Fall 11-6-2012

Diagonal And Triangular Matrices

Hamdan Al Alsulaimani
al_rqai@hotmail.com

DIAGONAL AND TRIANGULAR MATRICES

by

Hamdan Alsulaimani

B.S., Kuwait University, 2008

A Research Paper
Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree

Department of Mathematics
in the Graduate School
Southern Illinois University Carbondale
December 2012

RESEARCH PAPER APPROVAL

DIAGONAL AND TRIANGULAR MATRICES

By

Hamdan Alsulaimani

A Research Paper Submitted in Partial

Fulfillment of the Requirements

for the Degree of

Master of Science

in the field of Mathematics

Approved by:

Robert Fitzgerald

Philip Feinsilver

John McSorley

Graduate School
Southern Illinois University Carbondale
November 1, 2012

AN ABSTRACT OF THE RESEARCH PAPER OF

HAMDAN ALSULAIMANI, for the Master of Science in Mathematics, presented on NOV 1 2012, at Southern Illinois University Carbondale.

TITLE: Diagonal (Triangular) Matrices

PROFESSOR: Dr. R. Fitzgerald

I present the Triangularization Lemma which says that let P be a set of properties, each of which is inherited by quotients. If every collection of transformations on a space of dimension greater than 1 that satisfies P is reducible, then every collection of transformations satisfying P is triangularizable. I also present Burnside's Theorem which says that the only irreducible algebra of linear transformations on the finite-dimensional vector space $\mathcal{V}$ of dimension greater than 1 is the algebra of all linear transformations mapping $\mathcal{V}$ into $\mathcal{V}$. Moreover, I introduce McCoy's Theorem which says that the pair $\{A,B\}$ is triangularizable if and only if p(A,B)(AB-BA) is nilpotent for every noncommutative polynomial p. And then I show the relation between McCoy's Theorem and Lie algebras.

TABLE OF CONTENTS

INTRODUCTION

This paper shows important definitions and theorems for matrices. The main purpose in writing this paper is to explain the concepts of simultaneous diagonalization (triangularization) of matrices and how they are related to Lie algebra. Moreover, this paper may help the students and the mathematical researchers in the area of matrices.

Chapter 1 reviews some basic concepts which help us to understand some definitions and to prove some theorems in next chapters. Also it contains some examples as a review for reader.

Chapter 2 presents some theorems which are related with the simultaneously diagonalizable matrices and it presents the proofs of these theorems. Also, it shows that two matrices are simultaneously similar to diagonal matrices if and only if they commute and each is similar to a diagonal matrix.

Chapter 3 introduces the triangularizablility of transformations. Also this chapter shows some important theorems which are related with concepts of the triangularizablility as Burnside's Theorem which says that the only irreducible algebra of linear transformations on the finite-dimensional vector space $\mathcal{V}$ of dimension greater than 1 is the algebra of all linear transformations mapping $\mathcal{V}$ into $\mathcal{V}$ and McCoy's Theorem which says that the pair {A,B} is triangularizable if and only if p(A,B)(AB-BA) is nilpotent for every noncommutative polynomial p.

Chapter 4 presents the definition of a Lie algebra and how it is related with the simultaneously diagonalization (triangularization) of matrices. For instance, this chapter proves the theorem which says that the matrices A and B are simultaneously diagonal if and only if $[\mathcal{A},\mathcal{A}]=0$ where $\mathcal{A}$ is the lie algebra generated by A, B and the matrices A, B are diagonalizable. Moreover, it shows that the matrices A and B are simultaneously triangularizable if and only if the Lie algebra $\mathcal{A}$ is solvable.

CHAPTER 1

BACKGROUND

1.1 REVIEW

This chapter reviews some basic concepts of linear algebra needed for the later chapters. This material is adapted from [1]. Throughout $\mathcal{V}$ denotes a vector space over $\mathbb{C}$.
Let $\mathcal{V}^*$ be all linear tranformations $T : \mathcal{V} \longrightarrow \mathbb{C}$ and let $\mathcal{V}^{**}$ be all linear transformations $S : \mathcal{V}^* \longrightarrow \mathbb{C}$.

Definition. Consider the evaluation map is

$$e : \mathcal{V} \longrightarrow \mathcal{V}^{**}$$

$$e(v) = S_v \quad S_v(T) = T(v)$$

$\mathcal{V}$ is $reflexive$ if e is 1-1 and onto. In this case, every linear $S : \mathcal{V}^* \longrightarrow \mathbb{C}$. Looks like $S = S_v$ for some $v \in \mathcal{V}$.

Theorem 1.1.1. *If $\mathcal{V}$ is finite dimensional, then $\mathcal{V}$ is reflexive.*

Theorem 1.1.2. ***(Dimension Theorem)***
Let $\mathcal{V}$ and $\mathcal{W}$ be a vector spaces, and let $T : \mathcal{V} \longrightarrow \mathcal{W}$ be linear. If $\mathcal{V}$ is finite-dimensional, then

$$nullity(T) + rank(T) = dim(\mathcal{V})$$

Definition. Let λ be an eigenvalue of $\mathbf{A}$, then any non-zero vector $\mathbf{X}$ which satisfies the relation

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{0}$$

(i.e $\mathbf{AX} = \lambda\mathbf{X}$) is called the *eigenvector* of $\mathbf{A}$, and it is said to be associated with the eigenvalue λ.

Definition. Let T be a linear operator on a vector space $\mathcal{V}$, and λ be an eigenvalue of T. Define $E_\lambda = \{x \in \mathcal{V} : T(x) = \lambda x\} = N(T - \lambda I)$. The set E_λ is called the *eigenspace* of T corresponding to the eigenvalue λ. Namely, the set of all eigenvectors is the eigenspace E_λ.

Example 1.1.1. Let

$$A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$$

First, we have to find characteristic polynomial(det($t\mathbf{I} - \mathbf{A}$)) to find eigenvalues. Now, CP=det($t\mathbf{I} - \mathbf{A}$)= $t^2 - 5t + 6$=(t-2)(t-3), then $\lambda_1 = 2$ and $\lambda_2 = 3$. Thus λ_1 and λ_2 are the eigenvalues of $\mathbf{A}$. Second, we have to find null space of $\mathbf{A} - \lambda\mathbf{I}$ to find the eigenvectors of $\mathbf{A}$. When $\lambda_1 = 2$, then $(\mathbf{A} - 2\mathbf{I})x = 0$, so the eigenvector of $\mathbf{A}$ is

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

.When $\lambda_2 = 3$, then $(\mathbf{A} - 3\mathbf{I})x = 0$, so the eigenvector of $\mathbf{A}$ is

$$\begin{pmatrix} 1/2 \\ 1 \end{pmatrix}$$

Theorem 1.1.3. *Let T be a linear operator on a finite-dimensional vector space $\mathcal{V}$, and let λ be an eigenvalue of T having multiplicity m. Then $1 \leq dim(E_\lambda) \leq m$*

Definition. An eigenvalue of a matrix is *regular* if its multiplicity is equal to the $dim(E_\lambda)$.

Theorem 1.1.4. *A matrix is similar to a diagonal matrix if and only if all its eigenvalues are regular.*

This theorem states, in fact, that $\mathbf{A}$ is similar to a diagonal matrix if and only if $dim\mathbf{R}(\lambda\mathbf{I} - \mathbf{A}) = n - m_\lambda(\mathbf{A})$ for every value of λ.

Theorem 1.1.5. *Let $\mathcal{V}$ be a vector space with dimension n. Then every linearly independent subset of $\mathcal{V}$ can be extend to a basis for $\mathcal{V}$.*

Definition. A matrix $\mathbf{A} \in \mathbf{M}_{n \times n}(\mathbb{C})$ is called *nilpotent* if, for some positive integer k, $\mathbf{A}^k = 0$, where 0 is the $n \times n$ zero matrix.

CHAPTER 2

SIMULTANEOUS DIAGONALIZATION

2.1 DIAGONAL MATRICES

In general, this chapter studies matrices which are simultaneously similar to diagonal matrices. We show that two matrices are simultaneously similar to diagonal matrices if and only if they commute and each is similar to a diagonal matrix. This material is adapted from [4].

Definition. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, ...$ are *simultaneously similar to diagonal* (*triangular*) *matrices* if there exists a matrix S such that $\mathbf{S}^{-1}\mathbf{AS}, \mathbf{S}^{-1}\mathbf{BS}, \mathbf{S}^{-1}\mathbf{CS}$,...are all diagonal (triangular) matrices.

Definition. A *matrix of type* $(r_1, ..., r_k)$ is a matrix of order $r_1 + ... + r_k$ having the block diagonal form $\mathbf{dg}(\mathbf{A}_1, ..., \mathbf{A}_k)$ where $\mathbf{A}_1, ..., \mathbf{A}_k$ are of order $\mathbf{r}_1, ..., \mathbf{r}_k$ respectively.

Example 2.1.1. The type of matrix is not defined uniquely. Thus the matrix

$$\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

may equally well be said to be of type (2,1) or of type (3).

Theorem 2.1.1. *Let $\lambda_1, ..., \lambda_k$ be distinct numbers, and write $r_1 + ... + r_k = n$. Then an $n \times n$ matrix commutes with*

$$\boldsymbol{D}\text{=}\boldsymbol{dg}(\lambda_1 I_{r_1}, ..., \lambda_k I_{r_k})$$

if and only if it is of type $(r_1, ..., r_k)$.

Proof. $(\Rightarrow)$ Let $\mathbf{AD} = \mathbf{DA}$ and write A in the partitioned form

$$A = \begin{pmatrix} A^{(11)} & A^{(12)} & \cdots & A^{(1k)} \\ A^{(21)} & A^{(22)} & \cdots & A^{(2k)} \\ \vdots & \vdots & \ddots & \vdots \\ A^{(k1)} & A^{(k2)} & \cdots & A^{(kk)} \end{pmatrix}, \ \ and \ \ let \ \ D = \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{r_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k I_{r_k} \end{pmatrix},$$

where $A^{(ij)}$is an $r_i \times r_j$ matrix. Then

$$AD = \begin{pmatrix} A^{(11)} & A^{(12)} & \cdots & A^{(1k)} \\ A^{(21)} & A^{(22)} & \cdots & A^{(2k)} \\ \vdots & \vdots & \ddots & \vdots \\ A^{(k1)} & A^{(k2)} & \cdots & A^{(kk)} \end{pmatrix} \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{r_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k I_{r_k} \end{pmatrix},$$

Then

$$AD = \begin{pmatrix} \lambda_1 A^{(11)} & \lambda_2 A^{(12)} & \cdots & \lambda_k A^{(1k)} \\ \lambda_1 A^{(21)} & \lambda_2 A^{(22)} & \cdots & \lambda_k A^{(2k)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1 A^{(k1)} & \lambda_2 A^{(k2)} & \cdots & \lambda_k A^{(kk)} \end{pmatrix}$$

$$DA = \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{r_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k I_{r_k} \end{pmatrix} \begin{pmatrix} A^{(11)} & A^{(12)} & \cdots & A^{(1k)} \\ A^{(21)} & A^{(22)} & \cdots & A^{(2k)} \\ \vdots & \vdots & \ddots & \vdots \\ A^{(k1)} & A^{(k2)} & \cdots & A^{(kk)} \end{pmatrix},$$

Then

$$DA = \begin{pmatrix} \lambda_1 A^{(11)} & \lambda_1 A^{(12)} & \cdots & \lambda_1 A^{(1k)} \\ \lambda_2 A^{(21)} & \lambda_2 A^{(22)} & \cdots & \lambda_2 A^{(2k)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_k A^{(k1)} & \lambda_k A^{(k2)} & \cdots & \lambda_k A^{(kk)} \end{pmatrix},$$

and therefore $\lambda_i \mathbf{A}^{(ij)} = \lambda_j \mathbf{A}^{(ij)}$ $(i, j = 1, ..., k)$. This implies that $\mathbf{A}^{(ij)} = \mathbf{O}$ when $i \neq j$, then

$$A = \begin{pmatrix} A^{(11)} & 0 & \cdots & 0 \\ 0 & A^{(22)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A^{(kk)} \end{pmatrix},$$

and thus $\mathbf{A}$ is of type $(r_1, ..., r_k)$.

$(\Leftarrow)$ Conversely, let $\mathbf{A}$ is of type $(r_1, ..., r_k)$, then

$$A = \begin{pmatrix} A^{(11)} & 0 & \cdots & 0 \\ 0 & A^{(22)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A^{(kk)} \end{pmatrix}, \ \ and \ \ let \ \ D = \begin{pmatrix} \lambda_1 I_{r_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{r_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_k I_{r_k} \end{pmatrix},$$

then we have

$$AD = \begin{pmatrix} A^{(11)}\lambda_1 & 0 & \cdots & 0 \\ 0 & A^{(22)}\lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A^{(kk)}\lambda_k \end{pmatrix} = \begin{pmatrix} \lambda_1 A^{(11)} & 0 & \cdots & 0 \\ 0 & \lambda_2 A^{(22)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k A^{(kk)} \end{pmatrix} = DA,$$

then $\mathbf{A}$ of type $(r_1, ..., r_k)$ commutes with $\mathbf{D}$. $\square$

Theorem 2.1.2. ***(Rank-multiplicity theorem)***

For every $n \times n$ matrix $\boldsymbol{A}$ and every number w we have

$$dim\boldsymbol{R}(w\boldsymbol{I} - \boldsymbol{A}) \geq n - m_w(\boldsymbol{A}).$$

Proof. By (Dimension Theorem) we have

$$n - dim\mathbf{R}(w\mathbf{I} - \mathbf{A}) = dimN(w\mathbf{I} - \mathbf{A}), \tag{2.1}$$

and by theorem(1.1.3) we have

$$dimN(w\mathbf{I} - \mathbf{A}) \leq m_w(\mathbf{A}). \tag{2.2}$$

Then, form (2.1) and (2.2) we have

$$dim\mathbf{R}(w\mathbf{I} - \mathbf{A}) \geq n - m_w(\mathbf{A}).$$

□

Theorem 2.1.3. *If a matrix $\boldsymbol{A}$ of type $(r_1, ..., r_k)$ is similar to a diagonal matrix, then there exists a matrix $\boldsymbol{S}$ of type $(r_1, ..., r_k)$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is diagonal.*

Proof. Write $r_1 + ... + r_k = n$ and $\mathbf{A} = \mathbf{dg}(\mathbf{A}_1, ..., \mathbf{A}_k)$ where $\mathbf{A}_1, ..., \mathbf{A}_k$ are of order $\mathbf{r}_1, ..., \mathbf{r}_k$ respectively. If $\mathbf{X}$ is any $p \times p$ matrix and w any number, put

$$f_w(\mathbf{X}) = dim\mathbf{R}(w\mathbf{I}_p - \mathbf{X}) + m_w(\mathbf{X}) - p$$

By the rank-multiplicity theorem, we have, for all $\mathbf{X}$ and w,

$$f_w(\mathbf{X}) \geq 0$$

Moreover, by theorem (1.1.4) ,

$$f_w(\mathbf{X}) = 0 \quad (for\ all\ w)$$

if and only if $\mathbf{X}$ is similar to a diagonal matrix. Now clearly

$$f_w(\mathbf{A}) = \sum_{i=1}^{k} f_w(\mathbf{A}_i),$$

and so

$$\sum_{i=1}^{k} f_w(\mathbf{A}_i) = 0$$

For all w by hypothesis. Hence $f_w(\mathbf{A}_i) = 0$ for $i = 1, ..., k$ and all w. Each matrix $\mathbf{A}_i$ is therefore similar to a diagonal matrix. For $i = 1, ..., k$ let $\mathbf{S}_i$ be a non-singular matrix and $\mathbf{D}_i$ a diagonal matrix , both of order $\mathbf{r}_i$, such that $\mathbf{S}_i^{-1}\mathbf{A}_i\mathbf{S}_i = \mathbf{D}_i$. writing $\mathbf{S} = \mathbf{dg}(\mathbf{S}_1, ..., \mathbf{S}_k)$, we obtain at once

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{dg}(\mathbf{D}_1, ..., \mathbf{D}_k),$$

and the theorem therefore proved.

□

The following is the main theorem of this chapter.

Theorem 2.1.4. *Two matrices are simultaneously similar to diagonal matrices if and only if they commute and each is similar to a diagonal matrix.*

Proof. ($\Rightarrow$) Let **A, B** be given matrices. If there exists a matrix $\mathbf{S}$ such that $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, $\mathbf{S}^{-1}\mathbf{B}\mathbf{S}$ are both diagonal, then these two matrices commute and therefore $\mathbf{A}$ and $\mathbf{B}$ are commute. Let we explain this part of proof. Let $D_1 = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ and let $D_2 = \mathbf{S}^{-1}\mathbf{B}\mathbf{S}$. As we know $D_1D_2 = D_2D_1$, then L.H.S=$D_1D_2 = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}\mathbf{S}^{-1}\mathbf{B}\mathbf{S} = \mathbf{S}^{-1}\mathbf{A}\mathbf{B}\mathbf{S}$ and R.H.S=$D_2D_1 = \mathbf{S}^{-1}\mathbf{B}\mathbf{S}\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{S}^{-1}\mathbf{B}\mathbf{A}\mathbf{S}$. Since $D_1D_2 = D_2D_1$, then $\mathbf{S}^{-1}\mathbf{A}\mathbf{B}\mathbf{S} = \mathbf{S}^{-1}\mathbf{B}\mathbf{A}\mathbf{S}$, so $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$.

($\Leftarrow$) Suppose, on the other hand, that $\mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{A}$ and that $\mathbf{A}$ and $\mathbf{B}$ are both similar to diagonal matrices. Let $\lambda_1, ..., \lambda_k$ be the distinct eigenvalues of $\mathbf{A}$ and let their multiplicities be $r_1, ..., r_k$ respectively. There exists , then, a matrix $\mathbf{P}$ such that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}\mathbf{=}\mathbf{dg(}\lambda_1 I_{r_1}, ..., \lambda_k I_{r_k}\ )$$

Now, in view of our hypothesis, $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ commutes with $\mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ and hence, by Theorem (2.1.1), $\mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ is of type $(r_1, ..., r_k)$. Since $\mathbf{B}$ is similar to a diagonal matrix; therefore, by Theorem (2.1.3) there exists a matrix $\mathbf{Q}$, of type $(r_1, ..., r_k)$ such that $\mathbf{Q}^{-1}\mathbf{P}^{-1}\mathbf{B}\mathbf{P}\mathbf{Q}$ is a diagonal. Moreover, again by Theorem (2.1.1), $\mathbf{Q}$ commutes with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, and therefore

$$\mathbf{Q}^{-1}\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{Q} = \mathbf{D}\mathbf{=}\mathbf{dg(}\lambda_1 I_{r_1}, ..., \lambda_k I_{r_k}\ )$$

Thus $(\mathbf{PQ})^{-1}\mathbf{A}(\mathbf{PQ})$ and $(\mathbf{PQ})^{-1}\mathbf{B}(\mathbf{PQ})$ are both diagonal, and the theorem is proved. $\square$

CHAPTER 3

THE TRIANGULARIZATION

3.1 TRIANGULARIZATION AND REDUCIBILITY

There are many known sufficient conditions that a collection of linear transformations be triangularizable. An important preliminary result is Burnside's Theorem on existence of invariant subspaces for algebras of linear transformation and Burnside's Theorem says that the only irreducible algebra of linear transformations on the finite-dimensional vector space $\mathcal{V}$ of dimension greater than 1 is the algebra of all linear transformations mapping $\mathcal{V}$ into $\mathcal{V}$. This chapter is adapted from [5].

Throughout this chapter we restrict our attention to collection of linear transformations on a finite-dimensional vector space over $\mathbb{C}$.

Definition. A subspace $\mathcal{W}$ is *invariant* for a collection C of linear transformations if $Ax \in \mathcal{W}$ whenever $x \in \mathcal{W}$ and $A \in C$. A subspace is *nontrivial* if it is different from $\{0\}$ and from the entire space. A collection of linear transformations is *reducible* if it has a nontrivial invariant subspace and is *irreducible* otherwise.

The central definition is the following.

Definition. A collection of linear transformations is *triangularizable* if there is a basis for the vector space such that all transformations in the collection have upper triangular matrix representations with respect to that basis.

It is clear that triangularizablility is equivalent to the existence of a chain of invariant subspaces

$$\{0\} = \mathcal{W}_0 \subset \mathcal{W}_1 \subset \mathcal{W}_2 \subset ... \subset \mathcal{W}_n = \mathcal{V},$$

with dimension of $\mathcal{W}_j$ equal to j for each j and with $\mathcal{V}$ the entire vector space. Namely, if the collection is triangularizable with respect to the basis $\{e_1, e_2, ..., e_n\}$. Let $\mathcal{W}_j$ be the linear span of $\{e_1, e_2, ..., e_j\}$ for each j. Any such chain is called a *triangularizing chain* for the collection.

Quotient spaces will play an important role in this study.

Definition. : Let $\mathcal{V}$ be a vector space and T a linear transformation

$$T : \mathcal{V} \longrightarrow \mathcal{V}.$$

Suppose N is a T-invariant subspace (i.e if $n \in N$ then $T(n) \in N$). The *quotient space*: $\mathcal{V}/N = \{v+N : v \in \mathcal{V}\}$. Define $\hat{T} : \mathcal{V}/N \longrightarrow \mathcal{V}/N$ by $\hat{T}(v+N) = T(v)+N \ \ for\ all\ v \in \mathcal{V}$. This is well defined : if $v_1 + N = v_2 + N$ need to show $\hat{T}(v_1 + N) = \hat{T}(v_2 + N)$. Let $v_1+N = v_2+N$, then $v_1-v_2 \in N$. Since N is a T-invariant subspace, then $T(v_1-v_2) \in N$. And since T is a linear transformation, then $T(v_1)-T(v_2) \in N$, so $T(v_1)+N = T(v_2)+N$, then

$$\hat{T}(v_1 + N) = \hat{T}(v_2 + N)$$

is well defined.

Definition. : A property P is *inherited by quotients* if P is true for $(T, \mathcal{V})$, then P is also true for $(\hat{T}, \hat{\mathcal{V}})$.

Theorem 3.1.1. ***(The Triangularization Lemma)*** *let P be a set of properties, each of which is inherited by quotients. If every collection of transformations on a space of dimension greater than 1 that satisfies P is reducible, then every collection of transformations satisfying P is triangularizable.*

Proof. The proof is by induction on dim$\mathcal{V}$. If dim$\mathcal{V}$=1 then the chain $\{0\} = \mathcal{V}_0 \subset \mathcal{V}_1 = \mathcal{V}$ satisfies condition for triangularizability. Namely, $\mathcal{V}_0$ and $\mathcal{V}_1$ are C-invariant and dim$\mathcal{V}_i = i$.

Suppose $dim\mathcal{V} > 1$. Choose a maximal chain of C-invariant subspaces of $\mathcal{V}$ (since C is a collection of T's).

$$\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \mathcal{M}_2 \subset ... \subset \mathcal{M}_n = \mathcal{V},$$

We need to show $\dim \mathcal{M}_k/\mathcal{M}_{k-1} = 1$ for all k (since the $\dim \mathcal{M}_i = i$). Suppose instead that for some k, $\dim \mathcal{M}_k/\mathcal{M}_{k-1} > 1$. Each $T \in C$ induces $T_0 : \mathcal{M}_k \longrightarrow \mathcal{M}_k$ and $\hat{T}_0 : \mathcal{M}_k/\mathcal{M}_{k-1} \longrightarrow \mathcal{M}_k/\mathcal{M}_{k-1}$ as $\mathcal{M}_k$ and $\mathcal{M}_{k-1}$ are C-invariant. Let C_0 be the collection of T_0's and $\hat{C}_0$ the collection of $\hat{T}_0$'s). Now, P is true for $(\hat{C}_0, \mathcal{M}_k/\mathcal{M}_{k-1})$ by hypothesis. So $\mathcal{M}_k/\mathcal{M}_{k-1}$ is reducible, by hypothesis. So there exists a non-trivial $\hat{C}_0$-invariant subspace $L = \{x + M_{k-1} : x \in M_k\}$. Define $L^* = \{x \in M_k : x + M_{k-1} \in L\}$.

Claim a: L^* is T-invariant.

Pick $x \in L^*$ (we want to show that $T(x) \in L^*$). Now, $x + \mathcal{M}_{k-1} \in L$. Since $\mathcal{M}_k$, $\mathcal{M}_{k-1}$ are T-invariant subspace and L is $\hat{T}_0$-invariant subspace, so $T(x) \in M_k$ such that $\hat{T}_0(x + M_{k-1}) \in L$, and $T(x) \in M_k$ such that $T(x) + M_{k-1} \in L$, then $T(x) \in L^*$, so L^* is T-invariant.

Claim b: $\mathcal{M}_{k-1} \subsetneq L^* \subsetneq \mathcal{M}_k$.

First, we want to prove claim 1, claim 2, claim 3, and claim 4 to prove $\mathcal{M}_{k-1} \subsetneq L^* \subsetneq \mathcal{M}_k$.

Claim 1: $\mathcal{M}_{k-1} \subset L^*$.

Let $x \in \mathcal{M}_{k-1}$, then $x + \mathcal{M}_{k-1} = 0 + \mathcal{M}_{k-1} \in L$, since L is subspace of $\mathcal{M}_k/\mathcal{M}_{k-1}$, thus $x \in L^*$.

Claim 2: $\mathcal{M}_{k-1} \neq L^*$.

We want to prove that there exists $x \in L^*$, but $x \notin \mathcal{M}_{k-1}$, Since $L \neq 0$, so there exists $x + \mathcal{M}_{k-1} \neq 0 + \mathcal{M}_{k-1}$, and $x + \mathcal{M}_{k-1} \in L$. Then $x \notin \mathcal{M}_{k-1}$, and $x \in L^*$.

Claim 3: $L^* \subset \mathcal{M}_k$.

Let $x \in L^*$, then $x \in \mathcal{M}_k$ (by definition of L^*).

Claim 4: $L^* \neq \mathcal{M}_k$.

We want to prove that there exists y $\in \mathcal{M}_k$, but $y \notin L^*$. Since L is nontrivial invariant subspace of $\mathcal{M}_k/\mathcal{M}_{k-1}$, so $L \neq \mathcal{M}_k/\mathcal{M}_{k-1}$. Then there exists $y + \mathcal{M}_{k-1} \in \mathcal{M}_k/\mathcal{M}_{k-1}$, and $y + \mathcal{M}_{k-1} \notin L$. Then $y \in \mathcal{M}_k$, but $y \notin L^*$.

From claim 1, claim 2, claim 3, and claim 4, then we have claim b (i.e $\mathcal{M}_{k-1} \subsetneq L^* \subsetneq \mathcal{M}_k$), but this contradicts with maximality of the chain $\{\mathcal{M}_j\}$. So, $\dim \mathcal{M}_k/\mathcal{M}_{k-1} = 1$ for all k . Then

$$\{0\} = \mathcal{M}_0 \subset \mathcal{M}_1 \subset \mathcal{M}_2 \subset ... \subset \mathcal{M}_n = \mathcal{V},$$

satisfies condition for triangularizability. □

Theorem 3.1.2. *Every commutative collection of linear transformations is triangularizable.*

Proof. Let C be a collection of commuting T. We check that commutativity is inherited by quotient. Say $W \subset \mathcal{V}$ is a invariant subspace under all $T \in C$. By the definition of a quotient space

$$\hat{T} : \mathcal{V}/W \longrightarrow \mathcal{V}/W$$

$$\hat{T}(v + W) = T(v) + W \quad for\ all\ v \in \mathcal{V}$$

We want to prove that $\hat{T}_1\hat{T}_2$=$\hat{T}_2\hat{T}_1$

$$\begin{aligned}
\hat{T}_1\hat{T}_2(v + W) &= \hat{T}_1(T_2(v) + W) \\
&= T_1T_2(v) + W \\
&= T_2T_1(v) + W \\
&= \hat{T}_2(T_1(v) + W) \\
&= \hat{T}_2\hat{T}_1(v + W)
\end{aligned}$$

Then $\hat{T}_1\hat{T}_2$=$\hat{T}_2\hat{T}_1$

Our main goal is proving that C is reducible, this means we want to prove that there exists a non-trivial invariant subspace. First, we need to prove this if all T's are a multiple of the identity. Second, if there exists $T_1 \in C$, and T_1 is not multiple of identity, then we want to to prove that there exists W which is invariant under C.

Case 1: All T's are a multiple of the identity

Pick any $W \subset V$. If $w \in W$, then $T(w) = \lambda w \in W$. Hence $\mathcal{W}$ is invariant, so every subspace is invariant.

Case2: There exists $T_1 \in C$, and T_1 is not a multiple of identity.

Let λ be any eigenvalue of T_1 and let W be corresponding eigenspace. If $T_2 \in C$, and $x \in W$, then $T_1T_2(x) = T_2T_1(x) = T_2(\lambda x) = \lambda T_2(x)$,

so $T_2(x)$ is an eigenvector for T_1, and so $T_2(x) \in W$. Then W is invariant under C.

It is non-trivial for two reasons :

(a) λ is an eigenvalue so there is eigenvector and so $W \neq 0$.

(b) If $W = V$, Then $T_1\lambda = \lambda v$, for all $v \in W$, So T_1 is multiple of identity.

This is contradiction, then $W \neq V$.

From Case 1, Case 2, (a) and (b), we conclude that W is a nontrivial invariant subspace under C, so C is reducible. Then by Theorem (3.1.1) C is triangularizable. □

Definition. : A *noncommutative polynomial* in the linear transformations $\{A_1, A_2, ..., A_k\}$ is any linear combination of words in the transformations.

We use $\sigma(A)$ to denote the spectrum (which in the present, finite-dimensional, case is simply the set of eigenvalues) of A.

Theorem 3.1.3. ***(Spectral Mapping Theorem)***

Suppose $\{A_1, ..., A_k\}$ is a triangularizable collection of linear transformations, and if p is any noncommutative polynomial in $\{A_1, ..., A_k\}$, then

$$\sigma(p(A_1, ..., A_k)) \subset p(\sigma(A_1), ..., \sigma(A_k)),$$

where $p(\sigma(A_1), ..., \sigma(A_k))$ denotes the set of all $p(\lambda_1, ..., \lambda_k)$ such that $\lambda_j \in \sigma(A_j)$ for all j.

Proof. . This follows immediately from the facts that

(i) the eigenvalues of triangular matrices are the entries on the main diagonal,

(ii) each of the diagonal entries of a product of given matrices is a product of diagonal entries, and

(iii) each of the diagonal entries of a sum of given matrices is a sum of diagonal entries of the given matrices. $\square$

Example 3.1.1. Let

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \ and \ B = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix},$$

then

$$AB = \begin{pmatrix} 1 & 10 \\ 0 & 12 \end{pmatrix}, \ and \ A + B = \begin{pmatrix} 2 & 4 \\ 0 & 7 \end{pmatrix},$$

then we have $\sigma(AB) = \{1, 12\} \subset \sigma(A)\sigma(B) = \{1, 3\}.\{1, 4\} = \{1, 3, 4, 12\}$ and also we have $\sigma(A + B) = \{2, 7\} \subset \sigma(A) + \sigma(B) = \{1, 3\} + \{1, 4\} = \{2, 4, 5, 7\}$.

Theorem 3.1.4. *Every linear transformation on a finite-dimensional space is the sum of transformations of rank 1.*

Proof. Recall rank(T)=1 means dim$R(T)$=1. Also if linear

$$T : \mathcal{V} \longrightarrow \mathbb{C}$$

then the matrix of T is a row $1 \times n$ matrix. Pick any linear

$$S : \mathcal{V} \longrightarrow \mathcal{V}$$

and pick basis B then $[S]_B$: is a $n \times n$ matrix.

$$[S]_B = \begin{pmatrix} r_1 & \cdots \\ 0 & \cdots \\ \vdots & \vdots \\ 0 & \cdots \end{pmatrix} + \begin{pmatrix} 0 & \cdots \\ r_2 & \cdots \\ \vdots & \vdots \\ 0 & \cdots \end{pmatrix} + \cdots + \begin{pmatrix} 0 & \cdots \\ 0 & \cdots \\ \vdots & \vdots \\ r_n & \cdots \end{pmatrix} \quad all \ \ rank \ \ 1$$

where r_i is the i^{th} row of $[S]_B$, so the sum of matrices is rank 1.

$\square$

Definition. An *algebra* of linear transformations is a collection of linear transformations that is closed under addition, multiplication, and multiplication by scalars. An algebra is *unital* if it contains the identity transformation. In a unital algebra with identity I, we use the notation λ is an abbreviation for λI. The notation $\mathcal{B}(\mathcal{V})$ is used to denote the algebra of all linear transformations mapping $\mathcal{V}$ into $\mathcal{V}$. (The notation $\mathcal{L}(\mathcal{V})$ is also common.)

If $\mathcal{A}$ is an algebra of linear transformations and x is any given vector, then $\{Ax : A \in \mathcal{A}\}$ is easily seen to be an invariant subspace for $\mathcal{A}$. However, for some $\mathcal{A}$ and x, $\{Ax : A \in \mathcal{A}\}$ is the entire space (in which case x is said to be a *cycle vector* for $\mathcal{A}$).

Theorem 3.1.5. ***(Burnside's Theorem)*** *The only irreducible algebra of linear transformations on the finite-dimensional vector space $\mathcal{V}$ of dimension greater than 1 is the algebra of all linear transformations mapping $\mathcal{V}$ into $\mathcal{V}$.*

Proof. Let $\mathcal{A}$ be an irreducible algebra. We first show that $\mathcal{A}$ contains a transformation of rank 1. For this, let T_0 be a transformation in $\mathcal{A}$ with minimal nonzero rank. We must show that this rank is 1. If rank $(T_0) >1$, then $\dim R(T_0)>1$. So, there would be vectors x_1 and x_2 such that $\{T_0x_1, T_0x_2\}$ is linearly independent set.
claim: $\{AT_0(x_1) : A \in \mathcal{A}\} = \mathcal{V}$

Set $\mathcal{W} = \{AT_0(x_1) : A \in \mathcal{A}\}$. We want to check that $\mathcal{W}$ is invariant under $\mathcal{A}$. Pick $AT_0(x_1) \in \mathcal{W}$, and pick $\hat{A} \in \mathcal{A}$, then $\hat{A}(AT_0(x_1)) = \hat{A}AT_0(x_1) = (\hat{A}A)T_0(x_1)$,
since $\mathcal{A}$ is an algebra, then $\hat{A}A \in \mathcal{A}$, so $(\hat{A}A)T_0(x_1) \in \mathcal{W}$. Moreover, $\mathcal{W} \neq 0$ because $x_1 \neq 0$. And as we know that $\mathcal{A}$ is irreducible, then only invariant subspace is 0 or $\mathcal{V}$. Therefore, $\mathcal{W} = \mathcal{V}$ which proves the claim. Since $\{AT_0(x_1) : A \in \mathcal{A}\} = \mathcal{V}$, then there is $A_0 \in \mathcal{A}$ such that $A_0T_0x_1 = x_2$. Then $\{T_0A_0T_0x_1, T_0x_1\}$ is linearly independent and $T_0A_0T_0 - \lambda T_0 \neq 0$

for all scalars λ (since if $T_0A_0T_0 - \lambda T_0 = 0$ then multiplying by x_1 would give $T_0A_0T_0x_1$ and T_0x_1 dependent). Let we take the restriction

$$T_0A_0|_{T_0(v)} : T_0(\mathcal{V}) \longrightarrow T_0(\mathcal{V})$$

Then $T_0A_0(T_0v) = T_0(A_0T_0v) \in T_0(\mathcal{V})$

let λ_0 be an eigenvalue of $T_0A_0|_{T_0(v)}$, then there exists $0 \neq z \in T_0(\mathcal{V})$ such that $T_0A_0(z) = \lambda_0 z$, and $(T_0A_0|_{T_0(v)} - \lambda_0 I)(z) = 0$. Moreover, $T_0A_0|_{T_0(v)} - \lambda_0 I$ is not invertible. We want to explain this part of proof briefly

$$\mathcal{V} \xrightarrow{T_0} T_0(\mathcal{V}) \xrightarrow[\textit{is not invertible}]{T_0A_0|_{T_0(v)} - \lambda_0 I} T_0(\mathcal{V})$$

Then we have $R(T_0A_0|_{T_0(\mathcal{V})} - \lambda_0 I) \subsetneq T_0(\mathcal{V})$, so $R((T_0A_0 - \lambda_0 I)T_0) \subsetneq T_0(\mathcal{V})$. Thus $rank((T_0A_0 - \lambda_0 I)T_0) < rank T_0$ .This contradicts the minimality of the rank of T_0, then we conclude that T_0 has rank 1 (i.e $\dim(T_0) = 1$).

Pick a nonzero vector $y_0 \in R(T_0)$. Since $R(T_0) = \{\alpha y_0 : \alpha \in \mathbb{C}\}$, For any $x \in \mathcal{V}$ and $T_0(x) \in R(T_0)$ we have

$$T_0(x) = \alpha y_0 \ \ for \ \ some \ \ \alpha \in \mathbb{C}$$

Define

$$\phi_0 : \mathcal{V} \longrightarrow \mathbb{C}$$

by

$$\phi_0(x) = \alpha, \quad \alpha \in \mathbb{C}$$

We want to prove that ϕ_0 is linear. Now, $\phi_0(x+y)$ satisfies

$$T_0(x+y) = \phi_0(x+y)y_0$$

$$L.H.S = T_0(x) + T_0(y)$$

$$= \phi_0(x)y_0 + \phi_0(y)y_0$$

$$= [\phi_0(x) + \phi_0(y)]y_0$$

$$= \phi_0(x+y)y_0 = R.H.S$$

$\Rightarrow \phi_0$ is linear. so

$$T_0(x) = \phi_0(x)y_0 \tag{3.1}$$

Since every linear transformation of rank 1 has the form $x \to \phi(x)y$ for a vector y in $\mathcal{V}$ and linear functional ϕ, and by Theorem (3.1.4) every linear transformation on a finite-dimensional space is the sum of transformations of rank 1, it suffices to show that $\mathcal{A}$ contains every T of the form $T(x) = \phi(x)y$.

Let $\mathcal{F} = \{\phi \in \mathcal{V}^* : \text{if } S(x) = \phi(x)y_0 \text{ then } S \in \mathcal{A}\} \subset \mathcal{V}^*$. We want to prove that $\mathcal{F} = \mathcal{V}^*$. Suppose $\mathcal{F} \neq \mathcal{V}^*$. Pick $\psi \in \mathcal{V}^*$, $\psi \notin \mathcal{F}$, and pick a basis of $\mathcal{F}$: $\psi_2, ..., \psi_s$. Now, we have $\psi \notin \mathcal{F}$ and $\mathcal{F} = span\{\psi_2, ..., \psi_s\}$, then by theorem (1.1.5) $\psi, \psi_2, ..., \psi_s$ are independent and extend this to a basis of $\mathcal{V}^*$: $\psi, \psi_2, ..., \psi_s, \psi_{s+1}, ..., \psi_n$

Define

$$S : \mathcal{V}^* \longrightarrow \mathbb{C}$$

by

$$S(a\psi + a_2\psi_2 + ... + a_n\psi_n) = a$$

Since $\mathcal{V}$ is reflexive by Theorem (1.1.1), then $S = S_{x_0}$ for some x_0. Let $\phi \in \mathcal{F}$, then $\phi = a_2\psi_2 + ... + a_s\psi_s$ and we get $S(\phi) = S(a_2\psi_2 + ... + a_s\psi_s) = 0$. Since $S_{x_0}(\phi) = \phi(x_0)$, and $S(\psi) = 1$, and $S(\psi) = \psi(x_0)$, then we get $x_0 \neq 0$ with $\phi(x_0) = 0$ for all $\phi \in \mathcal{F}$.

We want to summarize what we got from previous paragraph. We assumed $\mathcal{F} \neq \mathcal{V}^*$ (in a proof by contradiction), and we got $x_0 \neq 0$, $\phi(x_0) = 0$ all $\phi \in \mathcal{F}$. $\mathcal{F}$ contains $\phi(x) = \phi_0(Ax)$ and all $A \in \mathcal{A}$. Now, we want to explain why $\mathcal{F}$ contains $\phi(x) = \phi_0(Ax)$ and all $A \in \mathcal{A}$. So, we need to show that if $S(x) = \phi(x)y_0$ then $S \in \mathcal{A}$. Let we say that $S(x) = \phi(x)y_0$, then we have $S(x) = \phi_0(Ax)y_0$ and we know that $T_0(x) = \phi_0(x)y_0$ and $T_0 \in \mathcal{A}$ by equation (3.1). And let $\mathcal{A}$ is algebra and $A \in \mathcal{A}$, then we have $T_0A \in \mathcal{A}$, so $S(x) = \phi_0(Ax)y_0 = T_0(Ax) = (T_0A)x$. Thus $S = T_0A \in \mathcal{A}$. Then we concluded $\phi_0(Ax_0) = 0$ and all $A \in \mathcal{A}$.

Since $T_0(x_3)$ is non-zero for some $x_3 \in \mathcal{V}$ and $T_0(x_3) = \phi_0(x_3)y$, so $\phi_0(x_3)$ is non-zero for some $x_3 \in \mathcal{V}$. Since $\{Ax_0 : A \in \mathcal{A}\} = \mathcal{V}$ by claim, and $x_3 = Ax_0$ for some $A \in \mathcal{A}$, then $0 \neq \phi_0(x_3) = \phi_0(Ax_0)$. This contradicts with $\phi_0(Ax_0) = 0$ all $A \in \mathcal{A}$. Hence $\mathcal{F} = \mathcal{V}^*$. Now, we have

(1) $y_0 \neq 0$ [because $0 \neq T_0x_1 = \phi_0(x_1)y_0$]

(2) Claim: $Z = \{Ay_0 : A \in \mathcal{A}\}$ is invariant subspace. We want to prove $\hat{A}(Ay_0) \in Z$. let we pick $Ay_0 \in Z$, and pick $\hat{A} \in \mathcal{A}$

$\hat{A}(Ay_0) = (\hat{A}A)y_0 \in Z$. Since $\hat{A}A \in \mathcal{A}$, $Z \neq 0$ by (1), and $\mathcal{V}$ is irreducible. So $Z = \mathcal{V}$

(3)

$$AT : \mathcal{V} \longrightarrow \mathbb{C}$$

So $AT \in \mathcal{V}^*$

From the previous paragraph, $\mathcal{V}^* = \mathcal{F}$. Now, our object is proving that $\mathcal{A}$ contains all rank-one transformations. Let we pick T a rank-one transformation, then there exists $y_1 \neq 0$, and $\phi \in \mathcal{V}^*$ with $T(x) = \phi(x)y_1$. We want to prove $T \in \mathcal{A}$. We have $\mathcal{F} = \{\phi \in \mathcal{V}^*\colon$ if $S(x) = \phi(x)y_0$ then $S \in \mathcal{A}\}$, and we have $\mathcal{F} = \mathcal{V}^*$ from previous paragraph. Let $y_1 = Ay_0$ by (2) for some $A \in \mathcal{A}$, then $T(x) = \phi(x)Ay_0$.

Define

$$\hat{T} = \phi(x)y_0$$

$A\hat{T}(x) = A\phi(x)y_0 = \phi(x)Ay_0 = \phi(x)y_1$.

Then $A\hat{T} = T$, and we have $\phi \in \mathcal{V}^* = \mathcal{F}$, so $\hat{T} \in \mathcal{A}$ and $A \in \mathcal{A}$, and $T = A\hat{T} \in \mathcal{A}$. Then $\mathcal{A}$ contains all rank-one transformations. Thus $\mathcal{A}$ contains all transformations by Theorem (3.1.4).

□

Theorem 3.1.6. ***(McCoy's Theorem)***

The pair {A,B} is triangularizable if and only if p(A,B)(AB-BA) is nilpotent for every noncommutative polynomial p.

Proof. (⇒)

If $\{A, B\}$ is triangularizable, then $\sigma(p(A, B)(AB - BA)) \subset p(\sigma(A), \sigma(B))(\sigma(AB) - \sigma(BA))$ by the Spectral Mapping Theorem. Let $\lambda_1, \lambda_2, ..., \lambda_n$ are eigenvalues of A, and let $\mu_1, \mu_2, ..., \mu_n$ are eigenvalues of B.

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}\begin{pmatrix} \mu_1 & * & \cdots & * \\ 0 & \mu_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix} - \begin{pmatrix} \mu_1 & * & \cdots & * \\ 0 & \mu_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{pmatrix}\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1\mu_1 & * & \cdots & * \\ 0 & \lambda_2\mu_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\mu_n \end{pmatrix} - \begin{pmatrix} \mu_1\lambda_1 & * & \cdots & * \\ 0 & \mu_2\lambda_2 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n\lambda_n \end{pmatrix} = \begin{pmatrix} 0 & * & \cdots & * \\ 0 & 0 & * & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

we have $\sigma(p(A, B)(AB - BA)) \subset \{0\}$, then we have $\sigma(p(A, B)(AB - BA)) = \{0\}$. So $p(A, B)(AB - BA)$ is nilpotent.

($\Leftarrow$) If $(AB - BA)v = 0$ for all v, then $AB - BA = 0$, and $AB = BA$, so the algebra $\mathcal{A}$ generated by A and B is triangularizable by Theorem (3.1.2). Suppose for some v such that $(AB - BA)v \neq 0$, call it $w = (AB - BA)v$, there exists C such that $Cw = v$. Namely, pick bases $w_1, w_2, \cdots, w_n$ of W and pick bases $v_1, v_2, \cdots, v_n$ of V. Define T by $T(\sum a_i w_i) = \sum a_i v_i$. Let C be a matrix for T. If $\mathcal{A}$ is irreducible, then by Burnside's Theorem, $C \in \mathcal{A}$, $C = p(A, B)$ for some p. Let $D = C(AB - BA)$, and we have

$$Dv = v \neq 0$$

$$D^2 v = D(Dv) = Dv = v$$

$$D^3 v = D(D^2 v) = Dv = v$$

$$\vdots$$

$$D^k v = v,$$

so no $D^k = 0$,then D is not nilpotent. So $\mathcal{A}$ is reducible and by The Triangularization Lemma then $\mathcal{A}$ is triangularizble. Hence the pair {A,B} is triangularizable. $\square$

Example 3.1.2. Let

$$A = \begin{pmatrix} -11 & 6 \\ -28 & 15 \end{pmatrix}, \ and \ B = \begin{pmatrix} -14 & 9 \\ -40 & 24 \end{pmatrix},$$

A and B are triangularizble since they have common eigenvector$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and let $p(A, B) = A^2 + 2AB + 3BA + 4B^2$

then we have

$$AB - BA = \begin{pmatrix} 12 & -6 \\ 24 & -12 \end{pmatrix}, \ and \ p(A, B) = \begin{pmatrix} -1169 & 627 \\ -2824 & 1497 \end{pmatrix},$$

$$p(A, B)(AB - BA) = \begin{pmatrix} 1020 & -510 \\ 2040 & -1020 \end{pmatrix}, and \ (p(A, B)(AB - BA))^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$\Rightarrow$ (p(A,B)(AB-BA)) is nilpotent with $k = 2$.

CHAPTER 4

LIE ALGEBRAS

4.1 DIAGONAL (TRIANGULAR) MATRICES AND LIE ALGEBRA

There are a strong relation between Lie algebras and simultaneous diagonalization and triangularization of matrices. This chapter shows the relation between Lie algebras and the matrices which are simultaneously similar to diagonal (triangular) matrices. The material on Lie algebras is from [2]. And the material which is connected with McCoy's theorem is already in McCoy's paper [3].

Definition. Let $\mathcal{V}$ be a vector space with product $\mathcal{V} \times \mathcal{V} \longrightarrow \mathcal{V}$. The product is called *bilinear* on $\mathcal{V}$ if

(1)$[\alpha a+b,c]=\alpha[a,c]+[b,c]$

(2)$[a,\alpha b+c]=\alpha[a,b]+[a,c]$,

for all a, b, c $\in \mathcal{V}$ and $\alpha \in \mathbb{C}$.

Definition. A *Lie algebra* is vector space $\mathcal{V}$ with product $\mathcal{V} \times \mathcal{V} \longrightarrow \mathcal{V}$, written as [a,b], such that

(1)The product is bilinear

(2)[b,a]=-[a,b]

(3)[[a,b],c]+[[b,c],a]+[[c,a],b]= 0,

for all a, b, c $\in \mathcal{V}$.

Example 4.1.1. $\mathcal{V}=\mathcal{R}^3$: [a,b]= a $\times$ b (cross product)

Example 4.1.2. If $\mathcal{A}$ is an algebra of matrices then [A,B]= AB - BA. Then from definition (2) is [B,A]= BA - AB = -(AB - BA) = -[A,B] and (3) is [[A,B],C]+[[B,C],A]+[[C,A],B] = (AB-BA)C - C(AB-BA) + (BC-CB)A - A(BC-CB) + (CA-AC)B - B(CA-AC) = ABC - BAC - CAB + CBA + BCA - CBA - ABC + ACB + CAB - ACB - BCA + BAC = 0

Theorem 4.1.1. *The matrices A and B are simultaneously diagonal if and only if [$\mathcal{A}$,$\mathcal{A}$]=0 where $\mathcal{A}$ is the lie algebra generated by A, B and the matrices A, B are diagonalizable.*

Proof. ($\Rightarrow$) Let A, B are simultaneously diagonal, then by theorem (2.1.4) we have AB = BA. So, [A,B]= AB - BA = 0. Thus [$\mathcal{A}$,$\mathcal{A}$]=0.
($\Leftarrow$)If A, B $\in \mathcal{A}$, then we have [A,B]= AB - BA. Since [$\mathcal{A}$,$\mathcal{A}$]= 0. So, [A,B]= AB - BA = 0. Thus AB = BA. By hypothesis, we have A and B are diagonalizable, then by Theorem (2.1.4), then we have that A and B are simultaneously diagonal. ☐

Definition. A Lie algebra $\mathcal{L}$ is *nilpotent* if $\mathcal{L}^k$=0 for some positive integer k.

$\mathcal{L}$ is algebra of linear transformation and $\mathcal{L}^{'}$=[$\mathcal{L}$,$\mathcal{L}$] is a commutator subalgebra.

Definition. Let $\mathcal{L}$ be a Lie algebra. Define

$$\mathcal{L}^{'} = [\mathcal{L}, \mathcal{L}]$$

$$\mathcal{L}^{''} = [\mathcal{L}^{'}, \mathcal{L}^{'}]$$

$$\vdots$$

$$\mathcal{L}^{(k)} = [\mathcal{L}^{k-1}, \mathcal{L}^{k-1}],$$

then $\mathcal{L}$ is *solvable* if $\mathcal{L}^{(k)}$=0 for some positive integer k.

Theorem 4.1.2. ***(Engel's Theorem)***

A Lie algebra $\mathcal{L}$ is nilpotent if and only if every matrix in $\mathcal{L}$ is nilpotent.

Theorem 4.1.3. ***(Lie's theorem)*** *If L is a solvable Lie algebra of a linear transformations in a finite-dimensional vector space $\mathcal{V}$ over $\mathbb{C}$, then the matrices of L can be taken in simultaneously triangular form.*

Theorem 4.1.4. *Let $\mathcal{L}$ be the Lie algebra generated by A and B. Then the following statements are equivalent:*

1) the matrices A and B are simultaneously triangularizable,

2) for any polynomial p(x,y) in the non commuting variables x and y, the matrix p(A,B)(AB-BA) is nilpotent,

3) the Lie algebra $\mathcal{L}$ is a solvable.

Proof. 1)$\Rightarrow$ 2) is proved by McCoy's Theorem. 2) $\Rightarrow$ 3) is proved by Engel's Theorem, $\mathcal{L}' = [\mathcal{L}, \mathcal{L}]$ is nilpotent, so $\mathcal{L}$ is solvable. 3) $\Rightarrow$ 1) is proved by Lie's Theorem. $\square$

REFERENCES

[1] Friedberg, S.H., Insel, A.J. and Spence, L.E., *Linear Algebra*, Pearson Education Inc., New Jersey, 2003.

[2] Jacobson, N., *Lie Algebra*, Dover Publications, Inc., New York 1979.

[3] McCoy, N.H., *On The Characteristic Roots of Matric Polynomials*, Bull. Amer. Math. Soc., 42 (1936).

[4] Mirsky, L., *An Introduction to linear Algebra*, Dover Publications, Inc., New York 1990.

[5] Radjavi, H. and Rosenthal, P., *Simultaneous Triangularization*, Springer-Verlag New York, Inc., New York 2000.

VITA

Graduate School
Southern Illinois University

Hamdan Alsulaimani Date of Birth: May 25, 1985

2010 Evergreen Terrace Dr. W. Apt 4, Carbondale, Illinois 62901

al_rqai@hotmail.com

Kuwait University
Bachelor of Science, Mathematics, May 2008

Research Paper Title:
Diagonal And Triangular Matrices

Major Professor: Dr. R. Fitzgerald