June 2022

On Simultaneous Similarity of d-tuples of Commuting Square Matrices

Corey Connelly
University of South Florida

On Simultaneous Similarity of d-tuples of Commuting Square Matrices

by

Corey Connelly

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts
Department of Mathematics and Statistics
College of Arts and Sciences
University of South Florida

Major Professor: Sherwin Kouchekian, Ph.D.
Boris Shekhtman, Ph.D.
Ivan Rothstein, Ph.D.

Date of Approval:
June 10, 2022

Keywords: Linear Algebra, Linear Operators, Finite Dimensional Vector Space, Cyclicity

Dedication

I would like to dedicate this to my family, my husband Robert, and my daughter Thea.

Dear Mom, Dad, Erin, my grandparents, and everyone in my family, including my dearest friends: thank you! I am lucky to be surrounded by so much love. I have inherited my passion for learning from you, and I hope to make all of you proud.

Robert, thank you for all of the unconditional love and support you've given me over the years. I could not have achieved this without you. Your academic journey directly inspired me to push forward and find what I am truly capable of. I am so fortunate to have found you, and will cherish you always.

Thea, I love you more than words can express. I am in awe watching you grow and learn, and I know you will go on to do incredible things. I hope that your father and I will succeed in passing along our joy for the subject of mathematics. Everything we do is for you.

Acknowledgments

I would first like to thank my thesis advisor, Dr. Sherwin Kouchekian, who has been a wonderful professor and mentor throughout my undergraduate and graduate academic career at USF. I owe so much to his instruction and guidance.

I would also like to extend special thanks to Dr. Boris Shekhtman for providing the problem at the heart of this thesis, and for lending his help and expertise over the course of my research.

Finally, I wish to express my gratitude to the entire faculty of the Mathematics department at USF for their quality instruction and support, and to my classmates for their collaboration, assistance, and friendship throughout my educational journey.

Table of Contents

Chapter 1:
Introduction

1.1 Background

Representation theory is a field of study which examines groups, algebras and rings through the lens of linear algebra. A *representation* is a description of all indecomposable modules within an algebra. A *module* is a generalization of the notion of a vector space. Instead of scalars defined by a field, a module draws its scalars from a ring. An *indecomposable module* is non-trivial and cannot be decomposed into a direct sum of two non-trivial submodules.

Finite dimensional algebras have either finite or infinite representation type, referring to the number of indecomposable modules which exist for that algebra. In 1972, P. Donovan and M.R. Freislich introduced the dichotomy of *tame* and *wild* representation types, and posited that every finite dimensional algebra with an infinite representation type belongs to one of these two disjoint classes [2]. Drozd confirmed their conjecture in 1979, see [4]. For further reading on the tame-wild dichotomy and how each representation type is defined, see [6] and [3].

Classification problems related to algebras of wild representation type contain the daunting problem of classifying pairs of square matrices up to simultaneous similarity. That is, if $A = (A_1, A_2)$ and $B = (B_1, B_2)$ are two pairs of $N \times N$ matrices, when does there exist a one-invertible matrix S such that $A_1 = S^{-1}B_1S$ and $A_2 = S^{-1}B_2S$? I.M. Gel'fand and V.A. Ponomarev [5] proved that a solution to this problem implies a solution to the problem of classifying d-tuples (for an arbitrarily chosen d) of square matrices up to simultaneous similarity.

It has been shown by B. Shekhtman in [9] that when any d-tuple $A = (A_1, \ldots, A_d)$ of pairwise commuting $N \times N$ matrices is cyclic, A is simultaneously similar to the d-tuple of pairwise commuting $N \times N$ matrices $B = (B_1, \ldots, B_d)$ if and only if B is cyclic, and the sets of polynomials in d variables which annihilate A and B are equivalent. This thesis offers a further generalization of this result for the case of n-cyclic d-tuples of pairwise commuting $N \times N$ matrices.

1.2 Preliminaries

In this section, we introduce the notion of a matrix as the representation of a linear operator on a finite dimensional vector space, define matrix similarity, and prove the equivalence of norms on finite dimensional vector spaces. First, we recall some elementary facts about linear transformations, particularly linear operators on finite dimensional vector spaces, see K. Hoffman and R. Kunze [7] for a thorough treatment of the topic. We begin with a few key definitions.

DEFINITION 1.1. Let V and W be vector spaces over a field $\mathbb{F}$. A function $T : V \to W$ is a *linear transformation* if for any $x, y \in V$ and $\alpha, \beta \in \mathbb{F}$

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y). \tag{1.1}$$

A linear transformation from a vector space into itself is called a *linear operator*.

Let V and W be vector spaces over $\mathbb{F}$, and denote the space of linear transformations from V into W by $\mathcal{L}(V, W)$. Since sums and scalar multiples of linear transformations are themselves linear transformations, it follows that $\mathcal{L}(V, W)$ is a vector space. Suppose that X is also a vector space over $\mathbb{F}$, then it is easy to see that the composition of linear transformations $T : V \to W$ and $U : W \to X$, where $UT(v) = U(T(v))$ for $v \in V$, is also a linear transformation from V into X.

REMARK 1.2. Clearly $\mathcal{L}(V, V)$, the space of linear operators on V, is a vector space and the composition of any two linear operators on V is itself a linear operator on V. It follows easily from these facts that for any polynomial p in one or several variables, $p(A)$, where $A \in \mathcal{L}(V, V)$, is also a linear operator on V.

DEFINITION 1.3. The *image* of a linear transformation $T : V \to W$, denoted by $\operatorname{Im} T$, is the set of vectors $w \in W$ such that $Tv = w$ for some $v \in V$. When V is a finite-dimensional vector space, the dimension of the image of T is called the *rank* of T.

DEFINITION 1.4. The *kernel* of a linear transformation $T : V \to W$, denoted by $\ker T$, is the set of vectors $v \in V$ such that $Tv = 0$. When V is a finite-dimensional vector space, the dimension of the kernel of T is called the *nullity* of T.

The Rank-Nullity Theorem is a standard and powerful result in linear algebra. The following proof is closely adapted from [7].

THEOREM 1.5. (The Rank-Nullity Theorem) *Let V and W be vector spaces over a field $\mathbb{F}$, where V is finite dimensional. If T is a linear transformation from V into W, then*

$$\dim V = \dim(\operatorname{Im} T) + \dim(\ker T). \tag{1.2}$$

Proof. Let $n = \dim V$, $r = \dim(\operatorname{Im} T)$, and $k = \dim(\ker T)$ and suppose $\{v_1, \ldots, v_k\}$ is a basis of the kernel of T. We may choose vectors $v_{k+1}, \ldots, v_n \in V$, if necessary, such that $\{v_1, \ldots, v_n\}$ forms a basis of V. We will show that the set $\{Tv_1, \ldots, Tv_n\}$ forms a basis of the image of T.

Clearly the vectors $Tv_1, \ldots, Tv_n$ span $\operatorname{Im} T$. More precisely, the vectors $Tv_{k+1}, \ldots, Tv_n$ span $\operatorname{Im} T$ since $Tv_j = 0$ for $j \leq k$. Thus it is left to show that $Tv_{k+1}, \ldots, Tv_n$ are linearly independent. Let c_i be scalars such that

$$\sum_{i=k+1}^{n} c_i(Tv_i) = 0. \tag{1.3}$$

It follows from the linearity of T that

$$T\left(\sum_{i=k+1}^{n} c_i v_i\right) = 0. \tag{1.4}$$

Put $v = \sum_{i=k+1}^{n} c_i v_i$. Clearly $v \in \ker T$, and Since $\{v_1, \ldots, v_k\}$ is a basis of $\ker T$ by assumption, then there exist scalars b_i such that $v = \sum_{i=1}^{k} b_i v_i$. Now it follows from (1.4) that

$$\sum_{i=1}^{k} b_i v_i - \sum_{i=k+1}^{n} c_i v_i = 0. \tag{1.5}$$

Since $\{v_1, \ldots, v_n\}$ are linearly independent, $b_i = 0$ for $1 \leq i \leq k$ and $c_i = 0$ for $k+1 \leq i \leq n$. Thus by equation (1.3), $\{Tv_{k+1}, \ldots, Tv_n\}$ is also a linearly independent set, and therefore a basis of $\operatorname{Im} T$. This implies that the rank of T is $n - k$. Thus $r = n - k$, which can be rearranged as $n = r + k$. □

DEFINITION 1.6. A linear transformation $T : V \to W$ is called *invertible* if there exists a mapping $T^{-1} : W \to V$ such that

$$T^{-1}T = I_V \quad \text{and} \quad TT^{-1} = I_W, \tag{1.6}$$

where $I_V : V \to V$ and $I_W : W \to W$ denote the identity mappings on V and W, respectively. T is invertible if and only if it is one-to one and onto. T^{-1} is unique, and is itself a linear transformation (see [7]).

LEMMA 1.7. *If V and W are vector spaces over a field $\mathbb{F}$ and T is a linear transformation from V into W, then T is one-to-one if and only if the kernel of T is trivial; i.e.,* $\ker T = \{0\}$.

Proof. First, suppose that T is one-to-one. Then there is at most one vector $v \in V$ such that $Tv = 0$. By the assumption that T is linear, $T(0) = 0$, so the kernel of T is $\{0\}$.

Suppose that the kernel of T is trivial, and let $Tx = Ty$ for some $x, y \in V$. Clearly $Tx - Ty = 0$. By the linearity of T, $T(x - y) = 0$, so $x - y \in \ker T$. Since the kernel of T is trivial by assumption, $x - y = 0$, which implies that $x = y$. □

THEOREM 1.8. *Let V and W be finite dimensional vector spaces over a field $\mathbb{F}$, where* $\dim V = \dim W$. *Let T be a linear transformation from V into W. Then the following are equivalent:*

(1) T is invertible.

(2) T is onto.

(3) T is one-to one.

Proof. If T is invertible, (2) and (3) follow immediately by definition. It is therefore sufficient to show that T is onto if and only if T is one-to-one.

Suppose that T is onto, which is to say that $\operatorname{Im} T = W$. Then $\dim(\operatorname{Im} T) = \dim W = \dim V$. It follows from the Rank-Nullity Theorem that $\dim(\ker T) = 0$ and thus $\ker T = \{0\}$. Utilizing Lemma 1.7, one finds that T is one-to-one.

Now suppose that T is one-to-one. As a result of Lemma 1.7 we have $\ker T = \{0\}$. The Rank-Nullity Theorem then implies that $\dim(\operatorname{Im} T) = \dim V$. By the assumption that $\dim V = \dim W$, it follows that $\dim(\operatorname{Im} T) = \dim W$. Since $\operatorname{Im} T$ is a subspace of W with the same dimension, clearly $\operatorname{Im} T = W$. □

We now examine the relationship between linear operators on the finite dimensional vector space $\mathbb{C}^N$ and $N \times N$ matrices with complex entries. Let $\mathcal{B} = \{e_1, \ldots, e_N\}$ be a basis of $\mathbb{C}^N$, $v \in \mathbb{C}^N$, and suppose T is a linear operator on $\mathbb{C}^N$. We may express v as a unique linear combination of basis vectors $v = \sum_{j=1}^{N} c_j e_j$ or as a coordinate vector with respect to $\mathcal{B}$:

$$[v]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}. \tag{1.7}$$

The image of each basis vector e_j under T may itself be expressed as a linear combination $Te_j = \sum_{i=1}^{N} a_{i,j} e_i$ or coordinate vector with respect to $\mathcal{B}$:

$$[Te_j]_{\mathcal{B}} = \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{N,j} \end{bmatrix}. \tag{1.8}$$

The matrix representing the linear operator T with respect to the basis $\mathcal{B}$ is

$$\begin{bmatrix} a_{1,1} & a_{12} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix}. \tag{1.9}$$

This matrix, denoted by $[T]_{\mathcal{B}}$, is an $N \times N$ matrix with complex entries. The j^{th} column of $[T]_{\mathcal{B}}$ is given by the coordinate vector $[Te_j]_{\mathcal{B}}$. We now observe, for the image of v under T,

$$\begin{aligned} Tv &= T\left(\sum_{j=1}^{N} c_j e_j\right) \\ &= \sum_{j=1}^{N} c_j (Te_j) \\ &= \sum_{j=1}^{N} c_j \left(\sum_{i=1}^{N} a_{i,j} e_i\right) \\ &= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} c_j a_{i,j}\right) e_i. \end{aligned} \tag{1.10}$$

Each coefficient $\sum_{j=1}^{N} c_j a_{i,j} = c_1 a_{i,1} + \ldots + c_N a_{i,N}$ represents a coordinate of the vector $[Tv]_{\mathcal{B}}$. But as we see below, this coordinate vector is precisely the column matrix $[T]_{\mathcal{B}}[v]_{\mathcal{B}}$.

$$\begin{aligned}
[T]_{\mathcal{B}}[v]_{\mathcal{B}} &= \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} \\
&= \begin{bmatrix} c_1 a_{1,1} + \ldots + c_N a_{1,N} \\ c_1 a_{2,1} + \ldots + c_N a_{2,N} \\ \vdots \\ c_1 a_{N,1} + \ldots + c_N a_{N,N} \end{bmatrix}.
\end{aligned} \tag{1.11}$$

Thus we have that $[Tv]_{\mathcal{B}} = [T]_{\mathcal{B}}[v]_{\mathcal{B}}$ for all $v \in \mathbb{C}^N$.

Now suppose that $\mathcal{B}' = \{e'_1, \ldots, e'_N\}$ is another basis of $\mathbb{C}^N$. Let $d_1, \ldots, d_N$ be the coordinates of v with respect to $\mathcal{B}'$. Then $v = \sum_{j=1}^{N} c_j e_j = \sum_{k=1}^{N} d_k e'_k$.

Each basis vector e'_k for $k = 1, \ldots, N$ can also be expressed in terms of the original basis $\mathcal{B}$ as $e'_k = \sum_{\ell=1}^{N} b_{\ell,k} e_\ell$, where $b_{1,k}, \ldots, b_{N,k}$ are the coordinates of e'_k. Then it follows that

$$\begin{aligned}
\sum_{j=1}^{N} c_j e_j &= \sum_{k=1}^{N} d_k \left(\sum_{\ell=1}^{N} b_{\ell,k} e_\ell \right) \\
&= \sum_{\ell=1}^{N} \left(\sum_{k=1}^{N} d_k b_{\ell,k} \right) e_\ell
\end{aligned} \tag{1.12}$$

and with a simple change of indices $j \to \ell$, we have

$$\sum_{j=1}^{N} c_j e_j = \sum_{j=1}^{N} \left(\sum_{k=1}^{N} d_k b_{j,k} \right) e_j. \tag{1.13}$$

Recall that the linear combination representing the vector v with respect to $\mathcal{B}$ is unique, so by equation (1.13), each coordinate of v is given by $c_j = d_1 b_{j,1} + \ldots + d_1 b_{j,N}$. Now consider the matrix with each column given by $[e'_k]_{\mathcal{B}}$ for $k = 1, \ldots, N$. This is the representation matrix with respect to $\mathcal{B}$ of the unique linear operator which maps e_k to e'_k (see [7]). Denoting this linear operator by U, it follows that

$$[U]_{\mathcal{B}} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N,1} & b_{N,2} & \cdots & b_{N,N} \end{bmatrix}. \tag{1.14}$$

When $[U]_{\mathcal{B}}$ is multiplied by the coordinate vector of v with respect to $\mathcal{B}'$, we obtain

$$\begin{aligned} [U]_{\mathcal{B}}[v]_{\mathcal{B}'} &= \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,N} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_{N,1} & b_{N,2} & \cdots & b_{N,N} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \\ &= \begin{bmatrix} d_1 b_{1,1} + \ldots + d_N b_{1,N} \\ d_1 b_{2,1} + \ldots + d_N b_{2,N} \\ \vdots \\ d_1 b_{N,1} + \ldots + d_N b_{N,N} \end{bmatrix} \\ &= \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} \end{aligned} \tag{1.15}$$

which is of course the coordinate vector of v with respect to $\mathcal{B}$. That is, $[U]_{\mathcal{B}}[v]_{\mathcal{B}'} = [v]_{\mathcal{B}}$. We can summarize the above observation in the following definition.

DEFINITION 1.9. If $\mathcal{B} = \{e_1, \ldots, e_N\}$ and $\mathcal{B}' = \{e'_1, \ldots, e'_N\}$ are two bases of $\mathbb{C}^N$, then there exists an $N \times N$ matrix S such that

$$[v]_{\mathcal{B}} = S[v]_{\mathcal{B}'} \tag{1.16}$$

for all $v \in \mathbb{C}^N$. S is called a *change of basis matrix.*

A change of basis matrix is unique and invertible. In fact any linear operator is invertible if and only if its representative matrix is invertible (see [7]). It is easy to see that $[U]_{\mathcal{B}}$ as defined above is a change of basis matrix. Next, we illustrate the relationship between the representation matrices of a single linear operator with respect to two different bases, and the change of basis matrix that exists for those bases. From this relationship, the notion of matrix similarity emerges.

LEMMA 1.10. *Let T be a linear operator on $\mathbb{C}^N$, and let $\mathcal{B} = \{e_1, \ldots, e_N\}$ and $\mathcal{B}' = \{e'_1, \ldots, e'_N\}$ be bases of $\mathbb{C}^N$. Let U be the invertible linear operator which maps $e_i \mapsto e'_i$ for $i = 1, \ldots, N$. Then*

$$[T]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}}. \tag{1.17}$$

Proof. Let $w \in \mathbb{C}^N$. Then

$$\begin{aligned} ([U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}})([w]_{\mathcal{B}'}) &= ([U]_{\mathcal{B}}^{-1}([T]_{\mathcal{B}}([U]_{\mathcal{B}}[w]_{\mathcal{B}'})) \\ &= [U]_{\mathcal{B}}^{-1}([T]_{\mathcal{B}}[w]_{\mathcal{B}}) \\ &= [U]_{\mathcal{B}}^{-1}[Tw]_{\mathcal{B}} \\ &= [Tw]_{\mathcal{B}'} \\ &= [T]_{\mathcal{B}'}[w]_{\mathcal{B}'}. \end{aligned} \tag{1.18}$$

Since this holds for any $w \in \mathbb{C}^N$, $[T]_{\mathcal{B}'} = [U]_{\mathcal{B}}^{-1}[T]_{\mathcal{B}}[U]_{\mathcal{B}}$. □

DEFINITION 1.11. Let A and B be $N \times N$ matrices with complex entries. A is *similar* to B, denoted by $A \sim B$, if there exists an invertible matrix S such that

$$A = S^{-1}BS. \tag{1.19}$$

Clearly in definition 1.11, the similar matrices A and B represent the same linear operator on $\mathbb{C}^N$ with respect to two potentially different bases, and S represents a change of basis matrix.

It is easy to show that similarity is an equivalence relation. When $A = B$, let $S = I$ where I is the (invertible) identity map of $\mathbb{C}^N$. Of course $A = I^{-1}BI$, and so similarity is reflexive. Whenever $A \sim B$, we have that $A = S^{-1}BS$. Then $B = SAS^{-1}$ as a consequence of the invertibility of S. Thus $B \sim A$ and similarity is symmetric. Finally, suppose A, B, and C are $N \times N$ matrices such that $A \sim B$ and $B \sim C$. Then $A = S^{-1}BS$ for some S, $B = R^{-1}CR$ for some P, and

$$\begin{aligned} A &= S^{-1}BS \\ &= (S^{-1}R^{-1})C(RS) \\ &= (RS)^{-1}C(RS). \end{aligned} \tag{1.20}$$

Thus $A \sim C$ and similarity is transitive.

Matrices which are similar to each other share some important properties that arise from the common linear operator that they represent. Notably, if $A \sim B$, then A and B have the same determinant, rank, trace, and eigenvalues (but not eigenvectors in general). They also share the same characteristic polynomial, Jordan Canonical Form, and Weyr Canonical Form (see [7], [8]).

Before proceeding to the next chapter, we must also briefly review a few facts about norms on vector spaces. First, recall the definition and properties of a norm.

DEFINITION 1.12. Let V be a vector space over a field $\mathbb{F}$. A *norm* on V is a function $||\cdot|| : V \to \mathbb{R}$ with the following properties:

(1) $||v|| \geq 0$ for all $v \in V$.

(2) $||v|| = 0$ if and only if $v = 0$.

(3) $||\alpha v|| = |\alpha|||v||$ for all $v \in V$ and all $\alpha \in \mathbb{F}$.

(4) $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$.

Let $\mathcal{B}$ be a basis of a vector space V, and let $[v]_{\mathcal{B}} = (c_1, \ldots, c_N) \in V$. The *sup norm* or *maximum norm* $||\cdot||_\infty : V \to \mathbb{R}$ is defined by

$$||v||_\infty = \max(|c_1|, \ldots, |c_N|). \tag{1.21}$$

We will verify that the maximum norm is indeed a norm. Let $[v]_{\mathcal{B}} = (c_1, \ldots, c_N) \in V$. If v is the zero vector in V, then $c_i = 0$ for $i = 1, \ldots, N$. It follows that $||v||_\infty = \max(|0|, \ldots, |0|) = 0$. Conversely, if $||v||_\infty = 0$, then $|c_i| \leq 0$ for all $i = 1, \ldots, N$. But the absolute value is non-negative so $|c_i| = 0$ and therefore $c_i = 0$ for all $i = 1, \ldots, N$. Clearly v is the zero vector of V.

Now let $[v]_{\mathcal{B}} = (c_1, \ldots, c_N) \in V$ and $\alpha \in \mathbb{F}$. Then $\alpha v = (\alpha c_1, \ldots, \alpha c_N)$ and $|\alpha|$ is clearly non-negative, so

$$\begin{aligned} ||\alpha v||_\infty &= \max(|\alpha c_1|, \ldots, |\alpha c_N|) \\ &= \max(|\alpha||c_1|, \ldots, |\alpha||c_N|) \\ &= |\alpha| \max(|c_1|, \ldots, |c_N|) \\ &= |\alpha|||v||_\infty. \end{aligned} \tag{1.22}$$

Let $[u]_{\mathcal{B}} = (b_1, \ldots, b_N) \in V$ and $[v]_{\mathcal{B}} = (c_1, \ldots, c_N) \in V$. We see that $[u+v]_{\mathcal{B}} = (b_1+c_1, \ldots, b_N+c_N)$. By definition, $||u+v||_\infty = |b_k+c_k|$ for some $1 \leq k \leq N$. By the triangle inequality, $|b_k+c_k| \leq |b_k|+|c_k|$, so

$$||u + v||_\infty \leq |b_k| + |c_k|. \tag{1.23}$$

But then of course $|b_k| \leq ||u||_\infty$ and $|c_k| \leq ||v||_\infty$, which gives

$$||u+v||_\infty \leq ||u||_\infty + ||v||_\infty. \tag{1.24}$$

DEFINITION 1.13. A *topology* $\mathcal{T}$ on a space V is a collection of sets in V which satisfy the following criteria:

(1) $\emptyset$ and V are elements of $\mathcal{T}$.

(2) Any union of elements of $\mathcal{T}$ belongs to $\mathcal{T}$.

(3) Any finite intersection of elements of $\mathcal{T}$ belongs to $\mathcal{T}$.

DEFINITION 1.14. Let $(V, \mathcal{T})$ be a topological space. A set $E \subseteq V$ is called *open*, or $\mathcal{T}$*-open*, if $E \in \mathcal{T}$. A set $F \subseteq V$ is called *closed*, or $\mathcal{T}$*-closed*, if its complement in V is open.

DEFINITION 1.15. $K \subseteq V$ is $\mathcal{T}$*-compact* if for every collection C of open sets in $\mathcal{T}$ which covers K,

$$K \subseteq \bigcup_{E \in C} E,$$

there exists a finite subcollection $C_f \subseteq C$ of open sets which covers K,

$$K \subseteq \bigcup_{E \in C_f} E.$$

The proofs of the next two results are standard. See J.B. Conway [1].

PROPOSITION 1.16. *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be topologies on a vector space V, where $\mathcal{T}_2 \subseteq \mathcal{T}_1$. If a subset $K \subseteq V$ is $\mathcal{T}_1$-compact, then K is $\mathcal{T}_2$-compact.*

PROPOSITION 1.17. *If $||\cdot||_1$ and $||\cdot||_2$ are norms on a vector space V, then $||v||_1 \leq ||v||_2$ for all $v \in V$ if and only if $||v||_1 < 1$ whenever $||v||_2 < 1$.*

DEFINITION 1.18. If $||\cdot||_1$ and $||\cdot||_2$ are norms on a vector space V and there exist constants $c_1, c_2 > 0$ such that

$$c_1\, ||v||_2 \leq ||v||_1 \leq c_2\, ||v||_2 \tag{1.25}$$

for all $v \in V$, then $||\cdot||_1$ and $||\cdot||_2$ are said to be *equivalent*.

Two equivalent norms on a vector space induce the same topology on that space.

THEOREM 1.19. *Any two norms on a finite dimensional vector space are equivalent.*

Proof. Let V be a finite dimensional vector space. It suffices to show that any norm $||\cdot||$ on V is equivalent to the maximum norm $||\cdot||_\infty$. For $v \in V$, there are coordinates $d_1, \ldots, d_N$ such that $v = \sum_{j=1}^{N} d_j e_j$. By property (3) of a norm (definition 1.12),

$$||v|| \leq ||d_1 e_1|| + \ldots + ||d_N e_N|| \tag{1.26}$$

and by property (2),

$$||v|| \leq |d_1|||e_1|| + \ldots + |d_N|||e_N||. \tag{1.27}$$

Then, since $|d_i| \leq ||v||_\infty$ for all $i = 1, \ldots, N$,

$$||v|| \leq ||v||_\infty ||e_1|| + \ldots + ||v||_\infty ||e_N||. \tag{1.28}$$

Put $c_2 = \sum_{j=1}^{N} ||e_j||$. Then

$$||v|| \leq c_2\, ||v||_\infty. \tag{1.29}$$

Now, let $\mathcal{T}_\infty$ be the topology on V induced by $||\cdot||_\infty$ and $\mathcal{T}$ the topology induced by $||\cdot||$. Let $B = \{v \in V : ||v||_\infty \leq 1\}$. Equation (1.29) implies that $\mathcal{T} \subseteq \mathcal{T}_\infty$. B is $\mathcal{T}_\infty$-compact, so by Proposition 1.16, B is $\mathcal{T}$-compact. Thus the relative topological spaces $(B, \mathcal{T}_\infty)$ and $(B, \mathcal{T})$ agree.

Let $D = \{v \in V : ||v||_\infty < 1\}$. The set $D \subset B$ is $\mathcal{T}_\infty$-open, and therefore open in $(B, \mathcal{T}_\infty)$. Then of course D is open in $(B, \mathcal{T})$, thus must be some open set E in $\mathcal{T}$ such that $E \cap B = D$. Clearly $0 \in E$ and for some real $r > 0$, $\{v \in V : ||v|| < r\} \subseteq E$. Hence if $||v|| < r$ and $||v||_\infty \leq 1$, then $||v||_\infty < 1$.

Now we show that $||v|| < r$ implies that $||v||_\infty < 1$. Let $||v|| < r$ and put $\beta = ||v||_\infty$. Then

$$\begin{aligned} 1 &= \frac{1}{\beta}||v||_\infty \\ &= ||v/\beta||_\infty \end{aligned} \tag{1.30}$$

which is to say that $||v/\beta||_\infty \in B$. Now suppose that $\beta \geq 1$. Then

$$||v/\beta||_\alpha < \frac{r}{\beta} \leq r. \tag{1.31}$$

Then it follows that $||v/\beta||_\infty < 1$. But this contradicts equation 1.30. Thus $||v||_\infty < 1$.

By Proposition 1.17, $||v||_\infty \leq \frac{1}{r}||v||_\alpha$ for all $v \in V$. Putting $c_1 = r$ gives

$$c_1 \, ||v||_\infty \leq ||v||_\alpha. \tag{1.32}$$

□

DEFINITION 1.20. For a linear transformation T on a normed space $(V, || \cdot ||_V)$, set

$$||T|| = \sup\{||Tv||_V : v \in V \text{ and } ||v||_V \leq 1\}. \tag{1.33}$$

This is called the *operator norm*. It is a norm on the space $\mathcal{L}(V, V)$.

It is not always the case that the norm of a linear transformation T is finite. In fact $||T||$ is bounded if and only if T is continuous. For the proof of the following result, see e.g. J.B. Conway [1].

PROPOSITION 1.21. *If V is a finite dimensional normed space and W is a normed space. If $T : V \to W$ is a linear transformation, then T is continuous.*

It follows immediately from Proposition 1.21 that if V is a finite dimensional normed space, any linear operator T in $\mathcal{L}(V, V)$ is continuous and $||T|| < \infty$.

LEMMA 1.22. *Suppose that V is a finite dimensional normed space. Let $|| \cdot ||_V$ be a norm on V and let $|| \cdot ||$ be the operator norm on $\mathcal{L}(V, V)$. If T is a linear operator on V, then*

$$||Tv||_V \leq ||T|| \, ||v||_V \tag{1.34}$$

for all $v \in V$.

Proof. Let $v \in V$ and put $||v||_V = \alpha$. Clearly $\alpha \geq 0$. Then $||v/\alpha||_V = 1$, and so by definition of the operator norm, $||Tv/\alpha||_V \leq ||T||$. This implies that

$$||Tv/\alpha||_V \alpha \leq ||T||\alpha. \tag{1.35}$$

Then clearly

$$||Tv||_V \leq ||T||\alpha. \tag{1.36}$$

And since $||v||_V = \alpha$,

$$||Tv||_V \leq ||T|| \, ||v||_V. \tag{1.37}$$

□

In Chapter 2, we introduce the concepts of simultaneous similarity and cyclicity, before reproducing and expanding upon the final result in [9]. Notably, we add another equivalent statement to the theorem by proving that two pairs of A and B of pairwise commuting $N \times N$ matrices with respective cyclic vectors u and v are simultaneously similar if and only if for any polynomial p in d variables, $||p(A)u||$ and $||p(B)v||$ differ only by a constant factor. Chapter 3 then defines n-cyclicity and generalizes the main result of chapter two to n-cyclic d-tuples of commuting matrices.

Chapter 2:
Simultaneous Similarity of Pairs of Cyclic Commuting Matrices

In this chapter, we examine the necessary and sufficient conditions for pairs of cyclic commuting matrices to be simultaneously similar. Chapter 1 defines similarity for two $N \times N$ matrices, but we must expand the notion of similarity to two finite sequences, or d-*tuples*, of matrices.

DEFINITION 2.1. Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of commuting $N \times N$ matrices with complex entries. A is *simultaneously similar* to B, denoted by $A \sim B$, if there exists an invertible matrix S such that

$$A_j = S^{-1} B_j S \tag{2.1}$$

for $j = 1, \ldots, d$.

The following lemma is a consequence of the simultaneous similarity of two d-tuples of matrices. Denote the algebra of polynomials in d variables with complex coefficients by $\mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_d]$.

LEMMA 2.2. *Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of commuting $N \times N$ matrices and suppose that $A \sim B$. Then*

$$p(A) = S^{-1} p(B) S \tag{2.2}$$

for all polynomials $p \in \mathbb{C}[x]$.

Proof. By the assumption that $A \sim B$, there exists an invertible matrix S such that $A_j = S^{-1} B_j S$ for $j = 1, \ldots, d$. For $k \geq 0$, it is clear that

$$\begin{aligned} A_j^k &= A_j \ldots A_j \\ &= (S^{-1} B_j S) \ldots (S^{-1} B_j S) \\ &= S^{-1} (B_j \ldots B_j) S \\ &= S^{-1} B_j^k S. \end{aligned} \tag{2.3}$$

Thus for any $k_1, \ldots, k_d \geq 0$

$$\begin{aligned} A_1^{k_1} \ldots A_d^{k_d} &= (S^{-1}B_1^{k_1}S) \ldots (S^{-1}B_d^{k_d}S) \\ &= S^{-1}(B_1^{k_1} \ldots B_d^{k_d})S. \end{aligned} \tag{2.4}$$

Now, choose $p \in \mathbb{C}[x]$ and let m denote the degree of p. For the d-tuple of variables $x = (x_1, \ldots, x_d)$, let $\ell_1 \ldots \ell_d$ be the (non-negative) exponents associated with each corresponding variable. Put $\ell = \ell_1 + \ldots + \ell_d$. Then $p(x)$ can be represented using a multi-index summation notation, as follows:

$$p(x) = \sum_{\ell \leq m} c_\ell (x_1^{\ell_1} \ldots x_d^{\ell_d}). \tag{2.5}$$

It follows from (2.4) and (2.5) that

$$\begin{aligned} p(A) &= \sum_{\ell \leq m} c_\ell (A_1^{\ell_1} \ldots A_d^{\ell_d}) \\ &= \sum_{\ell \leq m} c_\ell (S^{-1}B_1^{\ell_1} \ldots B_d^{\ell_d}S) \\ &= S^{-1} \left(\sum_{\ell \leq m} c_\ell (B_1^{\ell_1} \ldots B_d^{\ell_d}) \right) S \\ &= S^{-1}p(B)S. \end{aligned} \tag{2.6}$$

□

Recall from remark 1.2 that a polynomial of linear operators is itself a linear operator on the same space. In other words, the above lemma states that when $A \sim B$, then for any $p \in \mathbb{C}[x]$, $p(A)$ and $p(B)$ represent the same linear operator on $\mathbb{C}^N$ under a change of basis defined by the matrix S.

DEFINITION 2.3. Let $A = (A_1, \ldots, A_d)$ be a d-tuple of commuting $N \times N$ matrices. A is *cyclic* if there exists a vector $u \in \mathbb{C}^N$ such that

$$\{p(A)u : p \in \mathbb{C}[x]\} = \mathbb{C}^N. \tag{2.7}$$

The vector u is called a *cyclic vector* for A. Equivalently, we may say that A is cyclic if each vector in $\mathbb{C}^N$ is a linear combination of vectors of the form $(A_1^{k_1} \ldots A_d^{k_d})(u)$ where $k_1, \ldots, k_d \geq 0$.

DEFINITION 2.4. Denote the set of polynomials in d variables which annihilate A as:

$$J_A = \{p \in \mathbb{C}[x] : p(A) = 0\}. \tag{2.8}$$

In other words, J_A contains those polynomials which, when applied to the d-tuple of linear operators A, result in the zero operator in $\mathcal{L}(\mathbb{C}^N, \mathbb{C}^N)$.

The main result of this chapter is a theorem adapted from [9]. The original theorem applies to d-tuples of commuting $N \times N$ matrices with complex entries, where d is an arbitrary integer. In this chapter we only consider the case of pairs (2-tuples) of commuting matrices, but it is easy to generalize the argument from pairs to d-tuples.

Another notable difference in the theorem as presented in this chapter is the addition of the equivalent statement (2). We show that statements (1) and (3) hold if and only if, for any $p \in \mathbb{C}[x]$, the norms of the vectors $p(A)u$ and $p(B)v$, where u is a cyclic vector for A and v is a cyclic vector for B, differ only by a constant factor.

THEOREM 2.5. *Let $A = (A_1, A_2)$ and $B = (B_1, B_2)$ be pairs of $N \times N$ commuting matrices. Suppose that A is cyclic with a cyclic vector u. Then the following are equivalent:*

(1) $A \sim B$.

(2) B is cyclic, and there exists a cyclic vector v for B and a constant $c > 0$ such that

$$\frac{1}{c}\,||p(B)v|| \leq ||p(A)u|| \leq c\,||p(B)v|| \tag{2.9}$$

for any polynomial p in $\mathbb{C}[x]$.

(3) B is cyclic and $J_A = J_B$.

Proof. To show that (1) implies (2), suppose that $A \sim B$, and let S be the change of basis matrix which satisfies the definition of simultaneous similarity. By assumption, A has a cyclic vector u. Put $v = Su$. Clearly $v \in \mathbb{C}^N$ and $\{p(B)v : p \in \mathbb{C}[x]\} \subseteq \mathbb{C}^N$ since for any polynomial p and vector v, $p(B)v \in \mathbb{C}^N$.

Let $w_0 \in \mathbb{C}^N$. The matrix S is an invertible linear map on $\mathbb{C}^N$ and is therefore onto. Thus there exists some $w \in \mathbb{C}^N$ for which $S(w) = w_0$. By the assumption that A is cyclic, $w = p(A)u$ for some $p \in \mathbb{C}[x]$. Application of Lemma 2.2 yields:

$$\begin{aligned} w &= p(A)u \\ &= S^{-1}p(B)Su \\ &= S^{-1}p(B)v. \end{aligned} \tag{2.10}$$

It follows that:

$$\begin{aligned} w_0 &= S(w) \\ &= SS^{-1}p(B)v \\ &= p(B)v. \end{aligned} \tag{2.11}$$

Since w_0 is an arbitrarily chosen vector in $\mathbb{C}^N$, it is clear that $\mathbb{C}^N \subseteq \{p(B)v : p \in \mathbb{C}[x]\}$. Therefore, $\mathbb{C}^N = \{p(B)v : p \in \mathbb{C}[x]\}$. That is, B is cyclic with cyclic vector v.

Let $||\cdot||$ be a norm on $\mathbb{C}^N$. The choice of norm does not matter, as any norm on $\mathbb{C}^N$ is equivalent by Theorem 1.19. Then

$$\begin{aligned} ||p(A)u|| &= ||S^{-1}p(B)Su|| \\ &= ||S^{-1}p(B)v|| \\ &\leq ||S^{-1}||\,||p(B)v|| \text{ by ineq. (1.34).} \end{aligned} \tag{2.12}$$

That is to say the norm of a vector $p(A)u$ in $\mathbb{C}^N$ is less than or equal to the operator norm of S^{-1} times the norm of the image of $p(A)u$ under S^{-1}, which is the vector $p(B)v$. Similarly,

$$\begin{aligned} ||p(B)v|| &= ||Sp(A)S^{-1}v|| \\ &= ||Sp(A)S^{-1}Su|| \\ &= ||Sp(A)u|| \\ &\leq ||S||\,||p(A)u|| \text{ by ineq. (1.34)} \end{aligned} \tag{2.13}$$

which implies that

$$\frac{1}{||S||}\,||p(B)v|| \leq ||p(A)u|| \tag{2.14}$$

.

Clearly $||S|| \neq 0$, and $||S^{-1}|| \neq 0$, as S is an invertible and therefore non-zero matrix by the definition of simultaneous similarity. Set $c = \max\{||S||, ||S^{-1}||\}$. Then $c > 0$, and by inequalities (2.12) and (2.14),

$$\frac{1}{c}\,||p(B)v|| \leq ||p(A)u|| \leq c\,||p(B)v|| \tag{2.15}$$

Thus inequality (2.9) holds.

Next, to prove that (2) implies (3). B is cyclic by assumption, so it is sufficient to show that $J_A = J_B$. Let $p \in J_A$; i.e., $p(A) = 0$. It follows that $p(A)u = 0$, and thus $||p(A)u|| = 0$.

Using inequality (2.9), we get

$$\frac{1}{c}\,||p(B)v|| \leq ||p(A)u|| = 0. \tag{2.16}$$

Since $\frac{1}{c} > 0$, (2.16) implies $||p(B)v|| = 0$ and therefore $p(B)v = 0$. Because it is a cyclic vector, clearly v is non-zero, so we must have $p(B) = 0$; i.e., $p \in J_B$. Conversely, let $q \in J_B$. Then $||q(B)v|| = 0$ and again by inequality (2.9),

$$||q(A)u|| \leq c\,||q(B)v|| = 0. \tag{2.17}$$

This implies that $q(A)u = 0$, and thus $q(A) = 0$; i.e., $q \in J_A$. Therefore $J_A = J_B$.

Finally, we show that (3) implies (1). Since both A and B are cyclic by assumption, there exist cyclic vectors u and v such that

$$\{p(A)u : p \in \mathbb{C}[x]\} = \mathbb{C}^N \quad \text{and} \quad \{p(B)v : p \in \mathbb{C}[x]\} = \mathbb{C}^N. \tag{2.18}$$

Define the mapping $S : \mathbb{C}^N \to \mathbb{C}^N$ by $Sp(A)u = p(B)v$.

To show that S is well-defined, suppose that $p(A)u, q(A)u \in \mathbb{C}^N$ such that $p(A)u = q(A)u$. Then $(p - q)(A)u = 0$ and since u is a non-zero cyclic vector, it follows that $(p - q)(A) = 0$. Thus the polynomial $p - q$ is an element of J_A. By assumption, $p - q \in J_B$ as well, so we must have $(p - q)(B) = 0$ which implies that $(p - q)(B)v = 0$. The last equality implies $p(B)v = q(B)v$; i.e., $Sp(A)u = Sq(A)u$.

To show that S is linear, let $p(A)u, q(A)u \in \mathbb{C}^N$ and $\alpha, \beta \in \mathbb{C}$. Then

$$\begin{aligned} S(\alpha p(A)u + \beta q(A)u) &= S([\alpha p + \beta q](A)u) \\ &= [\alpha p + \beta q](B)v \\ &= \alpha p(B)v + \beta q(B)v \\ &= \alpha Sp(A)u + \beta Sq(A)u. \end{aligned} \tag{2.19}$$

Let $p(B)v \in \mathbb{C}^N$. Then $p(B)v = Sp(A)u$ where $p(A)u \in \mathbb{C}^N$. That is to say, $\operatorname{Im} S = C^N$. By Theorem 1.8, since S is a linear operator on the finite dimensional vector space $\mathbb{C}^N$ and is onto, S is invertible.

Finally, let $w \in \mathbb{C}^N$. Since A is cyclic, $w = p(A)u$, for some $p \in \mathbb{C}[x]$. Then, for $j = 1, 2$:

$$\begin{aligned} SA_j w &= SA_j p(A)u \\ &= S(x_j p)(A)u \\ &= (x_j p)(B)v \\ &= B_j p(B)v \\ &= B_j S p(A)u \\ &= B_j S w. \end{aligned} \tag{2.20}$$

This holds for all $w \in \mathbb{C}^N$, so $SA_j = B_j S$ on $\mathbb{C}^N$ for $j = 1, 2$. That is, $A \sim B$. □

We now proceed to the case of a d-tuple of pairwise commuting matrices which does not have a single cyclic vector, but has some collection of n vectors which generate cyclic subspaces of $\mathbb{C}^N$. The existence of multiple cyclic vectors creates a new set of problems when attempting to find the necessary and sufficient conditions for simultaneous similarity of two such d-tuples. We examine how the equivalent statements must be modified to generalize Theorem 2.5 in Chapter 3.

Chapter 3:

Simultaneous Similarity of d-tuples of n-cyclic Commuting Matrices

Now we take a closer look at the notion of cyclicity of d-tuples of $N \times N$ commuting matrices. Theorem 2.5 applies only to d-tuples with a single cyclic vector. Therefore, we need a general definition of cyclicity.

First, we require some notation. Let V be a vector space and let $\mathcal{E}_1, \ldots, \mathcal{E}_n \subseteq V$. We denote the sum of these subsets by

$$+_{i=1}^{n} \mathcal{E}_i = \left\{ \sum_{i=1}^{n} e_i : e_i \in \mathcal{E}_i \right\}.$$

Note that this is not necessarily a direct sum. That is to say, two or more of the subsets $\mathcal{E}_1, \ldots, \mathcal{E}_n$ may intersect non-trivially. This implies that the sum of vectors representing each element of $+_{i=1}^{n} \mathcal{E}_i$ may not necessarily be unique.

DEFINITION 3.1. A d-tuple of commuting $N \times N$ matrices $A = (A_1, \ldots, A_d)$ is called *n-cyclic* if there exists an n-tuple of vectors $u = (u_1, \ldots, u_n)$ such that

$$+_{i=1}^{n} \{p(A)u_i : p \in \mathbb{C}[x]\} = \mathbb{C}^N. \tag{3.1}$$

and

$$+_{i=1}^{n-1} \{p(A)u_{k_i} : \{k_1, \ldots, k_{n-1}\} \subseteq \{1, \ldots, n\} \text{ and } p \in \mathbb{C}[x]\} \neq \mathbb{C}^N \tag{3.2}$$

The n-tuple $u = (u_1, \ldots, u_n)$ is called a *cyclic n-tuple* for A.

It should be mentioned that the above definition requires n to be the least integer for which (3.1) holds for this particular choice of n-tuple $u = (u_1, \ldots, u_n)$. That is, no choice of $n-1$ vectors from $\{u_1, \ldots, u_n\}$ can serve as a cyclic $n-1$-tuple for A. In the special case when $n = 1$, we simply say that A is cyclic; this is the same notion of cyclicity from definition 2.3.

However, our definition of minimality in the sense of 3.1 and 3.2 does not prevent an n-cyclic d-tuple A from also being m-cyclic with some cyclic m-tuple $v = (v_1, \ldots, v_m)$, where $m \neq n$ and $1 \leq m \leq N$. It is important to note that $\{v_1, \ldots, v_m\} \not\subseteq \{u_1, \ldots, u_n\}$ and $\{u_1, \ldots, u_n\} \not\subseteq \{v_1, \ldots, v_m\}$.

For the sake of clarity, we give a few examples.

EXAMPLE 3.2. Consider the the 2×2 matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \tag{3.3}$$

and the 2-tuple of vectors $u = (u_1, u_2)$ where

$$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{3.4}$$

Denote the standard basis of $\mathbb{C}^2$ by $\{e_1, e_2\}$. Let $w = \alpha e_1 + \beta e_2 \in \mathbb{C}^2$. Put $p_1(t) = \alpha t^0$ and $p_2(t) = \beta t^0$. We see that

$$\begin{aligned} p_1(A)u_1 + p_2(A)u_2 &= \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= w. \end{aligned} \tag{3.5}$$

That is to say, $u = (u_1, u_2)$ is a cyclic 2-tuple for A. Now we consider the vector

$$v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{3.6}$$

Set $p(t) = (2\alpha - \beta)\, t^0 + (\beta - \alpha)\, t$. It follows easily that

$$\begin{aligned} p(A)v &= (2\alpha - \beta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (\beta - \alpha) \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= (2\alpha - \beta) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + (\beta - \alpha) \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} (2\alpha - \beta) + (\beta - \alpha) \\ (2\alpha - \beta) + 2\,(\beta - \alpha) \end{bmatrix} \\ &= w. \end{aligned} \tag{3.7}$$

Hence we see that A is also cyclic, with cyclic vector v.

EXAMPLE 3.3. Similar to the previous example, consider the 2-tuple $A = (A_1, A_2)$, where

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}. \tag{3.8}$$

Let the vector v and 2-tuple $u = (u_1, u_2)$ be defined as in example 3.2, and let $w = \alpha e_1 + \beta e_2 \in \mathbb{C}^2$. If we put $p_1(s,t) = \alpha s$ and $p_2(s,t) = \frac{1}{2}\beta t$, then

$$\begin{aligned} p_1(A)u_1 + p_2(A)u_2 &= \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\beta \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\beta \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &= w. \end{aligned} \tag{3.9}$$

If we put $p(s,t) = \alpha s + \frac{1}{2}\beta t$, then

$$\begin{aligned} p(A)v &= \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\beta \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2}\beta \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ &= w. \end{aligned} \tag{3.10}$$

That is, A is both cyclic and 2-cyclic.

It should also be noted that for an n-cyclic d-tuple of matrices A, it is not necessarily the case that A is m-cyclic for any $1 \leq m \leq n$. The following is a simple counterexample.

EXAMPLE 3.4. Let A be the 3×3 matrix

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}. \tag{3.11}$$

We claim that A cannot be cyclic with a single cyclic vector. Denote the standard basis of $\mathbb{C}^3$ by $\{e_1, e_2, e_3\}$ and let $w = \alpha e_1 + \beta e_2 + \gamma e_3 \in \mathbb{C}^3$. It is easy to see that $Aw = \alpha e_3 + \beta e_3$ and $A^2 w = 0$. Let $p \in \mathbb{C}[x]$, which has the form $p(t) = a_0 + a_1 t + a_2 t^2 \dots$. It follows that

$$\begin{aligned} p(A)w &= a_0 w + a_1 Aw + a_2 A^2 w \dots \\ &= a_0 (\alpha e_1 + \beta e_2 + \gamma e_3) + a_1 (\alpha e_3 + \beta e_3). \end{aligned} \tag{3.12}$$

But the vectors $\alpha e_1 + \beta e_2 + \gamma e_3$ and $\alpha e_3 + \beta e_3$ together do not form a linearly independent basis of $\mathbb{C}^3$, which implies that there is no $w \in \mathbb{C}^3$ such that $\{p(A)w : p \in \mathbb{C}[x]\} = \mathbb{C}^3$. That is, A is not cyclic. Now, consider the vectors

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{3.13}$$

Observe that $u_1 = e_1$, $u_2 = e_2$, and $Au_1 = Au_2 = e_3$. Again let $w = \alpha e_1 + \beta e_2 + \gamma e_3 \in \mathbb{C}^3$. Put $p_1(t) = \alpha t^0 + \gamma t$ and $p_2(t) = \beta t^0$. We see that

$$\begin{aligned} p_1(A)u_1 + p_2(A)u_2 &= \alpha u_1 + \gamma A u_1 + \beta u_2 \\ &= \alpha e_1 + \beta e_2 + \gamma e_3 \\ &= w. \end{aligned} \tag{3.14}$$

Thus A is n-cyclic for $n = 2$, but not for $n = 1$.

DEFINITION 3.5. If $A = (A_1, \dots, A_d)$ is a d-tuple of commuting $N \times N$ matrices which is n-cyclic with a cyclic n-tuple $u = (u_1, \dots, u_n)$, we set

$$U_i = \{p(A)u_i : p \in \mathbb{C}[x]\} \tag{3.15}$$

where $i = 1, \dots, n$. We call U_i a *cyclic set* for A corresponding to the vector u_i.

For $n \geq 2$, note that the space $\mathbb{C}^N$ is not itself spanned by vectors of the form $(A_1^{k_1} \dots A_d^{k_d})(u)$, where u is a single cyclic vector for A. Instead any vector in $\mathbb{C}^N$ may be represented by a sum of vectors, each an element of a set spanned by vectors of the form $(A_1^{k_1} \dots A_d^{k_d})(u_i)$, where u_i belongs to a cyclic n-tuple for A and $k_j \in \mathbb{N} \cup \{0\}$ for $1 \leq j \leq d$.

The advantage of condition (3.2) in our definition of n-cyclicity is that it ensures that the vectors $u_1, \dots, u_n$ in a cyclic n-tuple u are necessarily linearly independent.

PROPOSITION 3.6. *If $A = (A_1, \dots, A_d)$ is a d-tuple of commuting $N \times N$ matrices with a cyclic n-tuple $u = (u_1, \dots, u_n)$, the n vectors $u_1, \dots, u_n$ are linearly independent.*

Proof. Enumerating if necessary, suppose that u_n is a dependent vector; i.e., there are constants $c_1, \ldots, c_{n-1}$ such that

$$u_n = c_1 u_1 + \ldots + c_{n-1} u_{n-1}. \tag{3.16}$$

If $w \in \mathbb{C}^N$, then it follows from the definition of n-cylicity (3.1) that $w = \sum_{i=1}^{n} p_i(A) u_i$ for some $(p_1, \ldots, p_n)$. Since each matrix A_j is a linear transformation for $1 \leq j \leq d$, it follows that $p_i(A)$ is a linear transformation for each $1 \leq i \leq n$ (see remark 1.2). Then we have that

$$\begin{aligned} w_n &= \left(\sum_{i=1}^{n-1} p_i(A) u_i \right) + p_n(A) u_n \\ &= \left(\sum_{i=1}^{n-1} p_i(A) u_i \right) + \left(p_n(A) \sum_{i=1}^{n-1} c_i u_i \right) \\ &= \sum_{i=1}^{n-1} p_i(A) u_i + \sum_{i=1}^{n-1} c_i p_n(A) u_i \\ &= \sum_{i=1}^{n-1} (p_i + c_i p_n)(A) u_i \end{aligned} \tag{3.17}$$

Put $q_i = p_i + c_i p_n$. Clearly q_i is a polynomial in d variables for $1 \leq i \leq n-1$. Thus we have that $w = \sum_{i=1}^{n-1} q_i(A) u_i$ for the arbitrarily chosen vector $w \in \mathbb{C}^N$, which contradicts condition (3.2). Therefore $u_1, \ldots, u_n$ are linearly independent. $\square$

Finally, it should be noted that the main result of this chapter holds regardless of condition (3.2).

We now wish to generalize Theorem 2.5 to the case of n-cyclic d-tuples. It is easy to see that we are no longer able to compare the norms of the vectors $p(A)u$ and $p(B)v$ for cyclic vectors u and v and some polynomial p (see Theorem 2.5 (2)). We must now compare the norms of the vectors $\sum_{i=1}^{n} p_i(A) u_i$ and $\sum_{i=1}^{n} p_i(B) v_i$ for cyclic n-tuples $(u_i, \ldots, u_n)$ and $(v_i, \ldots, v_n)$ and an n-tuple of polynomials $(p_i, \ldots, p_n)$.

Next, we generalize definition 2.4 to n-cyclic d-tuples of commuting matrices. Observe that if A is a 1-cyclic d-tuple with cyclic vector u, then for any $p \in \mathbb{C}[x]$, $p(A) = 0$ if and only if $p(A)u = 0$. This follows immediately from the fact that cyclic vectors are non-zero. Using the same reasoning, it can easily be seen that when A is an n-cyclic d-tuple with a cyclic tuple $u = (u_1, \ldots, u_n)$, for any $p \in \mathbb{C}[x]$ and $1 \leq i \leq n$, then $p(A) = 0$ if and only if $p(A)u_i = 0$.

DEFINITION 3.7. Let $A = (A_1, \ldots, A_d)$ be a d-tuple of commuting $N \times N$ matrices with a cyclic n-tuple $u = (u_1, \ldots, u_n)$. For $i = 1, \ldots, n$, define:

$$J_{A,u_i} = \{p \in \mathbb{C}[x] : p(A)u_i = 0\}. \tag{3.18}$$

LEMMA 3.8. *Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of commuting $N \times N$ matrices. Suppose that A and B are n-cyclic with cyclic n-tuples $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, respectively. For any $1 \leq i \leq n$, if $J_{A,u_i} = J_{B,v_i}$, then $\dim U_i = \dim V_i$.*

Proof. Suppose that $J_{A,u_i} = J_{B,v_i}$ for some $1 \leq i \leq n$. Put $\dim U_i = \ell$ and $\dim V_i = m$, and without loss of generality, suppose $\dim U_i > \dim V_i$ (i.e., $\ell > m$). Let $\{p_1(A)u_i, \ldots, p_\ell(A)u_i\}$ be a basis of U_i. Consider $\{p_1(B)v_i, \ldots, p_\ell(B)v_i\} \subseteq V_i$, the set consisting of the same polynomials $p_1, \ldots, p_\ell$ on B, times the corresponding cyclic vector v_i. There can be at most m linearly independent vectors in this set, since $\dim V_i = m$. Therefore, there exists some k such that $1 \leq k < \ell$ and $p_k(B)v_i$ is equal to a linear combination of $p_1(B)v_i, \ldots, p_{k-1}(B)v_i$. Then

$$p_k(B)v_i = \sum_{j=1}^{k-1} c_j p_j(B)v_i. \tag{3.19}$$

Subtracting $p_k(B)v_i$ from each side of equation (3.19) gives

$$0 = \left(\sum_{j=1}^{k-1} c_j p_j(B)v_i\right) - p_k(B)v_i. \tag{3.20}$$

Now we may combine the terms of equation (3.20) into a single sum where $c_k = -1$ to obtain

$$0 = \sum_{j=1}^{k} c_j p_j(B)v_i. \tag{3.21}$$

The linear combination of polynomials $\sum_{j=1}^{k} c_j p_j$ is itself a polynomial in $\mathbb{C}[x]$, so by equation (3.21), $\sum_{j=1}^{k} c_j p_j \in J_{B,v_i}$. By assumption, $J_{A,u_i} = J_{B,v_i}$, so $\sum_{j=1}^{k} c_j p_j \in J_{A,u_i}$. It then follows that

$$0 = \sum_{j=1}^{k} c_j p_j(A)u_i. \tag{3.22}$$

After separating the k^{th} term from the sum in equation (3.22) (recalling that by $c_k = -1$), we obtain

$$0 = \left(\sum_{j=1}^{k-1} c_j p_j(A) u_i \right) - p_k(A) u_i; \tag{3.23}$$

or equivalently,

$$p_k(A) u_i = \sum_{j=1}^{k-1} c_j p_j(A) u_i. \tag{3.24}$$

In other words, $p_k(A)u_i$ is a linear combination of vectors $p_1(A)u_i, \ldots, p_{k-1}(A)u_i$. But this contradicts the assumption that $\{p_1(A)u_i, \ldots, p_\ell(A)u_i\}$ is a linearly independent set of vectors in U_i. Therefore $\dim U_i \leq \dim V_i$, and by an identical argument, $\dim U_i \geq \dim V_i$. Thus $\dim U_i = \dim V_i$ for $i = 1, \ldots, n$. □

Let A be a d-tuple with a cyclic n-tuple $u = (u_1, \ldots, u_n)$. Consider an n-tuple of polynomials for which $\sum_{i=1}^{n} p_i(A) u_i = 0$. Note that it is not necessary that $p_i(A)u_i = 0$ for all (or any) $1 \leq i \leq n$. With this observation in mind, we make the following definition.

DEFINITION 3.9. Define:

$$J_{A,u} = \left\{ (p_1, \ldots, p_n) \in \mathbb{C}[x] : \sum_{i=1}^{n} p_i(A) u_i = 0 \right\}. \tag{3.25}$$

Our motivation for defining a set in this way will become clear shortly.

LEMMA 3.10. *Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of commuting $N \times N$ matrices. Suppose that A and B are n-cyclic with cyclic n-tuples $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, respectively. If $J_{A,u} = J_{B,v}$, then $J_{A,u_i} = J_{B,v_i}$ for all $i = 1, \ldots, n$.*

Proof. Fix $1 \leq k \leq n$ and let $q \in J_{A,u_k}$. Let $p = (0, \ldots, q, \ldots, 0)$ with q as the k^{th} term. It follows by definition 3.9 that $p \in J_{A,u}$. Since $J_{A,u} = J_{B,v}$ by assumption, we have that $p \in J_{B,v}$. It is easily seen that $q \in J_{B,v_k}$. Thus $J_{A,u_k} \subseteq J_{B,v_k}$. By an identical argument, $J_{B,v_k} \subseteq J_{A,u_k}$. Since k was arbitrarily chosen, $J_{B,v_i} = J_{A,u_i}$ for all $i = 1, \ldots, n$. □

REMARK 3.11. The converse of Lemma 3.10 is not true in general. For $p = (p_1, \ldots, p_n) \in J_{A,u}$ it is not necessarily the case that each $p_i \in J_{A,u_i}$.

EXAMPLE 3.12. Let A be the 2-cyclic 3×3 matrix from example 3.4, with the same cyclic 2-tuple

$u = (u_1, u_2)$. If $p_1(t) = t$ and $p_2(t) = -t$, then

$$\begin{aligned} p_1(A)u_1 + p_2(A)u_2 &= Au_1 - Au_2 \\ &= e_3 - e_3 \\ &= 0 \end{aligned} \tag{3.26}$$

That is, $(p_1, p_2) \in J_{A,u}$ but $p_1(A)u_1 \neq 0$ and $p_2(A)u_2 \neq 0$.

What follows is an immediate corollary of Lemma 3.8 and Lemma 3.10.

COROLLARY 3.13. *Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of commuting $N \times N$ matrices. Suppose that A and B are n-cyclic with cyclic n-tuples $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$, respectively. If $J_{A,u} = J_{B,v}$, then $\dim U_i = \dim V_i$ for all $i = 1, \ldots, N$.*

We may now proceed to the main result of the chapter.

THEOREM 3.14. *Let $A = (A_1, \ldots, A_d)$ and $B = (B_1, \ldots, B_d)$ be d-tuples of $N \times N$ commuting matrices. Suppose that A is n-cyclic with a cyclic n-tuple of vectors $u = (u_1, \ldots, u_n)$. Then the following are equivalent:*

(1) $A \sim B$.

(2) B is n-cyclic and there exists a cyclic n-tuple $v = (v_1, \ldots, v_n)$ for B and a constant $c > 0$ such that

$$\frac{1}{c} \left\| \sum_{i=1}^{n} p_i(B)v_i \right\| \leq \left\| \sum_{i=1}^{n} p_i(A)u_i \right\| \leq c \left\| \sum_{i=1}^{n} p_i(B)v_i \right\| \tag{3.27}$$

for any n-tuple of polynomials $(p_1, \ldots, p_n)$ such that $p_i \in \mathbb{C}[x]$ for $i = 1, \ldots, n$.

(3) B is n-cyclic and there exists a cyclic n-tuple $v = (v_1, \ldots, v_n)$ for B such that $J_{A,u} = J_{B,v}$.

Proof. First we show (1) implies (2). Suppose that $A \sim B$. By assumption, A has a cyclic n-tuple $u = (u_1, \ldots, u_n)$. Put $v_i = Su_i$ for $i = 1, \ldots, n$. We will demonstrate that $v = (v_1, \ldots, v_n)$ is a cyclic n-tuple of vectors for B.

It is easily seen that $+_{i=1}^{n} V_i \subseteq \mathbb{C}^N$. Let $w_0 \in \mathbb{C}^N$. S is onto, so there exists some $w \in \mathbb{C}^N$ for which $S(w) = w_0$. By the assumption that A is n-cyclic, $w = w_1 + \ldots + w_n$, where $w_i \in U_i$ for $i = 1, \ldots, n$.

Each vector $w_i = p_i(A)u_i$ for some polynomial p_i. Then, applying Lemma 2.2,

$$\begin{aligned} w &= \sum_{i=1}^{n} p_i(A)u_i \\ &= \sum_{i=1}^{n} S^{-1}p_i(B)Su_i \\ &= \sum_{i=1}^{n} S^{-1}p_i(B)v_i \\ &= S^{-1}\sum_{i=1}^{n} p_i(B)v_i. \end{aligned} \tag{3.28}$$

The last step follows from the linearity of S^{-1}. As a consequence, we have

$$\begin{aligned} w_0 &= S(w) \\ &= S\left(S^{-1}\sum_{i=1}^{n} p_i(B)v_i\right) \\ &= \sum_{i=1}^{n} p_i(B)v_i. \end{aligned} \tag{3.29}$$

Of course $p_i(B)v_i \in V_i$ for $i = 1, \ldots, d$. Since w_0 was an arbitrarily chosen vector in $\mathbb{C}^N$, $\mathbb{C}^N \subseteq +_{i=1}^{n} V_i$. Therefore, B is n-cyclic with cyclic n-tuple $v = (v_1, \ldots, v_n)$.

Let $||\cdot||$ be a norm on $\mathbb{C}^N$. Then

$$\begin{aligned} \left|\left|\sum_{i=1}^{n} p_i(A)u_i\right|\right| &= \left|\left|\sum_{i=1}^{n} S^{-1}p_i(B)Su_i\right|\right| \\ &= \left|\left|\sum_{i=1}^{n} S^{-1}p_i(B)v_i\right|\right| \\ &= \left|\left|S^{-1}\sum_{i=1}^{n} p_i(B)v_i\right|\right| \\ &\leq ||S^{-1}|| \left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| \text{ by ineq. (1.34).} \end{aligned} \tag{3.30}$$

Similarly,

$$\begin{aligned}\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| &= \left|\left|\sum_{i=1}^{n} Sp_i(A)S^{-1}v_i\right|\right| \\ &= \left|\left|\sum_{i=1}^{n} Sp_i(A)S^{-1}Su_i\right|\right| \\ &= \left|\left|\sum_{i=1}^{n} Sp_i(A)u_i\right|\right| \\ &= \left|\left|S\sum_{i=1}^{n} p_i(A)u_i\right|\right| \\ &\leq ||S|| \left|\left|\sum_{i=1}^{n} p_i(A)u_i\right|\right| \text{ by ineq. (1.34)}\end{aligned} \tag{3.31}$$

which implies that,

$$\frac{1}{||S||}\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| \leq \left|\left|\sum_{i=1}^{n} p_i(A)u_i\right|\right|. \tag{3.32}$$

Since S is invertible, $||S||$ and $||S^{-1}||$ are non-zero. Set $c = \max\{||S||, ||S^{-1}||\}$. Then $c > 0$, and by inequalities (3.30) and (3.32),

$$\frac{1}{c}\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| \leq \left|\left|\sum_{i=1}^{n} p_i(A)u_i\right|\right| \leq c\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right|. \tag{3.33}$$

Next to prove (2) implies (3). Since B is n-cyclic by assumption, it suffices to show that $J_{A,u} = J_{B,v}$. Let $(p_1, \ldots, p_n) \in J_{A,u}$; that is, $\sum_{i=1}^{n} p_i(A)u_i = 0$. Thus $|| \sum_{i=1}^{n} p_i(A)u_i|| = 0$, and by inequality (3.27),

$$\frac{1}{c}\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| \leq \left|\left|\sum_{i=1}^{n} p_i(A)u_i\right|\right| = 0. \tag{3.34}$$

Since $\frac{1}{c} > 0$, it follows that $\left|\left|\sum_{i=1}^{n} p_i(B)v_i\right|\right| = 0$. Thus $\sum_{i=1}^{n} p_i(B)v_i = 0$; i.e., $(p_1, \ldots, p_n) \in J_{B,v}$. Conversely, let $(q_1, \ldots, q_n) \in J_{B,v}$. It follows that $\left|\left|\sum_{i=1}^{n} q_i(B)v_i\right|\right| = 0$. Once again by inequality (3.27),

$$\left|\left|\sum_{i=1}^{n} q_i(A)u_i\right|\right| \leq c\left|\left|\sum_{i=1}^{n} q_i(B)v_i\right|\right| = 0. \tag{3.35}$$

This implies that $\sum_{i=1}^{n} q_i(A)u_i = 0$ (i.e., $(q_1, \ldots, q_n) \in J_{A,u}$). Thus $J_{A,u} = J_{B,v}$.

Finally to show (3) implies (1), fix $1 \leq i \leq n$ and define the map $S_i : U_i \to V_i$ by

$$S_i p(A) u_i = p(B) v_i. \tag{3.36}$$

Since $\{p(A)u_i : p \in \mathbb{C}[x]\} = U_i$ and $\{p(B)v_i : p \in \mathbb{C}[x]\} = V_i$, clearly S_i is defined on U_i.

To show that S_i is well-defined, suppose that $p(A)u_i = q(A)u_i$. Then $(p-q)(A)u_i = 0$. The assumption that $J_{A,u} = J_{B,v}$ implies that $J_{A,u_i} = J_{B,v_i}$ by Lemma 3.10. Since $p - q \in J_{A,u_i}$, it follows that $p - q \in J_{B,u_i}$. That is, $(p-q)(B)v_i = 0$, which implies that $p(B)v_i = q(B)v_i$. Therefore $S_i p(A)u_i = S_i q(A)u_i$ by the definition of S_i.

To show that S_i is linear, let $p(A)u_i, q(A)u_i \in U_i$ and $\alpha, \beta \in \mathbb{C}$. Then consequently,

$$\begin{aligned} S_i(\alpha p(A)u_i + \beta q(A)u_i) &= S_i([\alpha p + \beta q](A)u_i) \\ &= [\alpha p + \beta q](B)v_i \\ &= \alpha p(B)v_i + \beta q(B)v_i \\ &= \alpha S_i p(A)u_i + \beta S_i q(A)u_i. \end{aligned} \tag{3.37}$$

Let $p(B)v_i \in V_i$. By the definition of S_i, $p(B)v_i = S_i p(A)u_i$ where $p(A)u_i \in U_i$. Thus for any arbitrarily chosen vector in V_i, there exists some element of U_i which maps to that vector. In other words, S_i is onto. Since $J_{A,u} = J_{B,v}$ by assumption, Corollary 3.13 implies that $\dim U_i = \dim V_i$. Therefore, by Theorem 1.8, S_i is invertible.

Now let $w_i \in U_i$. It follows that $w_i = p(A)u_i$ for some $p \in \mathbb{C}[x]$. Then for $j = 1, \ldots, d$:

$$\begin{aligned} S_i A_j w_i &= S_i A_j p(A) u_i \\ &= S_i (x_j p)(A) u_i \\ &= (x_j p)(B) v_i \\ &= B_j p(B) v_i \\ &= B_j S_i p(A) u_i \\ &= B_j S_i w_i. \end{aligned} \tag{3.38}$$

As a result, $S_i A_j = B_j S_i$ on U_i for $j = 1, \ldots, d$. Since i was arbitrarily chosen, this holds for all $i = 1, \ldots, n$.

We must now show that there exists an invertible linear operator S on $\mathbb{C}^N$ for which $SA_j = BS_j$ for

$j = 1, \ldots, d$. Since A is n-cyclic by assumption, any vector in $\mathbb{C}^N$ can be expressed as $\sum_{i=1}^{n} p_i(A)u_i$ where each term $p_i(A)u_i \in U_i$. Define $S : \mathbb{C}^N \to \mathbb{C}^N$ by

$$\begin{aligned} S\left(\sum_{i=1}^{n} p_i(A)u_i\right) &= \sum_{i=1}^{n} S_i p_i(A)u_i \\ &= \sum_{i=1}^{n} p_i(B)v_i. \end{aligned} \tag{3.39}$$

Since B is also n-cyclic, $\sum_{i=1}^{n} p_i(B)v_i \in \mathbb{C}^N$ where $p_i(B)v_i \in V_i$ for $i = 1, \ldots, n$. Thus, S is defined on $\mathbb{C}^N$. To show that S is well-defined, let $\sum_{i=1}^{n} p_i(A)u_i = \sum_{i=1}^{n} q_i(A)u_i$. It follows that

$$\sum_{i=1}^{n} p_i(A)u_i - \sum_{i=1}^{n} q_i(A)u_i = 0. \tag{3.40}$$

Combining the terms of (3.40) yields

$$\sum_{i=1}^{n} p_i(A)u_i - q_i(A)u_i = 0. \tag{3.41}$$

We may then consider each $p_i - q_i$ as a single polynomial on A. It follows that the n-tuple $(p_1 - q_1, \ldots, p_n - q_n) \in J_{A,u}$. That is,

$$\sum_{i=1}^{n} (p_i - q_i)(A)u_i = 0. \tag{3.42}$$

Since $J_{A,u} = J_{B,v}$ by assumption,

$$\sum_{i=1}^{n} (p_i - q_i)(B)v_i = 0. \tag{3.43}$$

Which is to say that $(p_1 - q_1, \ldots, p_n - q_n) \in J_{B,v}$. Rewriting equation (3.43) implies

$$\sum_{i=1}^{n} p_i(B)v_i = \sum_{i=1}^{n} q_i(B)v_i. \tag{3.44}$$

Then by the definition of the mapping S_i for each i,

$$\sum_{i=1}^{n} S_i p_i(A)u_i = \sum_{i=1}^{n} S_i q_i(A)u_i. \tag{3.45}$$

Using the definition of S, we have that the image of $\sum_{i=1}^{n} p_i(A)u_i$ and $\sum_{i=1}^{n} q_i(A)u_i$ under S are equal:

$$S\left(\sum_{i=1}^{n} p_i(A)u_i\right) = S\left(\sum_{i=1}^{n} q_i(A)u_i\right). \tag{3.46}$$

Therefore S is well-defined.

The linearity of S follows from the linearity of $S_1, \ldots, S_n$. Let $\sum_{i=1}^{n} p_i(A)u_i$, $\sum_{i=1}^{n} q_i(A)u_i \in \mathbb{C}^N$ and $\alpha, \beta \in \mathbb{C}$. Then

$$\begin{aligned}
S\left(\alpha\sum_{i=1}^{n} p_i(A)u_i + \beta\sum_{i=1}^{n} q_i(A)u_i\right) &= S\left(\sum_{i=1}^{n} \alpha p_i(A)u_i + \sum_{i=1}^{n} \beta q_i(A)u_i\right) \\
&= S\left(\sum_{i=1}^{n} \alpha p_i(A)u_i + \beta q_i(A)u_i\right) \\
&= S\left(\sum_{i=1}^{n} [\alpha p_i + \beta q_i](A)u_i\right) \\
&= \sum_{i=1}^{n} S_i[\alpha p_i + \beta q_i](A)u_i \\
&= \sum_{i=1}^{n} \alpha(S_i p_i(A)u_i) + \beta(S_i q_i(A)u_i) \\
&= \alpha\sum_{i=1}^{n} S_i p_i(A)u_i + \beta\sum_{i=1}^{n} S_i q_i(A)u_i \\
&= \alpha S\left(\sum_{i=1}^{n} p_i(A)u_i\right) + \beta S\left(\sum_{i=1}^{n} q_i(A)u_i\right).
\end{aligned} \tag{3.47}$$

To show that S is invertible, let $w_0 \in \mathbb{C}^N$. Since B is n-cyclic, there are polynomials $q_1, \ldots, q_n \in \mathbb{C}[x]$ such that

$$\begin{aligned}
w_0 &= \sum_{i=1}^{n} q_i(B)v_i \\
&= \sum_{i=1}^{n} S_i q_i(A)u_i \\
&= S\left(\sum_{i=1}^{n} q_i(A)u_i\right).
\end{aligned} \tag{3.48}$$

Clearly $\sum_{i=1}^{n} q_i(A)u_i \in \mathbb{C}^N$ since A is n-cyclic, with cyclic n-tuple $(u_1, \ldots, u_n)$. Therefore S is onto and Theorem 1.8 implies that S is invertible.

Finally, let $w \in \mathbb{C}^N$. Since A is n-cyclic, then $w = \sum_{i=1}^{n} p_i(A)u_i$ for some $p_1, \dots, p_n \in \mathbb{C}[x]$. For $j = 1, \dots, d$, it follows that:

$$\begin{aligned} SA_j w &= SA_j \left(\sum_{i=1}^{n} p_i(A)u_i \right) \\ &= S \left(\sum_{i=1}^{n} A_j p_i(A)u_i \right) \\ &= S \left(\sum_{i=1}^{n} (x_j p_i)(A)u_i \right) \\ &= \sum_{i=1}^{n} S_i (x_j p_i)(A)u_i \\ &= \sum_{i=1}^{n} (x_j p_i)(B)v_i \\ &= \sum_{i=1}^{n} B_j p_i(B)v_i \\ &= B_j \sum_{i=1}^{n} p_i(B)v_i \\ &= B_j \sum_{i=1}^{n} S_i p_i(A)u_i \\ &= B_j S \left(\sum_{i=1}^{n} p_i(A)u_i \right) \\ &= B_j S w. \end{aligned} \tag{3.49}$$

This implies that $SA_j = B_j S$ on $\mathbb{C}^N$ for $j = 1, \dots, d$. That is, $A \sim B$. □

3.1 Conclusion

In summary, we have identified the necessary and sufficient conditions for the simultaneous similarity of n-cyclic d-tuples of commuting $N \times N$ matrices with complex entries.

The careful definition of $J_{A,u}$ and $J_{B,v}$ in chapter 3 is the crucial piece which makes the generalization of Theorem 2.5 to Theorem 3.14 possible. This construction specifically ensures, in the final part of the theorem, that the linear operator S is well-defined even when two or more cyclic sets of A (and B) have a non-trivial intersection.

An interesting possibility for further research would be generalizing this theorem to d-tuples of linear operators on infinite dimensional vector spaces. Statement (2) of Theorem 3.14, regarding the comparability of norms, will likely be a useful condition for exploring this case.

References

[1] J.B. Conway. *A Course in Functional Analysis*. Springer-Verlag, 1985.

[2] P. Donovan and M.R. Freislich. “Some evidence for an extension of the Brauer-Thrall conjecture”. In: *Sonderforschungsbereich Theor. Math* 40 (1972), pp. 24–26.

[3] Peter Dräxler and Rainer Nörenberg. “Classification Problems in the Representation Theory of Finite-Dimensional Algebras”. In: *Computational Methods for Representations of Groups and Algebras*. Ed. by P. Dräxler, C. M. Ringel, and G. O. Michler. Birkhäuser Basel, 1999, pp. 3–28.

[4] Yu. A. Drozd. “Tame and wild matrix problems”. In: *Representations and Quadratic Forms* (1979), pp. 39–74.

[5] I.M. Gel’fand and V.A. Ponomarev. “Remarks on the classification of a pair of commuting linear transformations in a finite-dimensional space”. In: *Funktsional. Anal. i Prilozhen.* 3 (4 1969), pp. 81–82.

[6] Michiel Hazewinkel, Nadiya Gubareni, and V. V. Kirichenko. “Algebras, rings and modules. Vol. 2”. In: Springer, 2007.

[7] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice Hall Inc., 1971.

[8] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[9] Boris Shekhtman. “On Cyclicity and Regularity of Commuting Matrices”. In: *Mathematics and Statistics* 1 (Oct. 2013), pp. 157–161.