal{X}) \xrightarrow{D^1_{\mathbf{A}}} \cdots \xrightarrow{D^{k-1}_{\mathbf{A}}} \Lambda_n^k(\mathcal{X}) \xrightarrow{D^k_{\mathbf{A}}} \cdots \xrightarrow{D^{n-1}_{\mathbf{A}}} \Lambda_n^n(\mathcal{X}) \to 0.$

(here $\Lambda_n^k(\mathcal{X})$ is $X \otimes \Lambda_n^k$ and $D^k_{\mathbf{A}} \stackrel{\text{def}}{=} D_{\mathbf{A}}|_{\Lambda_n^k(\mathcal{X})}$). Then it is easy to see that

$$\sigma_{\mathrm{T}}(\mathbf{A},\mathcal{X}) = \{\lambda \in \mathbb{C}^n : K(\mathbf{A}-\lambda,\mathcal{X}) \text{ is not exact}\}. \tag{1.2.28}$$

For $n = 2$ this complex takes the form

$$0 \longrightarrow \mathcal{X} \xrightarrow{\delta_1} \mathcal{X} \oplus \mathcal{X} \xrightarrow{\delta_0} \mathcal{X} \longrightarrow 0, \tag{1.2.29}$$

where $\delta_1(x) = (-A_2x, A_1x)$ and $\delta_0(x_1,x_2) = A_1x_1 + A_2x_2$. Clearly $A_1A_2 = A_2A_1$ implies that $\delta_0 \, o \, \delta_1 = 0$ so that (1.2.28) is a chain complex. To say that (1.2.29) is exact means three things. We see from the definitions of the maps δ_1 and δ_0 that the exactness at the first stage and the third stage respectively mean that $\mathcal{N}(A_1) \cap \mathcal{N}(A_2) = 0$ and $\mathcal{R}(\mathbf{A}) = \mathcal{X}$. Exactness at the second stage means that every pair $(x_1,x_2) \in \mathcal{X} \oplus \mathcal{X}$ for which $A_1x_1 + A_2x_2 = 0$ has the form $(x_1,x_2) = (-A_2x, A_1x)$ for some $x \in \mathcal{X}$. This is the same as dictated by equations (1.2.25) and (1.2.26).

There are $n+1$ maps involved in the complex (1.2.27) and $K(\mathbf{A},\mathcal{X})$ is exact iff $\mathcal{R}(D^{k-1}_{\mathbf{A}}) = \mathcal{N}(D^k_{\mathbf{A}})$ for all $k = 1,\ldots,n-1$, $\mathcal{N}(D^0_{\mathbf{A}}) = 0$ and $\mathcal{R}(D^{n-1}_{\mathbf{A}}) = \Lambda_n^n(\mathcal{X})$. Let us analyse the last two conditions. Since $\Lambda_n^0(\mathcal{X})$ consists of vectors of the form $x \otimes e_0$, $\mathcal{N}(D^0_{\mathbf{A}}) = \{x \in \mathcal{X} : \sum_{i=1}^n A_ix \otimes e_i = 0\}$. Since the e_i's are the generators of the algebra Λ_n, $\sum_{i=1}^n A_ix \otimes e_i = 0$ iff $A_ix = 0$ for all $i = 1,\ldots,n$. So $\mathcal{N}(D^0_{\mathbf{A}}) = \mathcal{N}(A_1) \cap \mathcal{N}(A_2) \cap \cdots \cap \mathcal{N}(A_n)$. Thus exactness at the first stage means that

$$\mathcal{N}(A_1) \cap \mathcal{N}(A_2) \cap \cdots \cap \mathcal{N}(A_n) = \{0\}. \tag{1.2.30}$$

$\Lambda_n^{n-1}(\mathcal{X})$ consists of vectors of the form $\sum_{i=1}^n x_i \otimes e_1 \ldots \hat{e}_i \ldots e_n$, where $\hat{e}_i$ means that e_i is omitted. So $\mathcal{R}(D^{n-1}_{\mathbf{A}}) = \{(\sum_{i=1}^n A_ix_i) \otimes e_1 \ldots e_n : x_1,\ldots x_n \in \mathcal{X}\} = \mathcal{R}(\mathbf{A}) \otimes \Lambda_n$. Hence exactness at the last stage means that

$$\mathcal{R}(\mathbf{A}) = \mathcal{X}. \tag{1.2.31}$$

Obviously for the exactness of the complex $K(\mathbf{A}, \mathcal{X})$ these conditions are necessary but not sufficient. Note that

(1.3.32) $\sigma_\delta(\mathbf{A}, \mathcal{X}) = \{\boldsymbol{\lambda} \in \mathbb{C}^n : D^{n-1}_{\mathbf{A}-\boldsymbol{\lambda}} \text{ is not onto}\} = \{\boldsymbol{\lambda} \in \mathbb{C}^n : \mathcal{R}(\mathbf{A} - \boldsymbol{\lambda}) \neq \mathcal{X}\}$

so that σ_δ represents the singularity of the complex at the last stage. Now it is *natural to introduce the following definition:*

1.2.33. **Definition.** $\boldsymbol{\lambda} \in \mathbb{C}^n$ is called a *joint eigenvalue* for $\mathbf{A} \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ if there exists a nonzero vector $x \in \mathcal{X}$ such that $A_i x = \lambda_i x$ for all $i = 1, \ldots, n$. The *joint point spectrum* $\sigma_{\text{pt}}(\mathbf{A}, \mathcal{X})$ is the collection of all joint eigenvalues.

If $\mathbf{A}$ is not a tuple but a single bounded operator A then $\sigma_{\text{pt}}(\mathbf{A}, \mathcal{X}) = \sigma_{\text{pt}}(A)$, the orinary point spectrum of A. Obviously, $\sigma_{\text{pt}}(\mathbf{A}, \mathcal{X}) \subset \sigma_{\text{pt}}(A_1) \times \cdots \times \sigma_{\text{pt}}(A_n)$. So if $A_1, \ldots, A_n$ have empty point spectra, then $\sigma_{\text{pt}}(\mathbf{A}, \mathcal{X})$ is empty. Thus the set σ_{pt} is not in general a spectral system. In terms of the Koszul complex $\sigma_{\text{pt}}(\mathbf{A}, \mathcal{X}) = \{\boldsymbol{\lambda} \in \mathbb{C}^n : D^0_{\mathbf{A}-\boldsymbol{\lambda}} \text{ is not one-one}\}$. So σ_{pt} represents the singularity of the complex at the first stage. Hence $\sigma_{\text{pt}}(\mathbf{A}) \subseteq \sigma_{\text{T}}(\mathbf{A})$. In the next section we shall see that the Taylor spectrum of a tuple of matrices consists of the joint point spectrum only. The inclusion relations for different spectra are as follows.

(1.2.34) $$\sigma_T \subseteq \sigma' \subseteq \sigma'' \subseteq \sigma_H.$$

And $\sigma_H \subseteq \sigma_T$ if the underlying space is either a finite-dimensional space or a Hilbert space. In general, all the above inclusions can be proper. See [T1, T2].

The Taylor joint spectrum $\sigma_{\text{T}}(\mathbf{A})$ carries an analytic functional calculus. We state it without proof. A detailed proof can be found in [T2] and [Cu3].

1.2.35. **Theorem.** Let $\mathbf{A} \in \mathcal{L}(\mathcal{X})^n_{\text{com}}$ and let $\Omega \supseteq \sigma_{\text{T}}(\mathbf{A}, \mathcal{X})$. Let $\mathcal{U}(\Omega)$ be the algebra of germs of analytic functions on Ω. Then there is a unital continuous homomorphism

$$f : \mathcal{U}(\Omega) \to \mathcal{L}(\mathcal{X})$$

satisfying

(i) $z_i(\mathbf{A}) = A_i$ for $i = 1, \ldots, n$,

(ii) $f(\mathbf{A}) \in (\mathbf{A})''$,

(iii) If $f = 0$ on a neighbourhood of $\sigma_T(\mathbf{A}, \mathcal{X})$, then $f(\mathbf{A}) = 0$.

(iv) For every relatively compact open subset Φ of Ω containing $\sigma_T(\mathbf{A}, \mathcal{X})$, there exists a constant $C_\Phi > 0$ such that

$$\|f(\mathbf{A})\| \leq C_\Phi \sup\{|f(z)| : z \in \Phi\}$$

for all $f \in \mathcal{U}(\Omega)$.

1.2.36. **Theorem.** Let $\mathbf{A} \in \mathcal{L}(\mathcal{X})^n_{\text{com}}, \Omega \supseteq \sigma_T(\mathbf{A}, \mathcal{X})$, and let $f \in \mathcal{U}(\Omega)$. Then

$$\sigma_T(f(\mathbf{A}), \mathcal{X}) = f(\sigma_T(\mathbf{A}, \mathcal{X})).$$

Proof See [T2].

So σ_T is a spectral system possessing the projection property and the spectral mapping property for analytic functions.

1.3 The Joint Spectrum of Matrices

In this section we show that the Taylor (joint) spectrum of a tuple of matrices is the set of joint eigenvalues. The proofs presented in this section are on the lines of [Cu1]. Let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{C} = (C_1, \ldots, C_n)$ be two n-tuples of bounded operators on two Banach spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ respectively. For $B_i \in \mathcal{B}(\mathcal{X}_2, \mathcal{X}_1), i = 1, \ldots, n$ let

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{C} \end{pmatrix} = \left(\begin{pmatrix} A_1 & B_1 \\ 0 & C_1 \end{pmatrix}, \ldots, \begin{pmatrix} A_n & B_n \\ 0 & C_n \end{pmatrix} \right).$$

1.3.1. **Lemma.** Suppose $B_1, \ldots, B_n$ are such that $\left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{C} \end{smallmatrix}\right)$ is a commuting n-tuples of operators on $\mathcal{X}_1 \oplus \mathcal{X}_2$. If $\mathbf{A}$ and $\mathbf{C}$ are nonsingular, then $\left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{C} \end{smallmatrix}\right)$ is nonsingular. (Here nonsingularity is in the sense of Taylor, see (1.2.17), (1.2.18) and the discussion there.) Consequently, $\sigma_T\left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ 0 & \mathbf{C} \end{smallmatrix}\right) \subseteq \sigma_T(\mathbf{A}) \cup \sigma_T(\mathbf{C})$.

PROOF For notational convenience, let us denote the map $D_{\mathbf{A}}$, corresponding to any commuting tuple $\mathbf{A}$, as defined in (1.2.16), by $D(\mathbf{A})$. Recall that a tuple $\mathbf{A}$ is nonsingular if $\mathcal{R}(D(\mathbf{A})) = \mathcal{N}(D(\mathbf{A}))$. Hence we have to show that $\mathcal{R}(D(\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{smallmatrix})) = \mathcal{N}(D(\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{smallmatrix}))$. A trivial identification shows that $D(\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{smallmatrix}) = (\begin{smallmatrix}D(\mathbf{A}) & D(\mathbf{B})\\ 0 & D(\mathbf{C})\end{smallmatrix})$. So $D(\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{smallmatrix})(\begin{smallmatrix}x\otimes\xi\\ y\otimes\eta\end{smallmatrix}) = 0$ implies that $D(\mathbf{A})(x\otimes\xi)+D(\mathbf{B})(y\otimes\eta) = 0$ and $D(\mathbf{C})(y\otimes\eta) = 0$. Since $\mathbf{C}$ is invertible, there exists $z \otimes \zeta \in \Lambda_n(\mathcal{X}_2)$ such that $y \otimes \eta = D(\mathbf{C})(z \otimes \zeta)$. Thus $0 = D(\mathbf{A})(x \otimes \xi) + D(\mathbf{B})D(\mathbf{C})(z \otimes \zeta)$. But $D(\mathbf{B})D(\mathbf{C}) = -D(\mathbf{A})D(\mathbf{B})$ since $(D(\begin{smallmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{smallmatrix}))^2 = 0$. Hence $0 = D(\mathbf{A})([x \otimes \xi - D(\mathbf{B})(z \otimes \zeta)])$. Now by nonsingularity of $\mathbf{A}$, we have $x \otimes \xi - D(\mathbf{B})(z \otimes \zeta) = D(\mathbf{A})(u \otimes \gamma)$ for some $u \otimes \gamma \in \Lambda_n(\mathcal{H}_1)$. Hence

$$D\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ 0 & \mathbf{C}\end{pmatrix}\begin{pmatrix}u\otimes\gamma\\ z\otimes\zeta\end{pmatrix} = \begin{pmatrix}x\otimes\xi\\ y\otimes\eta\end{pmatrix}.$$

This completes the proof.

The next lemma is a very well-known result on simultaneous upper-triangularization of a commuting tuple of matrices.

1.3.2. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ be a commuting n-tuple of linear transformations on an N-dimensional vector space $\mathcal{X}$. Then there exist $N+1$ subspaces $L_0, L_1, \ldots, L_N$ satisfying:

(i) $\{0\} = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_N = \mathcal{X}$,

(ii) L_k is k-dimensional $(k = 1, \ldots, N)$,

(iii) each L_k is simultaneously invariant under $A_1, \ldots, A_n$.

PROOF It is elementary to see that for any commuting family $\mathcal{F}$ of operators on a finite-dimensional vector space, there exists a vector $x \in \mathcal{X}$ that is an eigenvector of every $T \in \mathcal{F}$ (for a proof see [HorJ], page 51). Applying this to the commuting tuple $\mathbf{A} = (A_1, \ldots, A_n)$, we get a common eigenvector x. Let L_1 be the one-dimenional space spanned by x. Then L_1 is invariant under $A_1, \ldots, A_n$. Next consider the vector space $\mathcal{W} = \mathcal{X}/L_1$ and the linear transformations $\tilde{A}_j$ in $\mathcal{W}$ defined as $\tilde{A}_j(x+L_1) = A_jx+L_1$. Then $\tilde{A}_1, \ldots, \tilde{A}_n$ is are commuting linear transformations on the finite-dimensional space $\mathcal{W}$. So they have a common eigenvector, say x_2+L_1. Thus for some scalar tuple $(\mu_1, \ldots, \mu_n)$, $\tilde{A}_j(x_2 + L_1) = \mu_j x_2 + L_1$ which means that

$A_j x_2 = \mu_j x_2 + z$ for some $z \in L_1$. Hence the subspace spanned by x_1 and x_2 is invariant under $A_1, \ldots, A_n$. We call this subspace L_2, it is two-dimensional, contains L_1 and is invariant.

Now applying the same reasoning to $\mathcal{X}/L_2$ and so on, we get for each $j = 1, \ldots, N-1$ the subspace L_j spanned by $x_1, x_2, \ldots, x_j$. These subspaces satisfy the conditions of the theorem. Finally, define $L_N = \mathcal{X}$ and that completes the proof. ∎

Let $x_1, \ldots, x_{N-1}$ and $L_1, \ldots, L_N$ be as in the proof of Theorem 1.3.2. Choose any x_N not in L_{N-1} arbitrarily. Then $\{x_1, \ldots, x_N\}$ is a basis for $\mathcal{X}$. In this basis the matrices for $A_1, \ldots, A_n$ are upper-triangular i.e., of the form

$$\begin{pmatrix} \lambda_1^{(1)} & a_{12}^{(1)} & \cdots & a_{1N}^{(1)} \\ 0 & \lambda_2^{(1)} & \cdots & a_{2N}^{(1)} \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_N^{(1)} \end{pmatrix}, \ldots\ldots, \begin{pmatrix} \lambda_1^{(n)} & a_{12}^{(n)} & \cdots & a_{1N}^{(n)} \\ 0 & \lambda_2^{(n)} & \cdots & a_{2N}^{(n)} \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_N^{(n)} \end{pmatrix}.$$

1.3.3. **Definition.** For $i = 1, \ldots, N$, let $\boldsymbol{\lambda}_i$ be the n-tuple consisting of the ith. diagonal entries in the matrices of $A_1, \ldots, A_n$ i.e. $\boldsymbol{\lambda}_i = (\lambda_i^{(1)}, \ldots, \lambda_i^{(n)})$. Each $\boldsymbol{\lambda}_i$ is called a *joint diagonal coefficient* of $\mathbf{A}$.

1.3.4. **Theorem.** If $\mathbf{A} = (A_1, \ldots, A_n)$ is an n-tuple of commuting linear transformations on an N-dimensional vector space $\mathcal{X}$, then

$$\sigma_{\mathrm{T}}(\mathbf{A}) = \sigma_{\mathrm{pt}}(\mathbf{A}) = \{\boldsymbol{\lambda}_i : i = 1, \ldots, N\}.$$

Proof The proof is achieved by repeated application of Lemma 1.3.1 to the simultaneous upper-triangularization of Theorem 1.3.2. It is easy to see that the n-tuples $\boldsymbol{\lambda}_i = (\lambda_i^{(1)}, \ldots, \lambda_i^{(n)})$ $i = 1, \ldots, N$ are joint eigenvalues of the tuple $\mathbf{A}$. So $\{\boldsymbol{\lambda}_i, i = 1, \ldots, N\} \subseteq \sigma_{\mathrm{pt}}(\mathbf{A}) \subseteq \sigma_{\mathrm{T}}(\mathbf{A})$. Now let $\mathcal{X}_2$ be the space spanned by $x_2, \ldots, x_N$. Then $\mathcal{X}_2$ is $N-1$ dimensional and $\mathcal{X} = L_1 \oplus \mathcal{X}_2$. For $j = 1, \ldots, n$, define the linear transformations C_j on $\mathcal{X}_2$ by the $(N-1) \times (N-1)$ matrix:

$$\begin{pmatrix} \lambda_2^{(j)} & \cdots & a_{2N}^{(j)} \\ & \ddots & \\ 0 & \cdots & \lambda_N^{(j)} \end{pmatrix}.$$

$\mathbf{C} = (C_1, \ldots, C_n)$ is a commuting tuple since $\mathbf{A}$ is so. By the result of Lemma 1.3.1, $\sigma_{\mathrm{T}}(\mathbf{A}) \subseteq \{\lambda_1\} \cup \sigma_{\mathrm{T}}(\mathbf{C})$. Repeating this argument N times, $\sigma_{\mathrm{T}}(\mathbf{A}) \subseteq \{\lambda_1, \ldots, \lambda_n\}$. Hence $\sigma_{\mathrm{T}}(\mathbf{A}) = \sigma_{\mathrm{pt}}(\mathbf{A}) = \{\lambda_1, \ldots, \lambda_n\}$. ∎

1.4 Partition of A Tuple

In this section we introduce the important concept of partition of a tuple of operators which will be used often in the later chapters. Following [McPR] we introduce the following notion.

1.4.1. **Definition.** For $\mathbf{A} \in \mathcal{L}(\mathcal{X})^n_{\mathrm{com}}$ the set $\gamma(\mathbf{A}) \subset I\!R^n$ is defined as

$$\gamma(\mathbf{A}) \stackrel{\mathrm{def}}{=} \{\lambda \in I\!R^n : \sum_{j=1}^{n} (A_j - \lambda_j)^2 \text{ is not invertible in } \mathcal{L}(\mathcal{X})\}.$$

If $n = 1$ then it is easily seen that $\gamma(A) = \sigma(A) \cap I\!R$ and hence $\gamma(A)$ can be empty. The relevance of the set $\gamma(A)$ can be readily seen from the following proposition.

1.4.2. **Proposition.** Let $\mathbf{A} \in \mathcal{L}(\mathcal{X})^n_{\mathrm{com}}$ and let $\sigma^*(\mathbf{A})$ be a subset of $\mathbf{C}^n$ with the property that $p(\sigma^*(\mathbf{A})) = \sigma(p(\mathbf{A}))$ for all polynomials $p : \mathbf{C}^n \to \mathbf{C}$. Then $\sigma^*(\mathbf{A}) \cap I\!R^n \subseteq \gamma(\mathbf{A})$ with equality if and only if $\sigma^*(\mathbf{A}) \subseteq I\!R^n$.

PROOF For each $\mu \in I\!R^n$ we define a polynomial $q_{\mu} : \mathbf{C}^n \to \mathbf{C}$ by the formula

$$q_{\mu}(z) = \sum_{i=1}^{n} (z_i - \mu_i)^2. \tag{1.4.3}$$

If $\lambda \in \sigma^*(\mathbf{A}) \cap I\!R^n$ then $0 \in q_{\lambda}(\sigma^*(\mathbf{A})) = \sigma(q_{\lambda}(\mathbf{A}))$ and hence $\lambda \in \gamma(\mathbf{A})$ by definition of $\gamma(\mathbf{A})$. If $\sigma^*(\mathbf{A}) \subseteq I\!R^n$ and $\lambda \in \gamma(\mathbf{A})$ then $0 \in q_{\lambda}(\sigma^*(\mathbf{A}))$ so that $q_{\lambda}(z) = 0$ for some $z \in \sigma^*(\mathbf{A}) \subseteq I\!R^n$. This is possible only if $z = \lambda$ and hence $\lambda \in \sigma^*(\mathbf{A})$. ∎

Since σ_{T} and σ_{H} possess the spectral mapping property for polynomials, the above proposition implies that

$$\sigma^*(\mathbf{A}) \cap I\!R^n \subseteq \gamma(\mathbf{A}) \tag{1.4.4}$$

where σ^* denotes either σ_{T} or σ_{H}. Equality holds if and only if $\sigma^*(\mathbf{A}) \subseteq I\!R^n$.

1.4.5. **Definition.** An n-tuple $\mathbf{A}$ of elements from $\mathcal{L}(\mathcal{X})$ is said to be *strongly commuting* if for each $1 \leq i \leq n$ there exist operators U_i and V_i, each with real spectrum, such that $A_i = U_i + iV_i$ and $\Pi(\mathbf{A}) = (U_1, \ldots, U_n, V_1, \ldots, V_n)$ is a commuting $2n$-tuple. $\Pi(\mathbf{A})$ is called a *partition* of $\mathbf{A}$.

In practice, the spectra $\sigma_{\mathrm{T}}, \sigma_{\mathrm{H}}, \sigma'$ and σ'' are often difficult to compute. One of the appealing aspects of the spectral set γ is that it readily lends itself to explicit computation. The following proposition shows how it helps in computing the Taylor spectrum and the Harte spectrum. This proposition will be of utmost interest in relation to perturbation of complex spectra.

1.4.6. **Proposition.** Let $\mathbf{A}$ be a *strongly commuting* n-tuple with a partition $\Pi(\mathbf{A})$ and let $p : \mathbf{C}^{2n} \to \mathbf{C}^n$ be the polynomial given by

$$p(z_1, \ldots, z_{2n}) = (z_1 + iz_{n+1}, \ldots, z_n + iz_{2n}).$$

Then

$$\sigma_{\mathrm{T}}(\mathbf{A}) = \sigma_{\mathrm{H}}(\mathbf{A}) = p(\gamma(\Pi(\mathbf{A}))).$$

PROOF We let σ^* denote either σ_{T} or σ_{H}. Then it follows from the remarks following Proposition 1.4.2 that $\sigma^*(\Pi(\mathbf{A})) = \gamma(\Pi(\mathbf{A}))$. Then we have

$$\sigma^*(\mathbf{A}) = \sigma^*(p(\Pi(\mathbf{A}))) = p(\sigma^*(\Pi(\mathbf{A}))) = p(\gamma(\Pi(\mathbf{A}))). \quad \blacksquare$$

The results of this section will be useful in Section 4.3 to find perturbation bounds for commuting tuples with complex spectra.

1.5 The Clifford Algebra

McIntosh and Pryde used the Clifford algebra and the Clifford operator to study joint spectrum in [McP]. We will be following their approach to obtain perturbation bounds on joint spectrum of commuting matrices in Chapter 4. The basic facts about the Clifford algebra are briefly described below.

1.5.1. **Definition.** The *Clifford algebra* $I\!R_{(n)}$, generated by $I\!R^n$ is a 2^n-dimensional linear space. Its basis elements are indexed by subsets of the set $M = \{1, \ldots, n\}$. Let $\mathcal{E} = \{h_S : S \subseteq M\}$ be a basis for $I\!R_{(n)}$. The vector space $I\!R_{(n)}$ is made into an algebra by defining a multipliction for these basis elements by

$$(1.5.2) \qquad h_S h_T = \prod_{s\in S, t\in T} (s,t) \ h_{S\Delta T}$$

where $(s,t) = -1$ if $s \leq t$ and $+1$ if $s > t$; and $S\Delta T$ is the symmetric difference of the sets S and T defined by $S\Delta T = (S \cup T) \setminus (S \cap T)$.

In particular $h_\emptyset$ is the identity for this algebra and

$$(1.5.3) \qquad h_{\{j\}}^2 = -h_\emptyset, \ \ h_{\{i\}}h_{\{j\}} = -h_{\{j\}}h_{\{i\}}; \quad i,j = 1, \ldots, n.$$

and if

$$(1.5.4) \qquad S = \{i_1, \ldots, i_k\} \quad i_1 < i_2 < \ldots < i_k \ \text{ then } \ h_S = h_{\{i_1\}} \cdots h_{\{i_k\}}$$

Let $\{h_1, \ldots, h_n\}$ be the standard basis of $I\!R^n$. Identifying h_j with $h_{\{j\}}$, $I\!R^n$ can be thought of as a subspace of $I\!R_{(n)}$. So $(x_1, \ldots, x_m)$ is identified with $\sum x_j h_{\{j\}}$. Since

$$(1.5.5) \qquad (\sum x_j h_{\{j\}})(\sum x_j h_{\{j\}}) = (-\sum x_j^2) h_\emptyset$$

and $h_\emptyset$ is the identity of $\mathcal{R}_{(n)}$, any non-zero element of $\mathcal{R}^n$ is invertible in $\mathcal{R}_{(n)}$.

For any two elements $\lambda = \sum \lambda_S h_S$ and $\mu = \sum \mu_S h_S$ of $I\!R_{(n)}$, let $< \lambda, \mu >$ $= \sum \lambda_S \mu_S$. This defines an inner product in which the basis $\{h_S : S \subset M\}$ is orthonormal.

Let $\mathcal{X}$ be a finite-dimensional vector space over $\mathbf{C}$. The tensor product $\mathcal{X} \otimes \mathcal{R}_{(n)}$ is a finite-dimensional vector space whose elements can be represented as $\sum x_S \otimes h_S$. If $\mathcal{X}$ has an inner product $< \cdot, \cdot >$ the space $\mathcal{X} \otimes \mathcal{R}_{(n)}$ naturally inherits it:

$$(1.5.6) \qquad < \sum x_S \otimes h_S, \sum y_S \otimes h_S >= \sum < x_S, y_S >$$

Let $I\!M_k$ be the algebra of all $k \times k$ matrices. Let $\{A_S : S \subset M\}$ be any collection of 2^n matrices from $I\!M_k$. Then $\sum A_S \otimes h_S$ is an element of $I\!M_k \otimes I\!R_{(n)}$. It can be thought of as a linear operator on $\mathbf{C}^k \otimes I\!R_{(n)}$ if we define its action on $\sum \lambda_T \otimes h_T$ by

$$(\sum_S A_S \otimes h_S)(\sum_T \lambda_T \otimes h_T) = \sum_{S,T} A_S \lambda_T \otimes h_S h_T \tag{1.5.7}$$

So $I\!M_k \otimes I\!R_{(n)}$ is a subalgebra of the algebra of all linear operators on $\mathbf{C}^k \otimes I\!R_{(n)}$ and $I\!M_k$ is a subalgebra of $I\!M_k \otimes I\!R_{(n)}$ by the identification of the matrix A with the operator $A \otimes h_\emptyset$ on $I\!M_k \otimes I\!R_{(n)}$.

1.5.8. **Definition.** Given an n-tuple $\mathbf{A} = (A_1, \ldots, A_n)$ of $k \times k$ matrices, the *Clifford operator* of $\mathbf{A}$ is an element of $I\!M_k \otimes I\!R_{(n)}$ defined as

$$\mathrm{Cliff}(\mathbf{A}) \stackrel{\mathrm{def}}{=} \imath \sum A_j \otimes h_{\{j\}}.$$

Basic properties of the Clifford operator are summarised below.

1.5.9. **Lemma.** If the tuple $\mathbf{A} = (A_1, \ldots, A_n)$ consists of commuting matrices, then

(i) $\mathrm{Cliff}(\mathbf{A})^2 = \sum A_j^2$.

(ii) For any $\lambda \in \mathbf{C}^n$ we define $\mathbf{A} - \lambda = (A_1 - \lambda_1, \ldots, A_n - \lambda_n)$. The operator $\mathrm{Cliff}(\mathbf{A} - \lambda)$ is invertible if and only if $\sum (A_j - \lambda_j)^2$ is invertible and in such a case $\mathrm{Cliff}(\mathbf{A} - \lambda)^{-1} = (\sum (A_j - \lambda_j)^2)^{-1} \mathrm{Cliff}(\mathbf{A} - \lambda)$.

(iii) If each A_j is self-adjoint, then $\mathrm{Cliff}(\mathbf{A})$ is self-adjoint and

$$||\mathrm{Cliff}(\mathbf{A})|| = r(\mathbf{A}).$$

PROOF (i)

$$\begin{aligned}
\mathrm{Cliff}(\mathbf{A})^2 &= (-1)(\sum_{i=1}^n A_i \otimes h_{\{i\}}) \times (\sum_{j=1}^n A_j \otimes h_{\{j\}}) \\
&= (-1) \sum_{i,j=1}^n A_i A_j \otimes h_{\{i\}} h_{\{j\}} \\
&= (-1) \sum_{j=1}^n A_j^2 \otimes h_\emptyset + \sum_{i<j} A_i A_j \otimes h_{\{i\}} h_{\{j\}}
\end{aligned}$$

The first term is identified with $\sum A_j^2$ and the second term is zero by the multiplication rule.

(ii) We note from part (i) that $\mathrm{Cliff}(\mathbf{A}-\boldsymbol{\lambda})^2 = \sum(A_j - \lambda_j)^2$. So $\mathrm{Cliff}(\mathbf{A}-\boldsymbol{\lambda})$ is invertible if and only if $\sum(A_j-\lambda_j)^2$ is invertible and in that case Cliff $(\mathbf{A}-\boldsymbol{\lambda})^{-1} = (\sum(A_j-\lambda_j)^2)^{-1}\mathrm{Cliff}(\mathbf{A}-\boldsymbol{\lambda})$.

(iii) A straightforward computation shows that if I is the $k\times k$ identity matrix and $j \in M$ is any index then $(I\otimes h_{\{j\}})^* = -I\otimes h_{\{j\}}$. So for any $B \in \mathbb{M}_k$ we have $(B\otimes h_{\{j\}})^* = ((B\otimes h_\emptyset)(I\otimes h_{\{j\}}))^* = -((I\otimes h_{\{j\}})(B^*\otimes h_\emptyset)) = -(B^*\otimes h_{\{j\}})$. Hence,

$$\mathrm{Cliff}(\mathbf{A})^* = (-i)\sum(-A_j^*)\otimes h_{\{j\}} = \mathrm{Cliff}(\mathbf{A}^*),$$

where $\mathbf{A}^*$ is the tuple $(A_1^*,\ldots,A_n^*)$. This proves the first part. For the second part note that $\|\mathrm{Cliff}(\mathbf{A})\|^2 = \|(\mathrm{Cliff}(\mathbf{A}))^2\| = \|\sum A_j^2\| = r(\mathbf{A})$. ∎

Notes and References. Algebraic joint spectra have been studied by Arens [Are], Calderon [ArCa], Waelbroeck [Wa] and others. For an excellent account of the development of the theory and its interplay with several variables complex analysis, see Wermer [We]. Among the many attempts to define spatial joint spectra (see especially [Das1], [Das2] and [Har1], [Har2], [Har3]), Taylor's definition turned out to be the most natural. Taylor, in his papers [T1], [T2], [T3] and [T4], developed joint spectral theory and the associated functional calculus. For alternative proofs of the fact that Taylor joint spectrum coincides with the joint eigenvalues in the finite-dimensional case, see [CT] and [McPR]. Clifford algebra and Clifford operators were first used to study joint spectra by McIntosh and Pryde in [McP].

Chapter 2

Commuting Compact Tuples

In this Chapter we study commuting n-tuples of compact operators on an infinite-dimensional Banach space $\mathcal{X}$. The results are mainly of two types and accordingly they are grouped into two sections. The first section deals with the spectral properties of the tuple. We first show that any non-zero point in the Taylor joint spectrum of a commuting compact tuple is a joint eigenvalue. Then we show that every non-zero joint eigenvalue of a commuting compact tuple has a finite algebraic multiplicity. In the second section we obtain a simultaneous upper-triangularization of a commuting compact tuple. Here we use many of the results of the first section. The joint diagonal coefficients in this joint upper-triangular form are then shown to be in one-one correspondence with non-zero joint eigenvalues. This is a generalization of well-known results for single operators (see [Ri3]) and also of the finite-dimensional results of Section 1.3.

2.1 Spectral Properties

In Section 1.3 we have seen that the joint spectrum of a commuting tuple of matrices is the same as the set of its joint eigenvalues. In the present section we are going to prove that the spectral properties of an n-tuple of commuting compact operators on an infinite dimensional Banach space resemble those of an n-tuple of matrices to a large extent. We begin with the fact that every non-zero element in the Taylor joint spectrum of a commuting compact n-tuple is a joint eigenvalue.

2.1.1. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an n-tuple of commuting compact operators on a complex Banach space $\mathcal{X}$. Let $\sigma_T(\mathbf{A})$ be the Taylor joint spectrum of the tuple $\mathbf{A}$. Let $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ be a non-zero n- tuple of scalars. If $\boldsymbol{\lambda} \in \sigma_T(\mathbf{A})$ then $\boldsymbol{\lambda}$ is a joint eigenvalue of $\mathbf{A}$.

PROOF The Taylor joint spectrum $\sigma_T(\mathbf{A})$ of an n-tuple of commuting compact operators is countable because $\sigma_T(\mathbf{A}) \subseteq \sigma(A_1) \times \cdots \times \sigma(A_n)$ and each $\sigma(A_j)$ is countable. Moreover, the only possible limit point of $\sigma_T(\mathbf{A})$ is $(0, 0, \ldots, 0)$. Indeed, if $\boldsymbol{\mu}$ is a limit point of $\sigma_T(\mathbf{A})$, then there exists a sequence $\boldsymbol{\mu}_m = (\mu_1^{(m)}, \ldots, \mu_n^{(m)})$ of points in $\sigma_T(\mathbf{A})$ such that $||\boldsymbol{\mu}_m - \boldsymbol{\mu}|| \to 0$. So $|\mu_j^{(m)} - \mu_j| \to 0$ for each j. Since $\mu_j^{(m)} \in \sigma(A_j)$, it follows that $\mu_j = 0$ for each j.

So $\boldsymbol{\lambda}$ is not a limit point of $\sigma_T(\mathbf{A})$ and we can find a neighbourhood N of $\boldsymbol{\lambda}$ containing no other point of $\sigma_T(\mathbf{A})$. Let $f : \mathbb{C}^n \to \mathbb{C}$ be a holomorphic function which is 1 on N and 0 on $\sigma_T(\mathbf{A}) \backslash \{\boldsymbol{\lambda}\}$. Then $f(\mathbf{A})$ is a projection which commutes with each A_j. So $\mathcal{R}(f(\mathbf{A}))$ is an invariant subspace for each A_j. And $\sigma_T(\mathbf{A}|_{\mathcal{R}(f(\mathbf{A}))}) = \{\boldsymbol{\lambda}\}$. It follows that $\sigma(A_j|_{\mathcal{R}(f(\mathbf{A}))}) = \{\lambda_j\}$ for all $j = 1, \ldots, n$. Since $\boldsymbol{\lambda} \neq 0$, at least one λ_j (say λ_{j_0}) is non-zero. By the projection property of the Taylor joint spectrum, $\sigma(A_{j_0}|_{\mathcal{R}(f(\mathbf{A}))}) = \{\lambda_{j_0}\}$. So $A_{j_0}|_{\mathcal{R}(f(\mathbf{A}))}$ is an invertible compact operator. Consequently, $\mathcal{R}(f(\mathbf{A}))$ is finite-dimensional. Since on a finite-dimensional Banach space the joint spectrum consists of the joint eigenvalues, we conclude that $\boldsymbol{\lambda}$ is a joint eigenvalue of $\mathbf{A}|_{\mathcal{R}(f(\mathbf{A}))}$, and hence of $\mathbf{A}$. ■

Let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{B} = (B_1, \ldots, B_m)$ be two tuples of commuting bounded operators on a Banach space $\mathcal{X}$. We define the product $\mathbf{AB}$ to be the nm-tuple whose entries are $A_iB_j, 1 \leq i \leq n, 1 \leq j \leq m$, arranged in lexicographic order. Using this multiplication rule, one can successively define the powers $\mathbf{A}^2, \mathbf{A}^3, \ldots$. Then $\mathbf{A}^m$ is an operator tuple with n^m entries; these are the products $A_{i_1} \ldots A_{i_m}$ where the indices are chosen from $\{1, \ldots, n\}$ with repetitions allowed, and are then arranged lexicographically.

The n-tuple $\mathbf{A}$ can be regarded as an operator from $\mathcal{X}^m$ to $\mathcal{X}^{nm}$, taking a vector $\mathbf{x} = (x_1, \ldots, x_m) \in \mathcal{X}^m$ to the nm -tuple $(A_1x_1, \ldots, A_nx_1, A_1x_2, \ldots, A_nx_m)$. When $m = 1$, the null space of this operator, denoted by $\mathcal{N}(\mathbf{A})$, is the set of all joint eigenvectors of $\mathbf{A}$. With the above definition of powers of a tuple, we define, for any $k \geq 1$, the subspaces $\mathcal{N}_k(\mathbf{A})$ of $\mathcal{X}$ by

$$\mathcal{N}_k(\mathbf{A}) = \mathcal{N}(\mathbf{A}^k). \tag{2.1.2}$$

2.1.3. **Lemma.** $\mathcal{N}_k(\mathbf{A}) \subseteq \mathcal{N}_{k+1}(\mathbf{A})$ for all $k = 1, 2, \ldots$. If $\mathcal{N}_{k+1}(\mathbf{A}) = \mathcal{N}_k(\mathbf{A})$ for some k, then $\mathcal{N}_{k+l}(\mathbf{A}) = \mathcal{N}_k(\mathbf{A})$ for all $l = 1, 2, \ldots$.

PROOF The first statement is evident. To prove the second, take any $x \in \mathcal{N}_{k+2}(\mathbf{A})$. Then $\mathbf{A}^{k+2}x = 0$. So $\mathbf{A}^{k+1}\mathbf{A}x = 0$. This implies that $A_jx \in \mathcal{N}_{k+1}(\mathbf{A})$ for all $j = 1, \ldots, n$. But $\mathcal{N}_{k+1}(\mathbf{A}) = \mathcal{N}_k(\mathbf{A})$. So $A_jx \in \mathcal{N}_k(\mathbf{A})$ for all $j = 1, \ldots, n$. In other words, $\mathbf{A}^{k+1}x = 0$. So $x \in \mathcal{N}_{k+1}(\mathbf{A})$. Hence $\mathcal{N}_{k+2}(\mathbf{A}) = \mathcal{N}_{k+1}(\mathbf{A})$. Now the rest follows by induction. ∎

2.1.4. **Lemma.** Let $\mathbf{A} = (A_1, \ldots A_n)$ be an n-tuple of commuting compact operators and $\boldsymbol{\lambda}$ a non-zero scalar n-tuple. Then $\mathcal{N}_k(\mathbf{A} - \boldsymbol{\lambda})$ is finite-dimensional for all $k \geq 1$. Moreover, there exists a positive integer ν such that

$$\{0\} = \mathcal{N}_0(\mathbf{A} - \boldsymbol{\lambda}) \subsetneq \mathcal{N}_1(\mathbf{A} - \boldsymbol{\lambda}) \subsetneq \cdots \subsetneq \mathcal{N}_\nu(\mathbf{A} - \boldsymbol{\lambda}) = \mathcal{N}_{\nu+1}(\mathbf{A} - \boldsymbol{\lambda}) = \cdots$$

PROOF Since $\boldsymbol{\lambda} \neq 0$, for at least one i, $\lambda_i \neq 0$. Assume without loss of generality that $\lambda_1 \neq 0$. By definition of the null space of a tuple,

$$\mathcal{N}_k(\mathbf{A}) \subseteq \mathcal{N}(A_1 - \lambda_1)^k.$$

Since A_1 is a compact operator and $\lambda_1 \neq 0$, it is known that $\dim \mathcal{N}(A_1 - \lambda_1)^k < \infty$ for all $k \geq 1$. Moreover, $\dim \mathcal{N}(A_1 - \lambda_1)^k$ is bounded as $k \to \infty$ (see [Ta], pages 278-280). That completes the proof in view of Lemma 2.1.3. ∎

Thus the integer ν occuring in Lemma 2.1.4 is the smallest n such that $\mathcal{N}_n(\mathbf{A}-\boldsymbol{\lambda}) = \mathcal{N}_{n+1}(\mathbf{A}-\boldsymbol{\lambda})$. We call the integer ν the *index* and $\dim(\mathcal{N}_\nu(\mathbf{A}-\boldsymbol{\lambda}))$ the ***algebraic multiplicity*** of $\boldsymbol{\lambda}$ with respect to $\mathbf{A}$. Theorem 2.1.1 and Lemma 2.1.4 together show that any non-zero point in the Taylor spectrum of a commuting compact tuple is a joint eigenvalue with finite algebraic multiplicity.

2.2 Simultaneous Upper-triangularization

We have seen in Section 1.3 that if $\mathbf{A} = (A_1, \ldots, A_n)$ is an n-tuple of commuting linear operators acting on a Banach space $\mathcal{X}$ with $\dim(\mathcal{X}) = N < \infty$ then there exist subspaces $L_0, L_1, \ldots, L_N$ of $\mathcal{X}$ such that

(i) $\{0\} = L_0 \subseteq L_1 \subseteq \cdots \subseteq L_N = \mathcal{X}$,

(ii) L_k is k-dimensional ($k = 1, \ldots, N$),

(iii) each L_k is simultaneously invariant under $A_1, \ldots, A_n$.

So a family of subspaces $\{L_1, \ldots, L_N\}$, which has the properties (i), (ii) and (iii) above, determines an upper-triangular representation of $A_1, \ldots A_n$. One can choose a basis $\mathcal{E} = \{x_1, \ldots, x_N\}$ of $\mathcal{X}$ which has the properties:

(i) each x_j lies in L_j but not in L_{j-1},

(ii) the matrix of the operators $A_1, \ldots, A_n$ with respect to the basis $\mathcal{E}$ are upper-triangular.

The joint diagonal coefficients are then defined with the help of this simultaneous upper-triangularization. It was shown in Theorem 1.3.4 that $\sigma_{\mathrm{T}}(\mathbf{A}) = \sigma_{\mathrm{pt}}(\mathbf{A}) =$ the set of all joint diagonal coefficients of $\mathbf{A}$. Now we will obtain an extension of this to commuting compact linear operators acting on infinite-dimensional spaces.

The theory for a single operator has been completely charted out in Ringrose[Ri3]. We shall briefly recapitulate it.

Throughout this section $\mathcal{X}$ stands for a complex infinite dimensional Banach space. The set $\mathcal{L}$ of all closed subspaces of $\mathcal{X}$ is a partially ordered set under inclusion. A completely ordered subset of this set is called a chain. The class $\mathcal{C}$ of all chains is again a partially ordered set by the inclusion relation on the subsets of $\mathcal{L}$. Let $\mathcal{C}_0$ be a *completely ordered subset of* $\mathcal{C}$. Define $\mathcal{F}_0 = \cup\{\mathcal{F} : \mathcal{F} \in \mathcal{C}_0\}$. Then $\mathcal{F}_0$ is a chain because if $L_1, L_2 \in \mathcal{F}_0$, then there exist $\mathcal{F}_1, \mathcal{F}_2 \in \mathcal{C}_0$ such that $L_1 \in \mathcal{F}_1$ and $L_2 \in \mathcal{F}_2$. Since $\mathcal{C}_0$ is completely ordered, we can assume $\mathcal{F}_1 \subseteq \mathcal{F}_2$. So $L_1, L_2 \in \mathcal{F}_2$. Since $\mathcal{F}_2$ is a chain, either $L_1 \subseteq L_2$ or $L_2 \subseteq L_1$.

The chain $\mathcal{F}_0$ is obviously an upper bound for the class $\mathcal{C}_0$. Moreover if $\mathcal{G}$ is any other upper bound, then $\mathcal{F}_0 \subseteq \mathcal{G}$. So $\mathcal{F}_0$ is the least upper bound of the class $\mathcal{C}_0$. So each totally ordered subset of $\mathcal{C}$ has a least upper bound. It follows from Zorn's lemma that $\mathcal{C}$ contains maximal elements, which we call *maximal chains.* Every chain is contained in at least one maximal chain.

Given a subfamily $\mathcal{F}_0$ of a chain $\mathcal{F}$, the set $\cap\{L : L \in \mathcal{F}_0\}$ is a closed subspace of $\mathcal{X}$, the same is true for $\overline{\cup\{L : L \in \mathcal{F}_0\}}$. Given $M \in \mathcal{F}$, we define the subspace M_- as

(2.2.1) $$M_- = \overline{\cup\{L \in \mathcal{F} : L \subsetneq M\}},$$

interpreting the right hand side as $\{0\}$ when there is no proper subspace of M in $\mathcal{F}$.

2.2.2. **Definition.** A chain $\mathcal{F}$ is called a *simple chain* if it satisfies the following conditions :

(i) $\mathcal{F}$ contains the subspaces $\{0\}$ and $\mathcal{X}$,
(ii) if $\mathcal{F}_0$ is a subfamily of $\mathcal{F}$ then $\cap\{L : L \in \mathcal{F}_0\}$ and $\overline{\cup\{L : L \in \mathcal{F}_0\}}$ are in $\mathcal{F}$,
(iii) for each $M \in \mathcal{F}$, $\dim(M/M_-)$ is at most one.

Note that condition (ii) implies that $M_- \in \mathcal{F}$ for each $M \in \mathcal{F}$.

2.2.3. **Theorem.** A simple chain is maximal.

PROOF. See [Ri3, page 167].

If $T \in \mathcal{L}(\mathcal{X})$, then a chain $\mathcal{F}$ is called invariant under T if each $M \in \mathcal{F}$ is an invariant subspace of T. It is well-known (see [AS]) that if T is compact, then T has a non-trivial closed invariant subspace i.e., there exists a closed subspace which is neither $\{0\}$ nor $\mathcal{X}$ and which is left invariant by T. This shows the existence of non-trivial invariant chains for a compact operator T. Let $\mathcal{C}_T$ denote the class of all invariant chains of T. The usual Zorn's lemma argument applied to $\mathcal{C}_T$ shows the existence of maximal elements of $\mathcal{C}_T$ which we call *maximal invariant chains*. It is not apparent that maximal invariant chains are also maximal chains i.e., elements which are maximal in $\mathcal{C}_T$ are maximal in $\mathcal{C}$. The following theorem (see [Ri3], page 169) shows that this is indeed the case.

2.2.4. **Theorem.** Let T be a compact operator on $\mathcal{X}$. Then every maximal invariant chain of T is simple.

A corollary of this theorem is that every maximal chain is simple. This is so because every maximal chain is a maximal invariant chain for the zero operator. Of course, for this operator one does not have to appeal to the Aronszajn-Smith theorem for the existence of maximal chains.

Now we know that there exists a simple chain $\mathcal{F}$ of closed subspaces of $\mathcal{X}$ such that each L in $\mathcal{F}$ is invariant under T. If $M \in \mathcal{F}$ then either $M = M_-$ or M/M_- has dimension one. In the later case, suppose $z_M \in M \setminus M_-$, so that M is the linear span of $\{z_M\} \cup M_-$. Since M is invariant under T, the vector $Tz_M \in M$, so that there exists a scalar α^M and a vector $y_M \in M_-$ such that

$$Tz_M = \alpha^M z_m + y_M.$$

The scalar α^M does not depend on the choice of z_M in $M \setminus M_-$. Since M_- is invariant under $T - \alpha^M$ and $(T - \alpha^M)z_M \in M_-$ it follows that

$$(T - \alpha^M)M \subseteq M_-. \tag{2.2.5}$$

When $M = M_-$ we define $\alpha^M = 0$. In this way we associate a scalar α^M with every subspace M in $\mathcal{F}$.

2.2.6. **Definition.** The scalar α^M as defined above is called the *diagonal coefficient* of T at M.

2.2.7. **Theorem.** (i) Let $\mathcal{F}$ be as above. Suppose that λ is a non-zero eigenvalue of T, x a non-zero eigenvector satisfying $Tx = \lambda x$ and

$$M = \cap\{L : L \in \mathcal{F}, x \in L\}.$$

Then $M \in \mathcal{F}, x \in M \backslash M_{-}$ and $\alpha^M = \lambda$.

(ii) If $M \in \mathcal{F}$ and $\alpha^M \neq 0$, then α^M is an eigenvalue of T. If α^M has index 1 relative to T there is a vector $x \in M \backslash M_{-}$ satisfying $Tx = \lambda x$.

PROOF See [Ri3, pages 172 - 173].

Now we have all the material needed to investigate the relation between the spectral properties and the simultaneous upper-triangularizability of a commuting compact tuple. We proceed along the route sketched above for a single compact operator. Just as the theorem of Aronszajn and Smith[AS], asserting the existence of proper closed invariant subspaces for compact operators, is the starting point of the whole theory in [Ri3], we have taken as our starting point the theorem, due to Lomonosov, that a compact linear operator T acting on a Banach space $\mathcal{X}$ has a proper closed hyperinvariant subspace. It follows that for a commuting n-tuple $\mathbf{A} = (A_1, \ldots, A_n)$ of compact operators there exists a proper closed simultaneously invariant subspace. This result, along with Zorn's lemma implies the existence of a maximal totally ordered family $\mathcal{F}$ of closed subspaces of $\mathcal{X}$, each of which is invariant under $A_1, \ldots, A_n$. We then obtain analogues of the finite dimensional results concerning the joint eigenvalues of $\mathbf{A}$.

To ensure the existence of a proper closed subspace which is simultaneously invariant under a familly of commuting compact operators we need the following theorem. The theorem is due to Lomonosov[L]. But the proof was considerably simplified by Hilden following Lomonosov's original ideas. Here we reproduce this elementary proof from [Mic].

A subspace M of $\mathcal{X}$ is called *hyperinvariant* under the operator T if M is left invariant by all bounded operators which commute with T.

2.2.8. **Theorem (Lomonosov).** *Every compact operator T on $\mathcal{X}$ has a proper* closed hyperinvariant subspace.

PROOF (Hilden) If T has a non-zero eigenvalue λ then the space $\{x \in \mathcal{X} : Tx = \lambda x\}$ is the required space. So assume T does not have any non-zero eigenvalue i.e., $\sigma(T) = \{0\}$.

Without loss of generality assume $\|T\| = 1$. Choose $x_0 \in \mathcal{X}$ such that $\|Tx_0\| > 1$. Let $\mathcal{D} = \{x \in \mathcal{X} : \|x - x_0\| \leq 1\}$. Obviously $0 \notin \mathcal{D}$. Moreover $0 \notin \overline{T\mathcal{D}}$. Indeed for any $x \in \mathcal{D}, \|Tx - Tx_0\| \leq \|T\|. \|x - x_0\| \leq 1$. So if $\lim_{n\to\infty} Tx_n = 0$ for some sequence $\{x_n\}$ from $\mathcal{D}$, then $\|Tx_0\| = \lim_{n\to\infty} \|Tx_n - Tx_0\| \leq 1$. But $\|Tx_0\| > 1$. So $0 \notin \overline{T\mathcal{D}}$.

For any $y \in \mathcal{X}$, let $\mathcal{M}_y = \{Ay : A \in \mathcal{L}(\mathcal{X})$ and $AT = TA\}$. For $y \neq 0, \mathcal{M}_y \neq \{0\}$. So it remains to show that for some y, $\mathcal{M}_y$ is not dense in $\mathcal{X}$. Suppose the contrary, i.e., $\mathcal{M}_y$ is dense for all non-zero $y \in \mathcal{X}$. Then for any $y \neq 0$ there exists an $A \in \mathcal{L}(\mathcal{X})$ such that $\|Ay - x_0\| < 1$. Let $\mathcal{U}(A) = \{y \in \mathcal{X} : \|Ay - x_0\| < 1\}$. Each $\mathcal{U}(A)$ is open, and the union of all $\mathcal{U}(A)$ such that A commutes with T is $\mathcal{X}\backslash\{0\}$. Since $\overline{T\mathcal{D}} \subseteq \mathcal{X}\backslash\{0\}$, so $\mathcal{U}(A)$'s form an open cover of $\overline{T\mathcal{D}}$. But $\overline{T\mathcal{D}}$ is compact because T is compact and $\mathcal{D}$ is bounded. So there are $A_1, \ldots, A_n$ such that $\overline{T\mathcal{D}} \subseteq \mathcal{U}(A_1) \cup \cdots \cup \mathcal{U}(A_n)$. So $T\mathcal{D} \subseteq \mathcal{U}(A_1) \cup \cdots \cup \mathcal{U}(A_n)$.

So, for any $x \in \mathcal{D}, Tx \in \mathcal{U}(A_1) \cup \cdots \cup \mathcal{U}(A_n)$. So, there exists i_1 such that $Tx \in \mathcal{U}(A_{i_1})$. This, by definition of $\mathcal{U}(A_{i_1})$, means that $A_{i_1}Tx \in \mathcal{D}$. So, there exists i_2 such that $TA_{i_1}Tx \in \mathcal{U}(A_{i_2})$. So, $A_{i_2}TA_{i_1}Tx \in \mathcal{D}$. Continuing this m times, $A_{i_m}T \ldots A_{i_1}Tx \in \mathcal{D}$. Let $c = \max\{\|A_1\|, \ldots, \|A_n\|\}$. Since all A_i's commute with T, we have $(c^{-1}A_{i_m}) \ldots (c^{-1}A_{i_1})(cT)^m x \in \mathcal{D}$. So,

$$\begin{aligned} & \|(c^{-1}A_{i_m}) \ldots (c^{-1}A_{i_1})(cT)^m x\| \\ \leq\ & \|c^{-1}A_{i_m}\| \ldots \|c^{-1}A_{i_1}\| \, \|(cT)^m\| \, \|x\| \\ \leq\ & \|(cT)^m\| \, \|x\| \quad \text{by definition of } c. \end{aligned}$$

This goes to 0 by the spectral radius formula. Thus $0 \in \mathcal{D}$. That is a contradiction. ■

It follows from Lomonosov's Theorem that a commuting family of compact operators has a common non-trivial closed invariant subspace. This implies the existence of non-trivial chains consisting of subspaces simultaneously invariant under $A_1, \ldots, A_m$. Let $\mathcal{C}_{\mathbf{A}}$ denote the class of all chains which are simultaneously invariant under each A_j. Using a procedure similar to the one used for establishing the existence of maximal invariant chains for a single compact operator, one can show the existence of maximal elements of $\mathcal{C}_{\mathbf{A}}$, which we call the *maximal simultaneously invariant chains*. Our next lemma shows that these are, in fact, simple chains.

2.2.9. **Lemma.** Each maximal simultaneously invariant chain is simple.

PROOF Suppose $\mathcal{F}$ is a maximal simultaneously invariant chain. Then $\mathcal{F}$ obviously contains the subspaces $\{0\}$ and $\mathcal{X}$. For any subfamily $\mathcal{F}_0$ of $\mathcal{F}$, let $N = \cap\{L : L \in \mathcal{F}_0\}$. Then N is a closed subspace of X. Since each L is simultaneously invariant under each A_j, the same is true for N. Let $M \in \mathcal{F}$. Since $\mathcal{F}$ is totally ordered, either $M \subseteq L$ for each $L \in \mathcal{F}_0$, and hence $M \subseteq N$, or $L \subseteq M$ for at least one L in $\mathcal{F}_0$, and hence $N \subseteq M$. It follows that $\mathcal{F} \cup \{N\}$ is totally ordered by inclusion and is therefore a simultaneously invariant chain. Since $\mathcal{F}$ is maximal, $N \in \mathcal{F}$. Similarly, we can see that the closed subspace $\overline{\cup\{L : L \in \mathcal{F}\}}$ is also a member of $\mathcal{F}$.

It remains to show that M/M_- has dimension one for each $M \in \mathcal{F}$. Suppose $\dim M/M_- > 1$ for some $M \in \mathcal{F}$. Consider the Banach space M/M_- and the n-tuple $\mathbf{A}_0 \in \mathcal{L}(M/M_-)$ defined by $(\mathbf{A}_0)_j(x + M_-) = A_j x + M_-$. Then $\mathbf{A}_0$ is an n-tuple of commuting compact operators. So there is a closed subspace N_0 of M/M_- such that $\{0\} \neq N_0 \neq M/M_-$ and $(\mathbf{A}_0)_j(N_0) \subseteq N_0$ for all $j = 1, \ldots, n$. It follows that if $N = \{x \in M : x + M_- \in N_0\}$ then N is a closed subspace of X, simultaneously invariant under each A_j and $M_- \subsetneqq N \subsetneqq M$. Given any subspace $L \in \mathcal{F}$, either $M \subseteq L$ and so $N \subsetneqq L$, or $L \subsetneqq M$ and so $L \subseteq M_- \subsetneqq N$. Hence $N \notin \mathcal{F}$ and $\mathcal{F} \cup \{N\}$ is a chain. This violates the maximality of $\mathcal{F}$ as a simultaneously invariant chain. So for each $M \in \mathcal{F}$, $\dim M/M_-$ is at most one. ∎

We can now define a joint diagonal coefficient. The above lemma implies the existence of a simple simultaneously invariant chain for the commuting compact tuple $\mathbf{A}$. Let $\mathcal{F}$ be such a chain.

2.2.10. **Definition.** The scalar n-tuple $\boldsymbol{\alpha}^M = (\alpha_1^M, \ldots, \alpha_n^M)$ will be called a *joint diagonal coefficient* of $\mathbf{A}$ at $M \in \mathcal{F}$ if α_j^M is the diagonal coefficient of A_j at M.

2.2.11. **Lemma.** If $\boldsymbol{\lambda}$ is a joint eigenvalue then it is a joint diagonal coefficient.

PROOF There exists a non-zero vector x such that $A_j x = \lambda_j x$ for all $j = 1, \ldots, n$. We define $M = \bigcap\{L \in \mathcal{F} : x \in L\}$. Then by part (i) of Theorem 2.2.7, λ_j is the diagonal coefficient of A_j at M. So $\boldsymbol{\lambda}$ is a joint diagonal coefficient. ∎

2.2.12. **Lemma.** If $\boldsymbol{\alpha}^M \neq 0$ is a joint diagonal coefficient of $\mathbf{A}$ at M then $\boldsymbol{\alpha}^M$ is a joint eigenvalue of $\mathbf{A}$.

PROOF Let $\mathbf{A}_M$ be the restriction of the tuple $\mathbf{A}$ to the invariant subspace M. Then $\mathbf{A}_M$ is an n-tuple of commuting compact operators on the Banach space M. We claim that $\boldsymbol{\alpha}^M$ is a joint eigenvalue of $\mathbf{A}_M$. If not then by Theorem 2.1.1 $\boldsymbol{\alpha}^M \notin \sigma_{\mathrm{T}}(\mathbf{A}_M)$. This means that the Koszul complex $K(\mathbf{A}_M - \boldsymbol{\alpha}^M, M)$ as defined in (1.2.27) is exact. We recall that the exactness of the complex in particular means that $\mathcal{R}(\mathbf{A}_M - \boldsymbol{\alpha}^M) = M$ (see (1.2.31)). But on the other hand, since $\boldsymbol{\alpha}^M \neq 0$ we have $M_- \neq M$ and $\mathcal{R}(A_j - \alpha_j^M) \subseteq M_-$ for all $j = 1, \ldots, n$. Hence $\mathcal{R}(\mathbf{A}_M - \boldsymbol{\alpha}^M) \subseteq M_-$. That is a contradiction. So $\boldsymbol{\alpha}^M$ is a joint eigenvalue of $\mathbf{A}_M$ and hence of $\mathbf{A}$. ∎

2.2.13. **Definition.** Let $\boldsymbol{\lambda}$ be a joint diagonal coefficient of $\mathbf{A}$. Consider the set $\{M \in \mathcal{F} : \boldsymbol{\lambda}$ is the joint diagonal coefficient of $\mathbf{A}$ at $M\}$. The *diagonal multiplicity* of $\boldsymbol{\lambda}$ is the cardinality of this set.

The next lemma relates the diagonal multiplicity of a diagonal coefficient to its algebraic multiplicity as a joint eigenvalue. This generalizes Lemma 4.3.9 of [Ri3]. The proof there can be adapted to the present situation.

2.2.14. **Lemma.** If $\boldsymbol{\alpha}^M = (\alpha_1^M, \ldots, \alpha_n^M)$ is a non-zero joint diagonal coefficient of $\mathbf{A}$ then its diagonal multiplicity is finite and is equal to its algebraic multiplicity as a joint eigenvalue of $\mathbf{A}$.

PROOF Let d denote the diagonal multiplicity, m the algebraic multiplicity and ν the index of α^M relative to $\mathbf{A}$. Then $\mathcal{N}(\mathbf{A}-\alpha^M)$ is m dimensional and

$$\{0\} = \mathcal{N}_0(\mathbf{A}-\alpha^M) \subsetneqq \mathcal{N}_1(\mathbf{A}-\alpha^M) \subsetneqq \cdots \subsetneqq \mathcal{N}_\nu(\mathbf{A}-\alpha^M) = \mathcal{N}_{\nu+1}(\mathbf{A}-\alpha^M) = \cdots.$$

We first reduce the problem to the case $\nu = 1$. For this, define the operator n^ν-tuple $\mathbf{B}$ and the scalar n^ν-tuple $\boldsymbol{\mu}$ by

$$\mathbf{B} - \boldsymbol{\mu} = (\mathbf{A} - \alpha^M)^\nu.$$

Then $\boldsymbol{\mu} = (\alpha^M)^\nu$ and $\mathbf{B}$, being a polynomial in the A_j's, is a tuple of commuting compact operators. For the same reason the invariant subspaces of $\mathbf{A}$ in $\mathcal{F}$ are also invariant under the tuple $\mathbf{B}$. Since $\mathbf{B} - \boldsymbol{\mu}$ and $(\mathbf{B} - \boldsymbol{\mu})^2$ have the same null space $\mathcal{N}_\nu(\mathbf{A}-\alpha^M) = \mathcal{N}_{2\nu}(\mathbf{A}-\alpha^M)$ of dimension m, it follows that $\boldsymbol{\mu}$ is a joint eigenvalue of $\mathbf{B}$ with index one and multiplicity m. So without loss of generality we can assume ν to be 1.

Suppose $d > m$. Then there exist subspaces $M(0), M(1), \ldots, M(m)$ in $\mathcal{F}$ satisfying $M(0) \subsetneqq M(1) \subsetneqq \cdots \subsetneqq M(m)$ and α^M is the joint diagonal coefficient of $\mathbf{A}$ at $M(k)$ for all $k = 1, \ldots, m$. Since $M(k-1) \subsetneqq M(k)$, it follows that $M(k-1) \subseteq M(k)_-$ for all $k = 1, \ldots, m$. There exist vectors $x_0, x_1, \ldots, x_m$ such that $A_j x_k = \alpha_j^M x_k$ and $x_k \in M(k) \setminus M(k)_-$; $k = 1, \ldots, m$. The vectors $x_0, x_1, \ldots x_m$ lie in the m dimensional null space of $\mathbf{A} - \alpha^M$ and are, therefore, linearly dependent. Hence some x_k is a linear combination of $x_0, x_1, \ldots x_{k-1}$. But $x_0, x_1, \ldots x_{k-1} \in M(k-1) \subseteq M(k)_-$. So $x_k \in M(k)_-$. That is a contradiction. So $d \leq m$.

Suppose $m > d$. There are exactly d distinct subspaces $M(1), \ldots, M(d)$, say, such that α^M is the joint diagonal coefficient of $\mathbf{A}$ at $M(k), k = 1, \ldots, d$. Each $M(k)/M(k)_-$ has dimension 1 and therefore there is a continuous linear functional ψ_k on $M(k)$ with kernel $\psi_k^{-1}(0)$ equal to $M(k)_-$. Extend ψ_k to the whole of $\mathcal{X}$. Call the extension φ_k. Then $M(k)_- = \{x \in M(k) : \varphi_k(x) = 0\}$. If $m > d$ there is a non-zero vector x_0 in the m dimensional space $\mathcal{N}_1(\mathbf{A} - \alpha^M)$ satisfying the d linear conditions

$$\varphi_k(x) = 0 \ , \ k = 1, \dots, d.$$

If $M = \cap\{L \in \mathcal{F} : x_0 \in L\}$ then $M \in \mathcal{F}, x_0 \in M \setminus M_-$ and the diagonal coefficient of $\mathbf{A}$ at M is α^M. So $M = M(k)$ for some k and hence $x_0 \in M(k) \setminus M(k)_-$ with $\varphi_k(x) = 0$. That is a contradiction. So $m \leq d$. Hence $m = d$. ∎

The results obtained in this section are summarised in the following theorem.

2.2.15. **Theorem.** Suppose $\mathbf{A} = (A_1, \dots, A_n)$ is an n-tuple of commuting compact linear operators on a complex Banach space X, and $\mathcal{F}$ is a simple chain of closed subspaces of X, each of which is invariant under each $A_j, j = 1, \dots, n$. Then,
(i) a non-zero scalar n-tuple $\boldsymbol{\lambda}$ is a joint eigenvalue of $\mathbf{A}$ if and only if there exists a subspace $M \in \mathcal{F}$ such that $\boldsymbol{\lambda}$ is the joint diagonal coefficient of $\mathbf{A}$ at M,
(ii) the diagonal multiplicity of $\boldsymbol{\lambda}$ is equal to its algebraic multiplicity as a joint eigenvalue of $\mathbf{A}$.

Notes and References. The results of this chapter are motivated by Ringrose's papers [Ri1] and [Ri2] where he obtained these results in the case of a single compact operator. His book [Ri3] gives a complete description of the whole theory. In this connection, see also the works of Brodskiĭ [Br1], [Br2] and [Br3]. Lomonosov's theorem is the culmination of many attempts to generalise the invariant subspace theorem of Aronszajn and Smith, see [ArvF], [Ber], [BerR], [DDP] and [Do]. For various results on invariant subspaces and triangularization see the book [RaRo] by Radjavi and Rosenthal.

Chapter 3

A Spectral Radius Formula

In this chapter we define a new joint spectral radius for an n-tuple of commuting matrices and prove a corresponding spectral radius formula. Our motivation is drawn from Chō and Huruya who proved the following result in [CH]:

Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an n-tuple of commuting matrices. Let $r(\mathbf{A}) = \max\{\|\lambda\| : \lambda \in \sigma_{\mathrm{pt}}(\mathbf{A})\}$. Let $\mathbb{Z}_+^n$ be the set of all multiindices $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\alpha_j \geq 0$ $(j = 1, \ldots, n)$. For such a multiindex, let $|\alpha| = \sum_{j=1}^n \alpha_j$, $\alpha! = \alpha_1! \ldots \alpha_n!$. Also let $\mathbf{A}^{\alpha} = A_1^{\alpha_1} \ldots A_n^{\alpha_n}$ and $\mathbf{A}^* = (A_1^*, \ldots, A_n^*)$. Then

$$r(\mathbf{A}) = \inf_m \left\| \sum_{\alpha \in \mathbb{Z}_+^n, |\alpha| = m} \frac{m!}{\alpha!} A^{*\alpha} A^{\alpha} \right\|^{1/(2m)}.$$

We generalize the above result to the setting of a finite-dimensional Banach space. The main result of this chapter is stated and proved in Section 3.1. The result of Chō and Huruya was extended to infinite dimensional Hilbert spaces by Müller and Soltysiak in [MüS]. Subsequently, our result has been generalized to arbitrary Banach spaces by Müller. See [Mü].

Section 3.2 is devoted to infinite dimensional Hilbert spaces. Here we obtain a simpler arrangement of the proof in [MüS]. We also derive an analogue of a theorem of Rota.

3.1 The p-spectral Radius

Let V_p, $1 \le p \le \infty$ be the d-dimensional complex vector space $\mathbb{C}^d$ equipped with the p-norm :

$$||x||_p = (\sum_{i=1}^{d} |x_i|^p)^{1/p} \quad , \quad x \in \mathbb{C}^d.$$

Let $\mathbf{A} = (A_1, \ldots, A_n)$ be a commuting n-tuple of $d \times d$ matrices. So there exists a unitary matrix U such that $U^* A_j U$ is upper-triangular for all $1 \le j \le n$, i.e.,

$$U^* A_j U = \begin{pmatrix} \lambda_1^{(j)} & & & * \\ & \lambda_2^{(j)} & & \\ & & \ddots & \\ 0 & & & \lambda_n^{(j)} \end{pmatrix}.$$

The joint point spectrum of $\mathbf{A}$ is then given by

$$\sigma_{\mathrm{pt}}(\mathbf{A}) = \{(\lambda_i^{(1)}, \ldots, \lambda_i^{(n)}) : i = 1, \ldots, d\}. \tag{3.1.1}$$

Note that by Theorem 1.3.7, the Taylor joint spectrum and the Harte joint spectrum as well as the left spectrum and the right spectrum of $\mathbf{A}$ are the same as $\sigma_{\mathrm{pt}}(\mathbf{A})$. Let $|\boldsymbol{\lambda}|_p$ denote the p-norm of a vector $\boldsymbol{\lambda}$ in $\mathbb{C}^n$. We define the *geometric spectral radius* of $\mathbf{A}$ as

$$r_p(\mathbf{A}) = \max\{|\boldsymbol{\lambda}|_p : \boldsymbol{\lambda} \in \sigma_{\mathrm{pt}}(\mathbf{A})\}. \tag{3.1.2}$$

The n-tuple $\mathbf{A}$ can be identified with an operator from V_p to the space V_p^n, the direct sum of n copies of V_p equipped with the natural p-norm. The norm of this operator is given by

$$||\mathbf{A}||_p = \sup_{||x||_p = 1} (\sum_{j=1}^{n} ||A_j x||_p^p)^{1/p}. \tag{3.1.3}$$

For $m \ge 2$ we defined the tuple $\mathbf{A}^m$ in Section 2.1. We recall that $\mathbf{A}^m$ is an n^m-tuple of matrices consisting of the products $A_{i_1} \ldots A_{i_m}$ where the indices are chosen from $\{1, \ldots, n\}$ with repetitions allowed, and are then arranged lexicographically. The *algebraic spectral radius* of the n-tuple $\mathbf{A}$ is defined as

$$\rho_p(\mathbf{A}) = \inf_m ||\mathbf{A}^m||_p^{1/m}, \qquad 1 \le p \le \infty. \tag{3.1.4}$$

One of the basic theorems in matrix theory is the *spectral radius formula* which asserts that for a (single) matrix T the infimum above is actually a limit, is independent of the norm $||\cdot||_p$, and is equal to the geometric spectral radius $r(T)$. See, e.g., [HorJ, page 299]. The result of this section is an analogue for the joint spectral radius :

3.1.5. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an n-tuple of commuting $d \times d$ matrices. Then

$$r_p(\mathbf{A}) = \rho_p(\mathbf{A}), \qquad 1 \le p \le \infty.$$

3.1.6. **Remark.** Chō and Huruya's theorem is a special case of Theorem 3.1.5 for $p = 2$. This is so because the components of $\mathbf{A}^m$ are $\mathbf{A}^{\alpha}$ where α varies over $\mathbb{Z}_+^n$ with $|\alpha| = m$. Thus the expression on the right hand side of Chō and Huruya's result is

$$\inf_m ||\mathbf{A}^{m*}\mathbf{A}^m||^{1/(2m)} = \inf_m ||\mathbf{A}^m||_2^{1/m} = \rho_2(\mathbf{A}).$$

And the left hand side is by definition $r_2(\mathbf{A})$.

The proof of the theorem is given below. One of the basic ideas of our proof lies in the introduction of a new operator $\widehat{\mathbf{A}}$ corresponding to any n-tuple $\mathbf{A}$. This is an operator on V_p^∞, the Banach space of all sequences $x = (x_1, x_2, \ldots)$ with $x_j \in V_p$ and $\sum_{j=1}^\infty ||x_j||_p^p < \infty$, equipped with its natural norm $||x||_p = (\sum_{j=1}^\infty ||x_j||_p^p)^{1/p}$. The operator $\widehat{\mathbf{A}}$ is defined as

$$\widehat{\mathbf{A}} = \begin{pmatrix} A_1 & 0 & \cdots & \cdots \\ A_2 & 0 & \cdots & \cdots \\ \vdots & \vdots & & \\ A_n & 0 & \cdots & \cdots \\ 0 & A_1 & \cdots & \cdots \\ 0 & A_2 & \cdots & \cdots \\ \vdots & \vdots & & \\ 0 & A_n & \cdots & \cdots \\ \vdots & \vdots & & \end{pmatrix}. \tag{3.1.7}$$

i.e., $\widehat{\mathbf{A}}$ is an infinite matrix each of whose columns contains one copy of $\mathbf{A}$ according to the rule

$$\widehat{\mathbf{A}}_{jk} = A_{j \ (\mathrm{mod}\ n)} , \quad (k-1)n + 1 \leq j \leq kn , \quad k = 1, 2, \ldots$$

3.1.8. **Lemma.** Let $\mathbf{A}$ be an n-tuple of commuting matrices and let $\widehat{\mathbf{A}}$ be the operator defined in (3.1.7). Then

(*i*) $||\mathbf{A}||_p = ||\widehat{\mathbf{A}}||_p$

(*ii*) $\widehat{\mathbf{A}^m} = (\widehat{\mathbf{A}})^m$

(*iii*) $||\mathbf{A}^m||_p \leq ||\mathbf{A}||_p^m$

(*iv*) $\rho_p(\mathbf{A})$ is the (ordinary) spectral radius of $\widehat{\mathbf{A}}$.

Proof. Let $x = (x_1, x_2, \ldots)$ be an element of V_p^∞. Then

$$\widehat{\mathbf{A}}x = (A_1x_1, \ldots, A_nx_1, A_1x_2, \ldots, A_nx_2, A_1x_3, \ldots) = (\mathbf{A}x_1, \mathbf{A}x_2, \ldots),$$

so that $||\widehat{\mathbf{A}}x||_p^p = \sum_{i=1}^\infty ||\mathbf{A}x_i||_p^p \leq ||\mathbf{A}||_p^p \sum_{i=1}^\infty ||x_i||_p^p = ||\mathbf{A}||_p^p ||x||_p^p$. Hence $||\widehat{\mathbf{A}}||_p \leq ||\mathbf{A}||_p$. On the other hand for $x \in V_p$, $||\mathbf{A}x||_p^p = \sum_{j=1}^n ||A_jx||_p^p = ||\widehat{\mathbf{A}}(x, 0, 0, \ldots)||_p^p \leq ||\widehat{\mathbf{A}}||_p^p \, ||x||_p^p$. So $||\mathbf{A}||_p \leq ||\widehat{\mathbf{A}}||_p$.

This proves (*i*).

The statement (*ii*) is an obvious consequence of the definition of $\mathbf{A}^m$. The statement (*iii*) follows from (*i*), (*ii*) and the fact that any operator norm is submultiplicative. To prove (*iv*) note that

$$\begin{aligned}\rho_p(\mathbf{A}) &= \inf\ \|\mathbf{A}^m\|_p^{1/m} \\ &= \inf\ \|\widehat{\mathbf{A}^m}\|_p^{1/m} \quad \text{by}(i) \\ &= \inf\ \|(\widehat{\mathbf{A}})^m\|_p^{1/m} \quad \text{by}(ii),\end{aligned}$$

and this is the (ordinary) spectral radius of $\widehat{\mathbf{A}}$. ∎

Let S be an invertible matrix. The tuple $S\mathbf{A}S^{-1}$ is defined as

(3.1.9) $$S\mathbf{A}S^{-1} = (SA_1S^{-1}, \ldots, SA_nS^{-1}).$$

3.1.10. **Lemma.** We have

$$\|S\mathbf{A}S^{-1}\|_p \leq \|S\|_p\ \|\mathbf{A}\|_p\ \|S^{-1}\|_p.$$

PROOF Let $\mathbf{R} = S\mathbf{A}S^{-1}$. Let $\mathrm{diag}(S, S, \ldots)$ be the infinite block-diagonal matrix with the diagonal blocks all equal to S. The operator $\widehat{\mathbf{R}}$ on V_p^∞ is then the same as the operator $\mathrm{diag}(S, S, \ldots)\widehat{\mathbf{A}}\mathrm{diag}(S^{-1}, S^{-1}, \ldots)$. So

$$\begin{aligned}\|\widehat{\mathbf{R}}\|_p &\leq \|\mathrm{diag}(S, S, \ldots)\|_p\|\widehat{\mathbf{A}}\|_p\|\mathrm{diag}(S^{-1}, S^{-1}, \ldots)\|_p \\ &= \|S\|_p\|\widehat{\mathbf{A}}\|_p\|S^{-1}\|_p.\end{aligned}$$

Now use Part (*i*) of Lemma 1. ∎

For $\boldsymbol{\lambda}$ in $\mathbb{C}^n$ define $\boldsymbol{\lambda}^m$ in the same way as $\mathbf{A}^m$.

3.1.11. **Lemma.** We have

$$\sigma_{\mathrm{pt}}(\mathbf{A}^m) = \{\boldsymbol{\lambda}^m : \boldsymbol{\lambda} \in \sigma_{\mathrm{pt}}(\mathbf{A})\}.$$

PROOF If $\boldsymbol{\lambda}$ is a joint eigenvalue of $\mathbf{A}$ with a joint eigenvector x then any product of the form $\lambda_{i_1} \ldots \lambda_{i_m}$ is an eigenvalue of $A_{i_1} \ldots A_{i_m}$ with the same eigenvector x. So $\boldsymbol{\lambda}^m$ is a joint eigenvalue of $\mathbf{A}^m$. ∎

3.1.12. **Lemma.** For any commuting n-tuple $\mathbf{A}$of matrices,

$$r_p(\mathbf{A}) \leq \|\mathbf{A}\|_p.$$

PROOF Let $\lambda \in \sigma_{\rm pt}(\mathbf{A})$, and let $x \in V_p$ be such that $A_j x = \lambda_j x$ for all $j = 1, \ldots, n$. Then

$$|\lambda_j|\ ||x||_p = ||A_j x||_p \text{ for all } j = 1, \ldots, n.$$

Hence

$$\sum_{j=1}^{n} |\lambda_j|^p = \frac{1}{||x||_p^p} \sum_{j=1}^{n} ||A_j x||_p^p \leq ||\mathbf{A}||_p^p.$$

This shows $r_p(\mathbf{A}) \leq ||\mathbf{A}||_p$. ∎

3.1.13. **Lemma.** Let $\mathbf{A}$ be a commuting n-tuple of matrices. Then

$$r_p(\mathbf{A}) \leq \rho_p(\mathbf{A}).$$

PROOF Applying Lemma 3.1.12. to the tuple $\mathbf{A}^m$ we get

$$r_p(\mathbf{A}^m) \leq ||\mathbf{A}^m||_p .$$

But $r_p(\mathbf{A}^m) = (r_p(\mathbf{A}))^m$ by Lemma 3.1.11. . Hence

$$r_p(\mathbf{A}) \leq ||\mathbf{A}^m||_p^{1/m} \quad \text{for all } m = 1, 2, \ldots$$

So $r_p(\mathbf{A}) \leq \rho_p(\mathbf{A})$. ∎

We have noted before that for a commuting n-tuple $\mathbf{A}$ there exists a unitary matrix U such that $U^* A_j U$ is upper-triangular for all $j = 1, \ldots, n$. We denote the diagonal part of $U^* A_j U$ by D_j and the strictly upper-triangular part by N_j. Then $D_j = \operatorname{diag}(\lambda_1^{(j)}, \ldots, \lambda_d^{(j)})$. Let $\mathbf{D} = (D_1, \ldots, D_n)$ and $\mathbf{N} = (N_1, \ldots, N_n)$.

3.1.14. **Lemma.** For the n-tuple $\mathbf{D}$ the geometric and algebraic spectral radii are equal i.e.,

$$\rho_p(\mathbf{D}) = r_p(\mathbf{D}) = r_p(\mathbf{A}).$$

PROOF Let $x = (x_1, \ldots, x_d) \in V_p$. Then $D_j x = (\lambda_1^{(j)} x_1, \ldots, \lambda_d^{(j)} x_d)$. The norm of $\mathbf{D}$ as an operator from V_p to V_p^n is given by

$$||\mathbf{D}||_p = \sup_{||x||_p = 1} \left(\sum_{j=1}^{n} ||D_j x||_p^p\right)^{1/p}$$

$$
\begin{aligned}
&= \sup_{\|x\|_p=1} \Big(\sum_{j=1}^{n}\sum_{i=1}^{d} |\lambda_i^{(j)} x_i|^p\Big)^{1/p} \\
&= \sup_{\|x\|_p=1} \Big(\sum_{i=1}^{d}\sum_{j=1}^{n} |\lambda_i^{(j)}|^p |x_i|^p\Big)^{1/p} \\
&\le r_p(\mathbf{D}) \sup_{\|x\|_p=1} \Big(\sum_{i=1}^{d} |x_i|^p\Big)^{1/p} \\
&= r_p(\mathbf{D}). \qquad (3.1.15)
\end{aligned}
$$

So $\rho_p(\mathbf{D}) = \inf \|\mathbf{D}^m\|_p^{1/m} \le \|\mathbf{D}\|_p \le r_p(\mathbf{D})$. By Lemma 3.1.13, $\rho_p(\mathbf{D}) = r_p(\mathbf{D})$. ■

For any real t, let C_t be the $d \times d$ diagonal matrix with entries $t, t^2, \ldots, t^d$.

3.1.16. **Lemma.** For any $\epsilon > 0$, there exists t such that

$$\|C_t U^* \mathbf{A} U C_t^{-1}\|_p < r_p(\mathbf{A}) + \epsilon$$

PROOF Let A be any $d \times d$ matrix. Then

$$
C_t A C_t^{-1} = \begin{pmatrix} a_{11} & t^{-1}a_{12} & \cdots & t^{-d+1}a_{1d} \\ t a_{21} & a_{22} & \cdots & t^{-d+2}a_{2d} \\ \vdots & \vdots & \cdots & \vdots \\ t^{d-1}a_{d1} & t^{d-2}a_{d2} & \cdots & a_{dd} \end{pmatrix}. \qquad (3.1.17)
$$

If A is strictly upper-triangular then for large t we can make the p-norm of $C_t A C_t^{-1}$ as small as we want. We apply this fact to N_j (the strictly upper-triangular part of $U^* A_j U$) for all $j = 1, \ldots, n$. We choose t large enough so that $\|C_t N_j C_t^{-1}\|_p < \epsilon/n$ for all $j = 1, \ldots, n$. Then

$$
\begin{aligned}
\|C_t U^* \mathbf{A} U C_t^{-1}\|_p &= \|C_t (\mathbf{D} + \mathbf{N}) C_t^{-1}\|_p \\
&\le \|C_t \mathbf{D} C_t^{-1}\|_p + \|C_t \mathbf{N} C_t^{-1}\|_p \\
&= \|\mathbf{D}\|_p + \|C_t \mathbf{N} C_t^{-1}\|_p \\
&< r_p(\mathbf{A}) + \epsilon
\end{aligned}
$$

■

The next lemma is the final step in the proof of the theorem.

3.1.18. **Lemma.** For a commuting n-tuple $\mathbf{A}$, if $r_p(\mathbf{A}) < 1$ then $\|\mathbf{A}^m\|_p \to 0$ as $m \to \infty$.

PROOF If $r_p(\mathbf{A}) < 1$ then by Lemma 3.1.17, there exists t such that $||C_tU^*\mathbf{A}UC_t^{-1}||_p < 1$. We have

$$\begin{aligned}
||\mathbf{A}^m||_p &= ||(UC_t^{-1})(C_tU^*\mathbf{A}^mUC_t^{-1})(UC_t^{-1})^{-1})||_p \\
&= ||(UC_t^{-1})(C_tU^*\mathbf{A}UC_t^{-1})^m(UC_t^{-1})^{-1})||_p \\
&\leq ||UC_t^{-1}||_p||(C_tU^*\mathbf{A}UC_t^{-1})^m||_p||(UC_t^{-1})^{-1}||_p \text{ by Lemma 3.1.10} \\
&\leq ||UC_t^{-1}||_p||(C_tU^*\mathbf{A}UC_t^{-1})||_p^m||(UC_t^{-1})^{-1}||_p \text{ by Lemma 3.1.8}(iii).
\end{aligned}$$

Now as $m \to \infty$ the middle term tends to zero because $||C_tU^*\mathbf{A}UC_t^{-1}||_p < 1$. ■

PROOF OF THEOREM 3.1.5 To complete the proof of the theorem define for any $\epsilon > 0$ a new n-tuple $\mathbf{S} = \frac{1}{r_p(\mathbf{A})+\epsilon}\mathbf{A}$. Then $||\mathbf{S}^m||_p = (\frac{1}{r_p(\mathbf{A})+\epsilon})^m||\mathbf{A}^m||_p$. Since $r_p(\mathbf{S}) < 1$ Lemma 3.1.18 says that $||\mathbf{S}^m||_p \to 0$. So for sufficiently large m, $||\mathbf{A}^m||_p < (r_p(\mathbf{A}) + \epsilon)^m$. Hence

$$\rho_p(\mathbf{A}) \leq r_p(\mathbf{A}).$$

In view of Lemma 3.1.13, this proves the theorem. ■

3.1.19. **Remark.** Let Σ be any bounded set of matrices and let Σ^m be the set consisting of products of matrices from Σ of length m. Let $||\cdot||$ be any operator norm on the space $\mathbb{C}^d$. The *Rota-Strang joint spectral radius* (see [RoS]) of Σ is defined as $\nu(\Sigma) = \limsup_m \nu_m(\Sigma)$ where $\nu_m(\Sigma) = \sup\{||A|| : A \in \Sigma_m\}$. In two recent papers, [BeW] and [E4], it has been shown that $\nu(\Sigma)$ is equal to the *generalized spectral radius* $\tau(\Sigma)$ (introduced in [DL]) defined by $\tau(\Sigma) = \limsup_m \tau_m(\Sigma)$ where $\tau_m(\Sigma) = \sup\{r(A) : A \in \Sigma_m\}$. If Σ is taken to be the set $\{A_1, \ldots, A_n\}$ and if the ∞ - norm is used then it is easy to see that $\tau_m(\Sigma) = (r_\infty(\mathbf{A}))^m$ and $\nu_m(\Sigma) = ||\mathbf{A}^m||_\infty$. Since $\nu(\Sigma)$ and $\tau(\Sigma)$ are equal, one gets $r_\infty(\mathbf{A}) = \rho_\infty(\mathbf{A})$. Hence this gives another proof of Theorem 3.1.5 for the special case $p = \infty$.

3.2 The Formula in Hilbert Spaces

Let $\mathcal{H}$ be a separable Hilbert space and let $\mathcal{L}(\mathcal{H})$ be the space of all bounded operators on $\mathcal{H}$. Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an n-tuple of commuting elements of $\mathcal{L}(\mathcal{H})$. As before, we consider $\mathbf{A}$ to be a bounded operator from $\mathcal{H}$ to $\mathcal{H}^n$, the direct sum of n copies of $\mathcal{H}$. The *geometric spectral radius* is then defined to be

$$r(\mathbf{A}) = \sup\{|\lambda|_2 : \lambda \in \sigma_T(\mathbf{A})\}. \tag{3.2.1}$$

Recall that the *joint approximate point spectrum* of $\mathbf{A}$ denoted by $\sigma_{\rm app}(\mathbf{A})$ is defined to be the set

$$\sigma_{\rm app}(\mathbf{A}) \overset{\rm def}{=} \{\lambda \in \mathbf{C}^n : \mathbf{A} - \lambda \text{ is not bounded below}\}$$

It is easy to see that

$$\begin{aligned}\sigma_{\rm app}(\mathbf{A}) = \{\lambda \in \mathbf{C}^n :\ & \text{There exists a sequence of unit vectors} \\ & x_m \text{ satisfying } (A_j - \lambda_j)x_m \to 0 \text{ as } m \to \infty \text{ for all } j = 1, \ldots, n\}.\end{aligned}$$

In [CZ], Chō and Żelazko have shown, in the more general context of a Banach space, that $r(\mathbf{A})$ does not change if in (3.2.1) σ_T is replaced by any other joint spectrum having the polynomial spectral mapping property. In particular,

$$r(\mathbf{A}) = \sup\{|\lambda|_2 : \lambda \in \sigma_{\rm app}(\mathbf{A})\}. \tag{3.2.2}$$

The *algebraic spectral radius* is defined as

$$\rho(\mathbf{A}) = \inf ||\mathbf{A}^m||^{1/m}. \tag{3.2.3}$$

where the norm on the right hand side is the operator norm of the tuple considered as an operator from $\mathcal{H}$ to $\mathcal{H}^n$ equipped with the natural norm:

$$||(x_1, \ldots, x_n)|| = (\sum_{i=1}^{n} ||x_i||^2)^{1/2} \ , \ x_1, \ldots, x_n \in \mathcal{H}.$$

The operator $\widehat{\mathbf{A}}$ defined as in (3.1.7) is now an operator on $\mathcal{H}^\infty$, equipped with the norm

$$||(x_1, x_2, \ldots)|| = (\sum_{i=1}^{\infty} ||x_i||^2)^{1/2} \ , \ x_1, x_2, \ldots \in \mathcal{H}.$$

for which all the facts proved in Lemma 3.1.8 remain true. So the infimum in (3.2 .3) is actually a limit.

Let $\mathbf{M_A} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be the operator defined as

$$\mathbf{M_A}(T) = \sum_{j=1}^{n} A_j^* T A_j. \tag{3.2.4}$$

We then have

3.2.5. **Theorem (Müller and Soltysiak).** Let $\mathbf{A}$ be a commuting n-tuple of Hilbert space operators. Then

$$r(\mathbf{A}) = r(\mathbf{M_A})^{1/2} = \rho(\mathbf{A}) = \lim_{m\to\infty} ||\mathbf{A}^m||^{1/m}. \tag{3.2.6}$$

Our next two remarks are directed towards a simplification of the proof in [MüS] even while following their essential ideas.

By a theorem of Curto [Cu1], the (ordinary) spectrum of $\mathbf{M_A}$ and the joint spetrum of $\mathbf{A}$ are related by

$$\sigma(\mathbf{M_A}) = \{\sum_{j=1}^{n} \bar{\lambda}_j \mu_j : \lambda, \mu \in \sigma_T\}. \tag{3.2.7}$$

Using this and the Cauchy-Schwarz inequality one sees that

$$r(\mathbf{M_A}) \leq r(\mathbf{A})^2. \tag{3.2.8}$$

From the definition of $r(\mathbf{A})$ it is also clear that $r(\mathbf{A})^2$ is a point in $\sigma(\mathbf{M_A})$. Hence

$$r(\mathbf{A})^2 \leq r(\mathbf{M_A}). \tag{3.2.9}$$

This proves the first equality in (3.2.6). Next note that the operator $\mathbf{M_A}$ is a completely positive map on $\mathcal{L}(\mathcal{H})$ (see [P]). Such maps attain their norm at the identity operator I. Applying this to all powers of $\mathbf{M_A}$ we see that

$$||\mathbf{M_A}^m|| = ||\mathbf{M_A}^m(I)|| = ||\mathbf{A}^m||^2. \tag{3.2.10}$$

So the ordinary spectral radius formula for $\mathbf{M_A}$ gives the second equality in (3.2 .6). ■

It is clear from 3.2.2 that

$$r(\mathbf{A}) \leq \|\mathbf{A}\|. \tag{3.2.11}$$

Let $\mathcal{L}_{\text{inv}}(\mathcal{H})$ and $\mathcal{L}_+(\mathcal{H})$ denote, respectively, the set of all invertible operators and the set of all positive definite operators. For any $S \in \mathcal{L}_{\text{inv}}(\mathcal{H})$ we have $\sigma(S\mathbf{A}S^{-1}) = \sigma(\mathbf{A})$ where $S\mathbf{A}S^{-1}$ is the tuple defined in (3.1.9). So it follows from (3.2.11) that

$$r(\mathbf{A}) \leq \inf\{\|S\mathbf{A}S^{-1}\| : S \in \mathcal{L}_{\text{inv}}(\mathcal{H})\}. \tag{3.2.12}$$

We can prove more : there is equality here; for a single operator this was proved by Rota [Ro].

3.2.13. **Theorem.** Let $\mathbf{A}$ be a commuting n-tuple of Hilbert space operators. Then

$$\begin{aligned} r(\mathbf{A}) &= \inf\{\|S\mathbf{A}S^{-1}\| : S \in \mathcal{L}_{\text{inv}}(\mathcal{H})\} \\ &= \inf\{\|S\mathbf{A}S^{-1}\| : S \in \mathcal{L}_+(\mathcal{H})\}. \end{aligned} \tag{3.2.14}$$

PROOF The proof for the case of a single operator T in [FN] can be modified for the present situation. To prove the theorem, we need to produce for each $\eta > r(\mathbf{A})$ an $S \in \mathcal{L}_+(\mathcal{H})$ such that $\|S\mathbf{A}S^{-1}\| < \eta$. By Theorem 3.2.5, given such an η we can find a positive integer m such that $\|\mathbf{A}^m\| < \eta^m$. This means that the operator

$$\begin{aligned} R &= \sum_{m=0}^{\infty} \eta^{-2m}(\mathbf{A}^*)^m \mathbf{A}^m \\ &= I + \frac{1}{\eta^2}\sum_j A_j^* A_j + \frac{1}{\eta^4}\sum_{i,j} A_i^* A_j^* A_j A_i + \cdots \end{aligned}$$

is well-defined. (Here in all summations all subscript indices vary over 1,2,...,n.) Note that $R \geq I$. Further

$$\begin{aligned} \sum A_j^* R A_j &= \sum A_j^* A_j + \frac{1}{\eta^2}\sum A_i^* A_j^* A_j A_i + \cdots \\ &= \eta^2(R - I) \leq \eta^2 R. \end{aligned}$$

Put $S = R^{1/2}$. Then

$$\begin{aligned} \|S\mathbf{A}S^{-1}\|^2 &= \|(S^{-1}\mathbf{A}^*S)(S\mathbf{A}S^{-1})\| \\ &= \|\sum S^{-1}A_j^*S^2A_jS^{-1}\| \\ &= \|S^{-1}(\sum A_j^*RA_j)S^{-1}\| \\ &\leq \eta^2\|S^{-1}RS^{-1}\| = \eta^2. \end{aligned}$$

■

3.2.15. **Remark.** We note that in the finite-dimensional case we have, from the discussion in Section 3.1, a version of Theorem 3.2.13 for all p-norms, $1 \leq p \leq \infty$. More precisely, we have for any commuting tuple $\mathbf{A} = (A_1, \ldots, A_n)$ of matrices

(3.2.16) $$r_p(\mathbf{A}) = \inf \|S\mathbf{A}S^{-1}\|_p$$

where the infimum is taken over all invertible matrices S.

Notes and References. Let $F(k,n)$ be the set of all functions from the set $\{1,\ldots,k\}$ to the set $\{1,\ldots,n\}$. For $f \in F(k,n)$ define $A_f = A_{f(1)}\ldots A_{f(k)}$. Then Bunce conjectured that the radius of the joint approximate point spectrum should be equal to $\inf\{\|\sum_{n\in F(k,n)} A_f^*A_f\|^{1/2k}\}$. (See [Bu2].) Note that this expression is the same as $\inf_m \left\|\sum_{\boldsymbol{\alpha}\in\mathbb{Z}_+^n, |\boldsymbol{\alpha}|=m} \frac{m!}{\boldsymbol{\alpha}!}A^{*\boldsymbol{\alpha}}A^{\boldsymbol{\alpha}}\right\|^{1/(2m)}$. In finite dimension, where joint approximate point spectrum, joint point spectrum and Taylor joint spectrum all are same, the conjecture was proved by Chō and Huruya in [CH]. Chō and Żelazko in [CZ] showed that $\sigma_{\mathrm{app}}(\mathbf{A})$ and σ_{T} have the same spectral radius so that Bunce's conjecture took the form $\sup\{\|\boldsymbol{\lambda}\| : \boldsymbol{\lambda} \in \sigma_{\mathrm{T}}(\mathbf{A})\} = \inf_m \left\|\sum_{\boldsymbol{\alpha}\in\mathbb{Z}_+^n, |\boldsymbol{\alpha}|=m} \frac{m!}{\boldsymbol{\alpha}!}A^{*\boldsymbol{\alpha}}A^{\boldsymbol{\alpha}}\right\|^{1/(2m)}$. Müller and Soltysiak proved the conjecture in [MüS]. The operator $M_{\mathbf{A}}$ used in the proof is also used for obataining Sz.-Nagy-Foias type dilation theory for several commuting Hilbert space operators. (See [MüV].)

Chapter 4

Perturbation Inequalities

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

For commuting tuples of operators, though the concept of joint spectrum has been developed over the last quarter of a century, not many perturbation inequalities seem to be known in this case. Davis [Da] drew special attention to this problem and its importance. After that McIntosh and Pryde [McP] 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 [BDaM]. This approach was developed further by them and Ricker [McPR].

In two recent papers [P1, P2] Pryde has initiated an interesting program: using the ideas of Clifford analysis to generalise some classical perturbation inequalities for single matrices to the case of joint spectra of commuting tuples of matrices. In [P1] he generalizes the classical Bauer-Fike Theorem for single matrices to commuting tuples. In the first section of this chapter, we obtain a similar extension of a well-known theorem of Henrici [Hen]. We follow the ideas of Pryde [P1]. We must emphasize that attempts to obtain similar generalizations of other inequalities [P2] 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 chapter is of interest in this context.

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., [Min]).

4.1 The Classical Henrici Theorem

To state the classical Henrici Theorem we need to define a *measure of non-normality* of a $d \times d$ 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 Λ is a diagonal matrix and N a strictly upper triangular matrix we have

$$U^*AU = T = \Lambda + N \tag{4.1.1}$$

Of course, neither U nor T are uniquely determined. The matrix Λ has as its diagonal entries the eigenvalues of A. The matrix A is normal iff the part N in any decomposition (4.1.1) of A is zero. Given a norm ν on matrices the ν *measure of non-normality* can be defined as

$$\Delta_\nu(A) = \inf \nu(N) \tag{4.1.2}$$

where the infimum is taken over all N occuring in decomposition (4.1.1) of A. A is normal iff $\Delta_\nu(A) = 0$.

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

$$||A|| = \sup_{||x||=1} ||Ax|| \tag{4.1.3}$$

This norm will be of special interest to us.

We can now state the Henrici theorem [Hen].

4.1.4. **Theorem (Henrici).** Let A be a non-normal matrix and let B be any other matrix, $B \neq A$. Let ν be any norm majorizing the operator norm. Let

$$y = \frac{\Delta_\nu(A)}{\nu(B-A)} \tag{4.1.5}$$

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

$$g + g^2 + \cdots + g^d = y. \tag{4.1.6}$$

Then for each eigenvalue β of B there exists an eigenvalue α of A such that

$$|\alpha - \beta| \leq \frac{y}{g_d(y)}\nu(B-A). \tag{4.1.7}$$

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

$$\delta = \min|\alpha - \beta| \tag{4.1.8}$$

where α varies over all eigenvalues of A. Then

$$\frac{\delta}{1 + \frac{\Delta_\nu(A)}{\delta} + \cdot + \frac{\Delta_\nu^{d-1}(A)}{\delta^{d-1}}} \leq \nu(B-A). \tag{4.1.9}$$

See [StSu, 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 n-tuples of matrices. The formulation of our result requires some facts from Clifford algebras which were explained in Section 1.5.

4.2 Perturbation of Real Spectra

In this section we consider n-tuples of matrices $\mathbf{A} = (A_1, \ldots, A_n)$ with real eigenvalues only. Our first result in this section concerns the Bauer-Fike theorem[BaF]. 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 $TAT^{-1} = \Lambda = \mathrm{diag}(\alpha_1, \ldots, \alpha_d)$ and B is any arbitrary matrix then $\sigma(B)$ is contained in the union of the balls

$B(\alpha_i, \varepsilon)$ where $\varepsilon = ||A - B||.||T||.||T^{-1}||$. See [B, p.114]. This was generalized to the case of n-tuples of commuting matrices by Pryde in [P1]. However as Stewart and Sun have pointed out in their recent book [StSu, p.177] Bauer and Fike proved a stronger result:

4.2.1. **Theorem (Bauer-Fike).** Let $A, B \in \mathcal{M}_d$ and $T \in \mathrm{GL}(d)$. Then for any $\beta \in \sigma(B) \setminus \sigma(A)$,

$$||T^{-1}(A-\beta)^{-1}T||^{-1} \leq ||T^{-1}(A-B)T||. \tag{4.2.2}$$

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

4.2.3. **Theorem.** Let $\mathbf{A} = (A_1, \dots, A_n)$ and $\mathbf{B} = (B_1, \dots, B_n)$ be two n-tuples of commuting matrices with real spectra. Let $\beta \in \sigma_{\mathrm{pt}}(B) \setminus \sigma_{\mathrm{pt}}(A)$. Let $T \in \mathrm{GL}(d)$. Then

$$||T^{-1}(\mathrm{Cliff}(A-\beta)^{-1}T||^{-1} \leq ||T^{-1}\mathrm{Cliff}(A-B)T|| \tag{4.2.4}$$

PROOF Since $\beta \notin \sigma_{\mathrm{pt}}(\mathbf{A})$ and all A_j have real spectra, $\sum(A_j - \beta_j)^2$ is invertible. So by part (ii) of Lemma 1.5.9, $\mathrm{Cliff}(A-\beta)$ is invertible.

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

$$\begin{aligned}
\mathrm{Cliff}(\mathbf{A}-\mathbf{B})(x \otimes h_\phi) &= \imath \sum (A_j - B_j)(x \otimes h_{\{j\}}) \\
&= \imath \sum (A_j - \beta_j)(x \otimes h_{\{j\}}) \\
&= \mathrm{Cliff}(A-\beta)(x \otimes h_\phi)
\end{aligned}$$

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

Hence

$$T^{-1}(x \otimes h_\phi) = (T^{-1}(\mathrm{Cliff}(\mathbf{A}-\beta))^{-1}T)(T^{-1}(\mathrm{Cliff}(\mathbf{A}-\mathbf{B}))T)(T^{-1}(x \otimes h_\phi))$$

After taking norms and cancelling $||T^{-1}(x \otimes h_\phi)||$ from both sides we have $||(T^{-1}(\mathrm{Cliff}(\mathbf{A}-\beta))^{-1}T)||^{-1} \leq ||(T^{-1}(\mathrm{Cliff}(\mathbf{A}-\mathbf{B}))T)||$ ∎

Now we will define the *measure of non-normality* of an n-tuple $\mathbf{A} = (A_1, \ldots, A_n)$ 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 [HorJ, p.81]. Write $T_j = \Lambda_j + N_j$ where the Λ_j are diagonal and the N_j are strictly upper-triangular. Let $\mathbf{N} = (N_1, \ldots, N_n)$. We can define the *measure of non-normality* of $\mathbf{A}$ as

$$\Delta(\mathbf{A}) = \inf \;||\mathrm{Cliff}(\mathbf{N})|| \tag{4.2.5}$$

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

4.2.6. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{B} = (B_1, \ldots, B_n)$ be two commuting n-tuples of matrices with real spectra . Let $\beta \in \sigma_{\mathrm{pt}}(\mathbf{B}) \setminus \sigma_{\mathrm{pt}}(\mathbf{A})$. Let $\delta = \min \{||\alpha - \beta|| : \alpha \in \sigma_{\mathrm{pt}}(\mathbf{A})\}$. Then

$$\frac{\delta}{1 + \frac{\Delta(A)}{\delta} + \cdot + \frac{\Delta^{d-1}(A)}{\delta^{d-1}}} \leq ||\mathrm{Cliff}(A - B)|| \tag{4.2.7}$$

Proof Let U be a unitary such that the infimum in the definition (4.2.5) of $\Delta(\mathbf{A})$ is attained. Then $U^*A_jU = \Lambda_j + \mathrm{N}_j$ where the Λ_j are diagonal and the N_j strictly upper-triangular. Let $\Lambda = (\Lambda_1, \ldots, \Lambda_n)$ and $\mathrm{N} = (N_1, \ldots, N_m)$. Then

$$\begin{aligned} U^*\mathrm{Cliff}(\mathbf{A}-\beta)U &= \imath U^*(\textstyle\sum(A_j - \beta_j) \otimes h_{\{j\}})U \\ &= \imath \textstyle\sum(\Lambda_j + N_j - \beta_j) \otimes h_{\{j\}} \\ &= \mathrm{Cliff}(\Lambda + N - \beta) \end{aligned} \tag{4.2.8}$$

$$\begin{aligned} \text{But } \mathrm{Cliff}(\Lambda + N - \beta) &= \mathrm{Cliff}(\Lambda - \beta) + \mathrm{Cliff}(\mathrm{N}) \\ &= \mathrm{Cliff}(\Lambda - \beta)(1 + (\mathrm{Cliff}(\Lambda - \beta))^{-1}\mathrm{Cliff}(\mathbf{N})) \end{aligned} \tag{4.2.9}$$

Let $\Lambda_j - \beta_j = D_j$ and let $\mathbf{D} = (D_1, \ldots, D_n)$. Note that $\beta \notin \sigma_{\mathrm{pt}}(\Lambda)$ since $\beta \notin \sigma_{\mathrm{pt}}(\mathbf{A})$. So $0 \notin \sigma_{\mathrm{pt}}(\mathbf{D})$. Moreover note that $\sigma_{\mathrm{pt}}(\mathbf{D}) \subset \mathbb{R}^n$. Hence by Proposition 1.4.2,

$$\sigma_{\mathrm{pt}}(\mathbf{D}) = \{\boldsymbol{\lambda} \in \mathbb{R}^n : 0 \in \sigma\left(\textstyle\sum(D_j - \lambda_j)^2\right)\}.$$

So $\sum D_j^2$ is invertible. Since $\mathrm{Cliff}(\mathrm{D})^2 = \sum(D_j)^2$, hence Cliff(D) is invertible and $(\mathrm{Cliff}(\mathbf{D}))^{-1} = \sum_k T_k \otimes h_{\{k\}}$ where $T_k = (\sum_i D_i^2)^{-1} \ D_k$. The T_k are all diagonal. So

$$(4.2.10)\qquad \begin{aligned} (\mathrm{Cliff}(A-\beta))^{-1}\mathrm{Cliff}(\mathrm{N}) &= (\sum_k T_k \otimes h_{\{k\}})(\sum_j N_j \otimes h_{\{j\}}) \\ &= \sum_{k,j} T_k N_j \otimes h_{\{k\}} h_{\{j\}} = S, \text{ say.} \end{aligned}$$

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 S various products of $T_k N_j$ appear. But any product of d strictly upper-triangular matrices is zero. So S^d is zero. Therefore

$$(4.2.11)\qquad (\mathrm{Cliff}(A+N-\beta))^{-1} = (1 - S + \cdots + (-1)^{d-1}S^{d-1})(\mathrm{Cliff}(A-\beta))^{-1}$$

$$\|(\mathrm{Cliff}(A+N-\beta))^{-1}\| \leq \|(\mathrm{Cliff}(A-\beta))^{-1}\| . \big(1 + \|(\mathrm{Cliff}(A-\beta))^{-1}\| . \|\mathrm{Cliff}(\mathrm{N})\| + \cdots + \|(\mathrm{Cliff}(A-\beta))^{-1}\|^{d-1} . \|\mathrm{Cliff}(\mathrm{N})\|^{d-1}\big)$$

Let $\eta = \|(\mathrm{Cliff}(A-\beta))^{-1}\|^{-1}$. Then

$$(4.2.12)\qquad \|(\mathrm{Cliff}(A+\mathrm{N}-\beta))^{-1}\| \leq \eta^{-1}(1 + \frac{\Delta(\mathrm{A})}{\eta} + \cdots + \frac{\Delta^{d-1}(A)}{\eta^{d-1}})$$

Taking $T = U$ in Theorem 4.2.3 we have

$$(4.2.13)\qquad \begin{aligned} \|(\mathrm{Cliff}(A+\mathrm{N}-\beta))^{-1}\|^{-1} &= \|U^*(\mathrm{Cliff}(\mathrm{A}-\beta))^{-1}U\|^{-1} \\ &\leq \|U^*\mathrm{Cliff}(\mathrm{A}-\mathbf{B})U\| \\ &\leq \|\mathrm{Cliff}(\mathrm{A}-\mathbf{B})\| \end{aligned}$$

By (4.2.12) and (4.2.13)

$$(4.2.14)\qquad \|\mathrm{Cliff}(\mathrm{A}-\mathbf{B})\| \geq \frac{\eta}{(1 + \frac{\Delta(\mathrm{A})}{\eta} + \cdots + \frac{\Delta^{d-1}(\mathrm{A})}{\eta^{d-1}})}$$

So the proof will be complete if we show that $\eta = \delta$. Recall that the D_j are diagonal matrices with real entries. Let $a_i^{(j)}$ be the (i,i)th. entry of D_j, i.e.

$D_j = \mathrm{diag}(a_1^{(j)}, \ldots, a_d^{(j)})$. Then $\sigma_{\mathrm{pt}}(\mathbf{D}) = \{(a_i{}^{(1)}, \ldots, a_i{}^{(n)}) : i = 1, \ldots, d\}$. Put $(a_i{}^{(1)}, \ldots, a_i{}^{(n)}) = a_i$. Then

$$\begin{aligned}
\eta &= \|(\mathrm{Cliff}(A - \beta))^{-1}\|^{-1} \\
&= \|(\sum D_j{}^2)^{-1}\mathrm{Cliff}(\mathbf{D})\|^{-1} \\
&= \|\mathrm{Cliff}(\sum D_j{}^2)^{-1}\mathbf{D}\|^{-1} \\
&= \max\{\frac{\|a_i\|}{\|a_i\|^2}, i = 1, \ldots, d\}^{-1} \\
&= \max\{\|a_i\|, i = 1, \ldots, d\} \\
&= \max\{\|z\|, z \in \sigma_{\mathrm{pt}}(A - \beta)\} \\
&= \max\{\|\alpha - \beta\| : \alpha \in \sigma_{\mathrm{pt}}(\mathbf{A})\} \\
&= \delta, \quad \text{by definition.}
\end{aligned}$$

That completes the proof. ■

The results of this section have been generalized later to simultaneously upper-trianglarizable matrices (see [P3]).

4.3 Perturbation of Complex Spectra

In this section we consider n-tuples $\mathbf{A} = (A_1, \ldots, A_n)$ of commuting matrices with no restriction on the spectrum of A_j. We recall that an n-tuple $\mathbf{A}$ of elements from $\mathcal{L}(\mathcal{X})$ is said to be *strongly commuting* if for each $1 \leq j \leq n$ there exist operators A_{1j} and A_{2j}, each with real spectrum, such that $A_j = A_{1j} + iA_{2j}$ and $\Pi(\mathbf{A}) = (A_{11}, \ldots, A_{1n}, A_{21}, \ldots, A_{2n})$ is a commuting $2n$-tuple. $\Pi(\mathbf{A})$ is called a partition of $\mathbf{A}$. It is well-known that a commuting tuple of matrices is strongly commuting (see [McPR]).. Since the A_{qj} commute for all $q = 1,2$ and for all $j = 1, \ldots, n$, there exists a unitary U such that $U^*A_{qj}U = \Lambda_{qj} + N_{qj}$ where Λ_{qj} is diagonal and N_{qj} upper-triangular. Let

$$\mathbf{N} = (N_{11}, \ldots, N_{1n}, N_{21}, \ldots, N_{2n}) \tag{4.3.1}$$

and in this case define the measure of non-normality of $\mathbf{A}$ by

$$\Delta(\mathbf{A}) = \Delta(\pi(\mathbf{A})) = \inf \ ||\mathrm{Cliff}(\mathbf{N})|| \tag{4.3.2}$$

where as before the infimum is taken over all $\mathbf{N}$ associated with $\mathbf{A}$ in the above construction. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{C}^n$ let $\alpha_{1j} = \mathrm{Re}(\alpha_j), \alpha_{2j} = \mathrm{Im}(\alpha_j)$. Now $\mathbf{A}$ is a strongly commuting n-tuple in the sense of Definition 1.4.5. Thus by applying Proposition 1.4.6, we get $\boldsymbol{\alpha} \in \sigma_{\mathrm{pt}}(\mathbf{A})$ if and only if $(\alpha_{11}, \ldots, \alpha_{1n}, \alpha_{21}, \ldots, \alpha_{2n}) \in \sigma_{\mathrm{pt}}(\pi(\mathbf{A}))$. Now let $\mathbf{A}$ and $\mathbf{B}$ be any two n-tuples of commuting matrices. Let $\boldsymbol{\beta} \in \sigma_{\mathrm{pt}}(\mathbf{B})$. Then $(\beta_{11}, \ldots, \beta_{1n}, \beta_{21}, \ldots, \beta_{2n}) \in \sigma_{\mathrm{pt}}(\pi(\mathbf{B}))$. By Theorem 4.2.6, there exists $(\alpha_{11}, \ldots, \alpha_{1n}, \alpha_{21}, \ldots, \alpha_{2n}) \in \sigma_{\mathrm{pt}}(\pi(\mathbf{A}))$ such that if

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

then

$$\frac{\delta}{1 + \frac{\Delta(\mathbf{A})}{\delta} + \cdot + \frac{\Delta^{d-1}(A)}{\delta^{d-1}}} \leq ||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))|| \tag{4.3.3}$$

Note that δ is also equal to $||\boldsymbol{\beta} - \boldsymbol{\alpha}||$, the distance between $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ in $\mathbb{C}^n$. So we have proved:

4.3.4. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{B} = (B_1, \ldots, B_n)$ be two n-tuples of commuting $d{\times}d$ matrices. Let $\boldsymbol{\beta} \in \sigma_{\mathrm{pt}}(\mathbf{B})$. Define $\delta = \min \{||\boldsymbol{\alpha}-\boldsymbol{\beta}|| : \boldsymbol{\alpha} \in \sigma_{\mathrm{pt}}(A)\}$. Then

$$\frac{\delta}{1 + \frac{\Delta(\mathbf{A})}{\delta} + \cdot + \frac{\Delta^{d-1}(\mathbf{A})}{\delta^{d-1}}} \leq ||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||$$

4.4 A Spectral Variation Bound

Consider two $d \times d$ matrices A and B. Let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of A and $\mu_1, \ldots, \mu_d$ the eigenvalues of B. Let $S_A(B) = \max_j \min_i |\lambda_i - \mu_j|$. In [BFr] Bhatia and Friedland proved that

$$S_A(B) \leq d^{1/d}(2M)^{1-1/d}||A - B||^{1/d} \tag{4.4.1}$$

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

The approach in [BFr] was through characteristic polynomials. In [E1] Elsner obtained the same result from Henrici's theorem. Using Elsner's approach we can obtain an analogue of (4.4.1) for commuting n-tuples. Let

$$S_{\mathbf{A}}(\mathbf{B}) = \max_{\mu\in\sigma_{pt}(\mathbf{B})} \min_{\lambda\in\sigma_{pt}(\mathbf{A})} \ ||\lambda-\mu|| \tag{4.4.2}$$

Define

$$S_d(\Delta, r) = \begin{cases} \frac{y}{g_d(y)}r & \text{where } y = \Delta/r \text{ for } r > 0 \\ 0 & \text{for r } = 0 \end{cases} \tag{4.4.3}$$

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

The following lemma can be found in Elsner [E1] and is crucial to the proof.

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

$$\gamma = (\delta^{d-1} + \delta^{d-2}\tau + \ldots + \tau^{d-1})^{1/d}$$

Then γ is the minimal number such that

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

Theorem 4.3.4 can be equivalently stated as

$$S_{\mathbf{A}}(\mathbf{B}) \ \leq \ S_d\big(\Delta(\mathbf{A}), ||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||\big) \tag{4.4.5}$$

Let $M = \max$ (|| Cliff$((\pi(\mathbf{A}))||$, || Cliff$((\pi(\mathbf{B}))||$) . Let U be a unitary such that $U^* A_{q_j} U = N_{q_j} + \Lambda_{q_j}$ and suppose the infimum in the definition (4.3.2) is attained for this choice i.e., $\Delta(\pi(\mathbf{A})) = ||\mathrm{Cliff}(\mathrm{N})||$ where N is as in (4.3.1). Then by definition

$$\Delta(\mathbf{A}) = ||\mathrm{Cliff}(\mathrm{N})|| \leq ||\mathrm{Cliff}(\mathrm{N} + \Lambda)|| + ||\mathrm{Cliff}(\Lambda)|| \tag{4.4.6}$$

The first term on the right hand side is $||\mathrm{Cliff}(\pi(\mathbf{A}))||$. And $||\mathrm{Cliff}(\Lambda)|| = r(\Lambda) = r(\pi(\mathbf{A})) \leq ||\mathrm{Cliff}(\pi(\mathbf{A}))||$. So $\Delta(\mathbf{A}) \leq 2M$. Hence by monotonicity of S_d in the first component

$$S_{\mathbf{A}}(\mathbf{B}) \leq S_d\,(2M, ||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||) \tag{4.4.7}$$

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

$$||\lambda - \mu|| \leq ||\lambda|| + ||\mu|| \leq ||\mathrm{Cliff}(\mathbf{A})|| + ||\mathrm{Cliff}(\mathbf{B})|| \tag{4.4.8}$$

so that $S_{\mathbf{A}}(\mathbf{B}) \leq 2M$. So

$$\begin{aligned} S_{\mathbf{A}}(\mathbf{B}) &\leq \min\{S_d(2M, ||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||), 2M\} && (4.4.9) \\ &\leq d^{1/d}(2M)^{1-1/d}||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||^{1/d} && (4.4.10) \end{aligned}$$

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

4.4.11. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{B} = (B_1, \ldots, B_n)$ be two n-tuples of commuting $d \times d$ matrices. Let $\pi(\mathbf{A})$ and $\pi(\mathbf{B})$ be partitions of $\mathbf{A}$ and $\mathbf{B}$ respectively. Let $M = \max(||\mathrm{Cliff}((\pi(\mathbf{A})||, ||\mathrm{Cliff}((\pi(\mathbf{B})||)$. Let $S_{\mathbf{A}}(\mathbf{B})$ be as defined in (4.4.2). Then we have the following bound on $S_{\mathbf{A}}(\mathbf{B})$.

$$S_{\mathbf{A}}(\mathbf{B}) \leq d^{1/d}(2M)^{1-1/d}||\mathrm{Cliff}(\pi(\mathbf{A}) - \pi(\mathbf{B}))||^{1/d} \tag{4.4.12}$$

4.5 The Hoffman-Wielandt Theorem

For an operator T on $\mathbf{C}^d$ let $||T||_2 = (\mathrm{tr}\, T^*T)^{1/2}$ denote its *Frobenius norm* (also called the *Hilbert-Schmidt norm* or the *Schatten 2-norm*).

In this section we shall prove:

4.5.1. **Theorem.** Let $\mathbf{A} = (A_1, \ldots, A_n)$ be an n-tuple of commuting normal operators on $\mathbf{C}^d$ and let $\boldsymbol{\alpha}_k = (\alpha_k^{(1)}, \ldots, \alpha_k^{(n)}), 1 \leq k \leq d$ be the joint eigenvalues of $\mathbf{A}$. Let $\mathbf{B} = (B_1, \ldots, B_n)$ be another such n-tuple with joint eigenvalues $\boldsymbol{\beta}_k = (\beta_k^{(1)}, \ldots, \beta_k^{(n)}), 1 \leq k \leq d$. Then there exists a permutation σ on d indices such that

$$\sum_{j=1}^{n}\sum_{k=1}^{d} |\alpha_k^{(j)} - \beta_{\sigma(k)}^{(j)}|^2 \leq \sum_{j=1}^{n} ||A_j - B_j||_2^2. \tag{4.5.2}$$

When $n = 1$ the inequality (4.5.2) reduces to a famous inequality of Hoffman and Wielandt [HofW]. The noteworthy feature of (4.5.2) is that the *same* permutation σ does the job for each of the n components.

If we think of an n-tuple $(T_1, \ldots, T_n)$ of (not necessarily commuting) operators on $\mathbf{C}^d$ as a column vector

$$\mathbf{T} = \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{pmatrix}$$

and consider it as an operator from $\mathbf{C}^d$ to $\bigoplus_{n \text{ copies}} \mathbf{C}^d$ then the 2-norm of $\mathbf{T}$ is given by

$$||\mathbf{T}||_2 = (\mathrm{tr}\mathbf{T}^*\mathbf{T})^{1/2} = (\sum_{j=1}^{n} ||T_j||_2^2)^{1/2}.$$

The inequality (4.5.2) can then be written as

$$\sum_{k=1}^{d} ||\alpha_k - \beta_{\sigma(k)}||_{\mathbf{C}^n}^2 \quad \leq \quad ||\mathbf{A} - \mathbf{B}||_2^2 \tag{4.5.3}$$

Our proof of Theorem 4.5.1 has two ingredients. One is the use of Birkhoff's Theorem on doubly stochastic matrices exactly as in the proof by Hoffman and Wielandt. The other is the use of Clifford operators. We start with a lemma concerning the Frobenius norm of a Clifford operator.

4.5.4. **Lemma.** Let $\mathbf{T} = (T_1, \ldots, T_n)$ be any n-tuple of operators in $\mathbf{C}^d$ and let Cliff($\mathbf{T}$) be the associated Clifford operator. Then

$$||\mathrm{Cliff}(\mathbf{T})||_2^2 = 2^n \sum_{j=1}^{n} ||T_j||_2^2.$$

PROOF Note that

$$\begin{aligned} \mathrm{Cliff}(\mathbf{T})^*\mathrm{Cliff}(\mathbf{T}) &= \left[\imath \sum_{j=1}^{n} T_j^* \otimes h_{\{j\}} \right] \left[\imath \sum_{k=1}^{n} T_k \otimes h_{\{k\}} \right] \\ &= -\sum_{j,k} T_j^* T_k \otimes h_{\{j\}} h_{\{k\}} \\ &= \sum_{j=1}^{n} T_j^* T_j \otimes h_{\phi} - \sum_{j \neq k} T_j^* T_k \otimes h_{\{j\}} h_{\{k\}} \end{aligned} \tag{4.5.5}$$

Let $u_l, l = 1, 2, \ldots, d$ be an orthonormal basis for $\mathbb{C}^d$ and $h_S, S \subseteq \{1, 2, \ldots, n\}$ the standard orthonormal basis for $I\!R_{(n)}$. Then the collection $u_l \otimes h_S$ is an orthonormal basis for $\mathbb{C}^d \otimes I\!R_{(n)}$. We have, therefore, for each j

$$\begin{aligned} \operatorname{tr}(T_j^* T_j \otimes h_\phi) &= \sum_l \sum_S < (T_j^* T_j \otimes h_\phi)(u_l \otimes h_S), u_l \otimes h_S > \\ &= \sum_l \sum_S < T_j^* T_j u_l, u_l > \\ &= 2^n \operatorname{tr} T_j^* T_j. \end{aligned} \tag{4.5.6}$$

On the other hand if $j \neq k$ then

$$\begin{aligned} \operatorname{tr}(T_j^* T_k \otimes h_{\{j\}} h_{\{k\}}) &= \sum_l \sum_S < T_j^* T_k u_l \otimes h_{\{j\}} h_{\{k\}} h_S, u_l \otimes h_S > \\ &= 0. \end{aligned} \tag{4.5.7}$$

The Lemma now follows from (4.5.5) , (4.5.6) and (4.5.7). ∎

PROOF OF THE THEOREM Choose orthonormal bases $u_1, \ldots, u_d$ and $v_1, \ldots, v_d$ for $\mathbb{C}^d$ such that

$$A_j u_k = \alpha_k^{(j)} u_k, \quad B_j v_k = \beta_k^{(j)} v_k \tag{4.5.8}$$

for $1 \leq j \leq n$, $1 \leq k \leq d$. If P_k and Q_k denote the orthoprojectors onto the 1-dimensional spaces spanned by u_k and v_k respectively, we can write the spectral resolutions

$$A_j = \sum_{k=1}^{d} \alpha_k^{(j)} P_k \ , \quad B_j = \sum_{k=1}^{d} \beta_k^{(j)} Q_k \tag{4.5.9}$$

Then,

$$\begin{aligned} \operatorname{Cliff}(\mathbf{A}) &= \imath \sum_{j=1}^{n} A_j \otimes h_{\{j\}} \\ &= \imath \sum_{j=1}^{n} \left(\sum_{k=1}^{d} \alpha_k^{(j)} P_k \right) \otimes h_{\{j\}} \\ &= \imath \sum_{k=1}^{d} \left(\sum_{j=1}^{n} \alpha_k^{(j)} I \otimes h_{\{j\}} \right) (P_k \otimes h_\phi) \end{aligned} \tag{4.5.10}$$

In the same way,

$$\operatorname{Cliff}(\mathbf{B}) = \imath \sum_{k=1}^{d} \left(\sum_{j=1}^{n} \beta_k^{(j)} I \otimes h_{\{j\}} \right) (Q_k \otimes h_\phi) \tag{4.5.11}$$

Note that for each k the operator $P_k \otimes h_\phi$ is an orthoprojector in $\mathbb{C}^d \otimes I\!R_{(n)}$ and its range is the 2^n-dimensional subspace $(\mathrm{Ran}\ P_k) \otimes I\!R_{(n)}$. In particular $(P_k \otimes h_\phi)(P_j \otimes h_\phi) = 0$, if $j \neq k$. Note that

$$\mathrm{tr}(P_k \otimes h_\phi) = 2^n \tag{4.5.12}$$

An analogous statement is true for the Q_k.

Now as before we will compute traces with respect to the orthonormal basis $u_r \otimes h_S$ in $\mathbb{C}^d \otimes I\!R_{(n)}$, where $1 \leq r \leq d$ and $S \subseteq \{1, 2, \ldots, n\}$.

If P is any operator on $\mathbb{C}^d$ and T a non-empty subset of $\{1, 2, \ldots, n\}$ then

$$\begin{aligned} \mathrm{tr}\ (P \otimes h_T) &= \sum_{r,S} < (P \otimes h_T)(u_r \otimes h_S), u_r \otimes h_S > \\ &= \sum_{r,S} < Pu_r \otimes h_T h_S, u_r \otimes h_S >= 0 \end{aligned} \tag{4.5.13}$$

by the definition of the inner product and the fact that $h_T h_S$ is never equal to h_S unless T is empty. So,

$$\begin{aligned} \mathrm{tr}\ \mathrm{Cliff}(\mathbf{A})^* \mathrm{Cliff}(\mathbf{B}) &= -\mathrm{tr} \textstyle\sum_{k,l} (\sum_i \overline{\alpha_k^{(i)}} I \otimes h_{\{i\}})(\sum_j \beta_l^{(j)} I \otimes h_{\{j\}})(P_k Q_l \otimes h_\phi) \\ &= -\mathrm{tr} \textstyle\sum_{k,l} \left[-(\sum_i \overline{\alpha_k^{(i)}} \beta_k^{(i)}) P_k Q_l \otimes h_\phi + (\sum_{i \neq j} \overline{\alpha_k^{(i)}} \beta_l^{(j)}) P_k Q_l \otimes h_{\{i\}} h_{\{j\}} \right] \\ &= \sum_{k,l,i} \overline{\alpha_k^{(i)}} \beta_k^{(i)} \mathrm{tr}(P_k Q_l \otimes h_\phi), \text{ by (4.5.13) above.} \end{aligned} \tag{4.5.14}$$

Let

$$d_{kl} = \mathrm{tr}(P_k Q_l \otimes h_\phi), \quad k, l = 1, 2, \ldots, d \tag{4.5.15}$$

Then the $d \times d$ matrix $\mathrm{D} = ((d_{kl}))$ is 2^n times a doubly stochastic matrix, because of the relation (4.5.12) and the corresponding fact for Q_l. Hence by Birkhoff's Theorem there exist constants $a_\sigma \geq 0, \sum_\sigma a_\sigma = 1$ such that

$$D = 2^n \sum_{\sigma \in S_d} a_\sigma \sigma, \tag{4.5.16}$$

where S_d is the group of permutations on d indices and each element of S_d is also identified with a permutation matrix.

The rest of the proof now proceeds as in [B, p.74] for the classical Hoffman-Wielandt theorem. Using (4.5.4), (4.5.14) and (4.5.16) we can write

$$\begin{aligned}
&||\mathrm{Cliff}(\mathbf{A}) - \mathrm{Cliff}(\mathbf{B})||_2^2 \\
&= \mathrm{tr}\ \mathrm{Cliff}(\mathbf{A})^*\mathrm{Cliff}(\mathbf{A}) + \mathrm{tr}\ \mathrm{Cliff}(\mathbf{B})^*\mathrm{Cliff}(\mathbf{B}) - 2\mathrm{Re}\ \mathrm{tr}\ \mathrm{Cliff}(\mathbf{A})^*\mathrm{Cliff}(\mathbf{B}) \\
&= ||\mathrm{Cliff}(\mathbf{A})||_2^2 + ||\mathrm{Cliff}(\mathbf{B})||_2^2 - 2\mathrm{Re}\ \mathrm{tr}\ \mathrm{Cliff}(\mathbf{A})^*\mathrm{Cliff}(\mathbf{B}) \\
&= 2^n \sum_\sigma a_\sigma \sum_{k,j} \left[|\alpha_k^{(j)}|^2 + |\beta_{\sigma(k)}^{(j)}|^2 - 2\mathrm{Re}\ \overline{\alpha_k^{(j)}} \beta_{\sigma(k)}^{(j)} \right] \\
&= 2^n \sum_\sigma a_\sigma \sum_{k,j} |\alpha_k^{(j)} - \beta_{\sigma(k)}^{(j)}|^2 \\
&= 2^n \sum_\sigma a_\sigma \sum_{k=1}^d ||\boldsymbol{\alpha}_k - \boldsymbol{\beta}_{\sigma(k)}||_{\mathbf{C}^n}^2
\end{aligned}$$

so that

$$||\mathrm{Cliff}(\mathbf{A}) - \mathrm{Cliff}(\mathbf{B})||_2^2 \geq 2^n \min_\sigma \sum_{k=1}^d ||\boldsymbol{\alpha}_k - \boldsymbol{\beta}_{\sigma(k)}||_{\mathbf{C}^n}^2. \tag{4.5.17}$$

The theorem now follows from (4.5.4) and (4.5.17). ∎

4.5.18. **Remarks. 1.** Note that by the same argument a reverse inequality can also be proved: there exists another permutation σ such that

$$||\mathbf{A} - \mathbf{B}||_2^2 \leq \sum_{k=1}^d ||\boldsymbol{\alpha}_k - \boldsymbol{\beta}_{\sigma(k)}||_{\mathbf{C}^n}^2. \tag{4.5.19}$$

2. Sun[Su] has proved a generalization of the Hoffman-Wielandt inequality valid for operators similar to diagonal ones. This result can also be extended to commuting tuples. More precisely, let $\mathbf{A} = (A_1, \ldots, A_n)$ and $\mathbf{B} = (B_1, \ldots, B_n)$ be two n-tuples of operators on $\mathbf{C}^d$. Let P, Q be invertible operators such that

$$PA_jP^{-1} = \mathrm{diag}(\alpha_1^{(j)}, \ldots, \alpha_d^{(j)})$$

$$QB_jQ^{-1} = \mathrm{diag}(\beta_1^{(j)}, \ldots, \beta_d^{(j)})$$

are diagonal matrices. Then $\boldsymbol{\alpha}_k, \boldsymbol{\beta}_k$, $1 \leq k \leq d$ are the joint eigenvalues of $\mathbf{A}$ and $\mathbf{B}$ respectively. We define the condition number of P as

$$\mathrm{cond}\ P = ||P||.||P^{-1}||,$$

where $||P||$ denotes the usual operator sup norm of P. It can be proved that under the above circumstances there exists a permutation σ on d indices such that

$$\sum_{k=1}^{d} ||\boldsymbol{\alpha}_k - \boldsymbol{\beta}_{\sigma(k)}||^2_{\mathbb{C}^n} \leq (\text{cond } P)^2 (\text{cond } Q)^2 ||\mathbf{A} - \mathbf{B}||_2^2 \tag{4.5.20}$$

Sun's result [Su] is a special case of this when $n = 1$. This can be proved by imitating Sun's proof [Su] (see also [StSu, p.216]) and using our Theorem 1.

Notes and References. For various important results on the perturbation theory for eigenvalues of a single matrix see [B] and [StSu]. For commuting tuples of matrices the program was initiated by Pryde [P1, [P2]. Subsequently the results of this chapter were published in two papers [BBh1, BBh2]. The proofs presented above were then simplified by Elsner [E2, E3]. Pryde generalized the above results to simultaneously upper-triagularizable but not necessarily commuting matrices [P3].

Bibliography

[Al] Albrecht, E., *On joint spectra*, Studia Math. 64(1979), 263-271.

[Are] Arens, R., *The analytic functional calculus in commutative topological algebras*, Pacific J. Math. 11(1961), 405-429.

[ArCa] Arens, R. and Calderón, A. P., *Analytic functions of several Banach algebra elements*, Annals of Math. 62(1955), 204-216.

[Ar] Artin, E., *Geometric Algebra*, Wiley Interscience, New York, 1957.

[AS] Aronszajn, N. and Smith, K. T., *Invariant subspaces of completely continuous operators*, Ann. Math. 60(1954), 345-350.

[ArvF] Arveson, W. B., and Feldman, J., *A note on invariant subspaces*, Mich. Math. J. 15(1968), 61-64.

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

[BeW] Berger, M. A. and Wang, Y., *Bounded semigroups of matrices*, Linear Algebra Appl. 166 (1992), 21-27.

[B] Bhatia, R., *Perturbation Bounds for Matrix Eigenvalues*, Longman Scientific and Technical, Essex, England, 1987.

[BBh1] Bhatia R. and Bhattacharyya, T., *A Henrici theorem for joint spectra of commuting matrices*, Proc. Amer. Math. Soc. 118 (1993), 5-14.

[BBh2] Bhatia R. and Bhattacharyya, T., *A generalization of the Hoffman-Wielandt theorem*, Linear Algebra Appl. 179 (1993), 11-17.

[BBh3] Bhatia R. and Bhattacharyya, T., *On the joint spectral radius of commuting matrices*, to appear in Studia Math.

[BDaM] Bhatia, R., Davis, Ch. and McIntosh, A., *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl. 52(1983), 45 - 67.

[Ber] Bernstein, A. R., *Invariant subspaces of polynomially compact operators on Banach spaces*, Pacif. J. Math. 21(1967), 445-464.

[BerR] Bernstein, A. R. and Robinson, A., *Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos*, Pacif. J. Math. 16(1966), 421-431.

[BFr] Bhatia, R. and Friedland, S., *Variation of Grassman powers and spectra*, Linear Algebra Appl. 40(1981), 1 - 18.

[Bh] Bhattacharyya, T., *On commuting tuples of compact operators*, submitted to Proc. Edinburgh Math. Soc.

[Br1] Brodskiǐ, M. S., *On the triangular representation of completely continuous operators with point spectra* (Russian), Usp. mat. Nauk., 16(1961), 135-141.

[Br2] Brodskiǐ, M. S., *An abstract triangular representation of completely continuous operators with a point spectrum* (Russian), Funkcional'nyǐ Analiz i ego Primenenie (Trudy 5 Konf. po Funkcional'nomu i ego Primeneniju), 25-28. Izdat. Akad. Nauk Azerbaǐdžan. SSSR, Baku, (1961).

[Br3] Brodskiǐ, M. S., *On the unicellularity of the integration operator and a theorem of Titchmarsh* (Russian), Usp. mat. Nauk, 20(1965), 189-192.

[Bu1] Bunce, J. W., *The joint spectrum of commuting non-normal operators*, Proc. Amer. Math. Soc. 29(1971), 499-505.

[Bu2] Bunce, J. W., *Models for n-tuples of commuting operators*, J. Funct. Anal. 57(1984), 21-30.

[CH] Chō, M. and Huruya, T., *On the joint spectral radius*, Proc. Roy. Irish Acad. Sect. A 91(1991), 39-44.

[CT] Chō, M. and Takaguchi, M., Joint spectra of matrices, Sci. Rep. Hirosaki Univ. 26(1991), 15-19.

[CZ] Chō, M. and Żelazko, W., *On geometric spectral radius of commuting n-tuples of operators*, Hokkaido Math. J. 21(1992), 251-258.

]bibitem[ChDa] *Choi, M.-D. and Davis, Ch., *The spectral mapping theorem for joint approximate spectrum*, Bull. Amer. Math. Soc. 80(1974), 317-321.

[Co] Conway, J. B., *A Course in Functional Analysis*, Graduate Texts in Mathematics 96, Springer-Verlag, New York, 1985.

[Cu1] Curto, R. E., *On the connectedness of invertible n-tuples*, Indiana Univ. Math. J. 29(1980), 393-406.

[Cu2] Curto, R. E., *The spectra of elementary operators*, Indiana Univ. Math. J. 32(1983), 193-197.

[Cu3] Curto, R. E., *Applications of Several Complex Variables to Multiparameter Spectral Theory*, in: Surveys of Some Reccent Results in Operator Theory, Vol. 2, J. B. Conway and B. B. Morrel(eds.), Longman Scientific and Technical, Essex, England, 1989, 25-80.

[Das1] Dash, A. T., *Joint spectra*, Studia Math. 45(1973), 225-237.

[Das2] Dash, A. T., *Joint essential spectra*, Pacific J. Math. 64(1976), 119-128.

[DL] Daubechies, I. and Lagarias, J. C., *Sets of matrices all finite products of which converge*, Linear Algebra Appl. 161(1992), 227-263.

[Da] Davis, Ch., *Perturbation of spectrum of normal operators and of commuting tuples*, Linear and Complex Analysis Problem Book (V. P. Havin, S. V. Hruschov and N. K. Nikol'ski, eds.), Lecture Notes in Math., Springer, Berlin and New York, 1984.

[DDP] Deckard, D., Douglas, R. G. and Pearcy, C., *On invariant subspaces of quasi-triangular operators*, Am. J. Math. 91(1969), 637-647.

[Do] Donoghue, W. F., *The lattice of invariant subspaces of a completely continuous quasi-nilpotent transformation*, Pacif. J. Math. 7(1957), 1031-1035.

[E1] Elsner, L., *On the variation of spectra of matrices*, Linear Algebra Appl. 47(1982), 127-138.

[E2] Elsner, L., *A note on the Hoffman-Wielandt theorem*, Linear Algebra Appl. 182(1993), 235-237.

[E3] Elsner, L., *Perturbation theorems for joint spectrum of commuting matrices: a conservative approach*, Linear Algebra Appl. 208/209(1994), 83-95.

[E4] Elsner, L., *The generalized spectral radius theorem, an analytic-geometric proof*, to appear in Linear Algebra Appl.

[FN] Furuta, T. and Nakamoto, R., *On the numerical range of an operator*, Proc. Japan Acad. 47(1971), 279-284.

[Gi] Gillespie, T. A., *An invariant subspace theorem of J. Feldman*, Pacif. J. Math. 26(1968), 67-72.

[GK] Gohberg, I. C. and Krein, M. G., *On the theory of triangular representations of non-self-adjoint operators*, Soviet Math. Dokl. 2(1961), 392-395.

[Har1] Harte, R. E., *Spectral mapping theorems*, Proc. Royal Irish Acad. 72A(1972), 89-107.

[Har2] Harte, R. E., *Tensor products, multiplication operators and the spectral mapping theorem*, Proc. Royal Irish Acad. 73A(1973), 285-302.

[Har3] Harte, R. E., *Invertibility, singularity and Joseph L. Taylor*, Proc. Royal Irish Acad. 81A(1981), 71-79.

[Hel] Helson, H., *Linear Algebra*, Hindustan Book Agency, New Delhi, 1994.

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

[HofW] Hoffman, A. J. and Wielandt, H. W., *The variation of the spectrum of a normal matrix*, Duke Math. J. 20: 37 - 39 (1953).

[HorJ] Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[K] Kreyszig, E., *Introductory Functional Analysis with Applications*, John Wiley & Sons Inc., New York, 1978.

[L] Lomonosov, V. I., *Invariant subspaces for the family of operators which commute with a completely continuous operator*, Funct. Anal. and Appl. 7: 213-214 (1973).

[McP] McIntosh, A. and Pryde, A., *A functional calculus for several commuting operators*, Indiana Univ. Math. J. 36(1987) 421-439.

[McPR] McIntosh, A., Pryde, A. and Ricker, W., *Comparison of joint spectra for certain classes of commuting operators*, Studia Math. 88(1988), 23-36.

[Mic] Michaels, A. J., *Hilden's simple proof of Lomonosov's theorem*, Advances in Math. 25(1977), 56-58.

[Min] Ming, F., *Garske's inequality for n-tuples of operators*, Integral Equations Operator Theory 14(1991), 787-793.

[Mü] Müller, V., *On the joint spectral radius*, preprint.

[MüS] Müller, V. and Soltysiak, A., *Spectral radius formula for commuting Hilbert space operators*, Studia Math. 103(1992), 329-333.

[MüV] Müller, V. and Vasilescu, F.-H., *Standard models for some commuting multioperators*, Proc. Amer. Math. Soc. 117(1993), 979-989.

[Pa] Paulsen, V. I., *Completely bounded maps and dilatios*, Longman Scientific and Technical, Essex, 1986.

[P1] Pryde, A., *A Bauer-Fike theorem for the joint spectrum of commuting matrices*, Linear Algebra Appl. 173(1992), 221-230.

[P2] Pryde, A., *Optimal matching of joint eigenvalues for normal matrices*, Monash University Analysis Paper 74 (March 1991).

[P3] Pryde, A., *Inequalities for the joint spectrum of simultaneously upper-triangularizable matrices*, Miniconference on probability and analysis (Sydney 1991), Proc. Centre Math. Austral. Nat. Univ., 29(1992), 196-207.

[RaRo] Radjavi, R. and Rosenthal, P., *Invariant Subspaces*, Springer-Verlag, Berlin, 1973.

[Ri1] Ringrose, J. R., *Super-diagonal form for compact linear operators*, Proc. Lond. Math. Soc.(3) 12(1962), 367-384.

[Ri2] Ringrose, J. R., *On the triangular representations of integral operators*, Proc. Lond. Math. Soc.(3) 12(1962), 385-399.

[Ri3] Ringrose, J. R., *Compact Non-self-adjoint Operators*, Van Nostrand Reinhold Company, London, 1971.

[Ro] Rota, G. C., *On models for linear operators*, Comm. Pure Appl. Math. 13(1960), 469-472.

[RoS] Rota, G. C. and Strang, W. G., *A note on the joint spectral radius*, Indag. Math. 22(1960), 379-381.

[SlZ] Slodkowski, Z. and Żelazko, W., *On joint spectra of commuting families of operators*, Studia Math. 50(1974), 127-148.

[StSu] Stewart, G. W. and Sun, J., *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[Su] Sun, J. G., *On the perturbation of the eigenvalues of a normal matrix*, Math. Numer. Sinica 6: 334 - 336 (1984).

[T1] Taylor, J. L., *A joint spectrum for several commuting operators*, J. Funct. Anal. 6: 172-191 (1970).

[T2] Taylor, J. L., *The analytic functional calculus for several commuting operators*, Acta Math. 125: 1-38 (1970).

[T3] Taylor, J. L., *A general framework for a multi-operator functional calculus*, Adv. in Math. 9(1972), 183-252.

[T4] Taylor, J. L., *Functions of several noncommuting variables*, Bull. Amer. Math. Soc. 79(1973), 1-34.

[Ta] Taylor, A. E., *Introduction to Functional Analysis*, John Wiley & Sons, Inc., New York, 1961.

[Wa] Waelbroeck, L., *Le calcul symbolique dans les algèbres commutatives*, J. Math. Pures et Appl. (9) 33(1954), 147-186.

[We] Wermer, J., *Banach Algebras and Several Complex Variables*, Graduate Texts in Mathematics 35, Springer-Verlag, New York, 1976.