Simultaneous Triangularization of Certain Sets of Matrices

Katie Gedeon
Advisor: Lance Small
Univeristy of California, San Diego
Department of Mathematics

June 6, 2012

Abstract

In an effort to better understand linear transformations, we look at triangularization of matrices. After a discussion of both nilpotent and unipotent matrices, we prove the Lie-Kolchin theorem by considering the nilpotent parts of unipotent matrices.

Acknowledgments

First and foremost, I would like to thank my advisor Lance Small. Without his patience, understanding nature, and guidance, this paper would not have been possible. He has been a great mentor for me this past year, both with this research and in my quest to apply for graduate school. I would also like to thank Dan Rogalski and Cristian Popescu, who instilled in me a deep love for Algebra and helped me find the drive that I needed to undertake difficult mathematics.

Next I would like to thank Jeffrey Rabin, who has helped me develop as a mathematics student over the past four years. I often went to him seeking advice and direction, especially when applying for graduate school. Thank you for always being willing to meet with me and answer my questions.

In general, I would like to thank the Mathematics Department staff at UCSD, especially Jan Bitmead. They are a very helpful group and possess a wealth of information vital to navigating life as an undergraduate math student.

Last, but certainly not least, I would like to thank my friends. I have developed an amazing support system over the past four years and it is no exaggeration when I say I would not be where I am today without it. To all of you, from the friends who I have taken math classes with every quarter to those who have never heard of a nilpotent matrix, but listen to me anyway as I practice my presentation of this thesis, thank you. This paper is dedicated to you.

Contents

1 First Remarks

This paper is essentially a walk through some topics in linear algebra, building up to the Lie-Kolchin Theorem. In particular, we will look at two special kinds of matrices, nilpotent and unipotent. We will heavily depend on the process of triangularization. This process relies upon a basic understanding of eigenvalues and eigenvectors, as well as some other ideas of linear algebra. Much of the needed material, if not all, will be covered in section 1.1.

Unless otherwise stated, all work will be done over the complex number field, $\mathbb{C}$. The complex number field has many nice properties. The properties that will be most useful to us are that $\mathbb{C}$ is an algebraically closed field, and that it has characteristic 0. The natural numbers, $\mathbb{N}$. will also be utilized.

1.1 Background Material

The aim of this paper is to be accessible to advanced undergraduate students. In this section you will find definitions and examples of material that is assumed throughout the paper. There is nothing more complicated here than what is found in a first-level linear algebra course. This section may therefore be used more as a reference for readers who have not practiced this topic in a while. I have done this in an effort to make the paper as self-contained as possible, though I avoid defining some basic notions and the definitions in this section are not as rigorous. It is expected that most readers will choose to skip this section, especially 1.1.1. For a more complete background, see the references.

1.1.1 Matrices

First, we will be concerned primarily with square matrices. A matrix A is simply a set of numbers arranged into rows and columns.

Example 1.

$$A_1 = \begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n} \end{pmatrix}$$

is an $n \times n$ matrix with an entry $\alpha_{i,j}$ in the ith row and jth column.

Definition 1.1. *The identity matrix, denoted by I, is a square matrix with 1's along the diagonal and 0's everywhere else. That is,*

$$I = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Definition 1.2. *A square matrix N is invertible if there exists a matrix B such that*

$$BN = NB = I$$

We denote the inverse of N as N^{-1}.

Definition 1.3. *A triangular matrix is a matrix with zeroes above or below the main diagonal.*

Example 2.

$$A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

is an upper triangular matrix.

Definition 1.4. *A strictly triangular matrix is a triangular matrix with zeroes along the main diagonal.*

Example 3.

$$A_3 = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

is a strictly upper triangular matrix.

Definition 1.5. *The trace of a matrix N is the sum of its diagonal entries, and is denoted tr(N).*

I will heavily rely on an understanding of the trace of a matrix later in the paper.

Claim. *The trace of matrices is a linear function.*

That is, for some matrices N_0 and N_1, and some scalar α,

1. $\mathrm{tr}(\alpha N_0) = \alpha \mathrm{tr}(N_0)$
2. $\mathrm{tr}(N_0 + N_1) = \mathrm{tr}(N_0) + \mathrm{tr}(N_1)$

Proof. The proof of this claim is given in [4]. □

1.1.2 Eigenvalues and Eigenvectors

As I noted earlier, we will need to understand eigenvalues and eigenvectors in order to triangularize matrices. Occasionally, especially in older texts, eigenvalues are referred to as characteristic values, and eigenvectors are referred to characteristic vectors (see [1] and [7]).

Definition 1.6. *Let N be a square matrix. Suppose there exists some scalar λ and a vector $\vec{v}$ such that*

$$N\vec{v} = \lambda\vec{v}$$

Then we say that N has eigenvector $\vec{v}$ with corresponding eigenvalue λ.

The process of finding eigenvalues and eigenvectors relies on the following:

Theorem. (Cayley-Hamilton Theorem) *Every square matrix over a commutative field satisfies its own characteristic polynomial.*

A proof of the Cayley-Hamilton Theorem is given in [1].

Definition 1.7. *The characteristic polynomial of a matrix N is*

$$det(N - \lambda I) = 0$$

I choose to not define the determinant, or *det*, of a matrix as it is much easier to see in practice, and I will not utilize determinants other than referencing the characteristic polynomial of a matrix. For more, see [4]. A more advanced discussion of the Cayley-Hamilton Theorem can be found in [3] as well as [1]. Now that we know about eigenvalues and eigenvectors, let us look at triangularization.

Definition 1.8. *A matrix N is triangularizable if it is similar to a triangular matrix. That is, if there exists an invertible matrix P such that:*

$$P^{-1}NP = T$$

where T is a triangular matrix.

Let us look at an example of a triangularizable matrix.

Example 4.

$$A_4 = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

A_4 has an eigenvalue of 0, and we get an eigenvector, $\vec{v_1}$

$$\vec{v_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

We want to find an invertible matrix P such that:

$$P^{-1}A_4P = T$$

where T is a triangular matrix.
Letting the first column of $P = \vec{v_1}$, we complete $\vec{v_1}$ to a basis to complete P. So let

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

Then

$$P^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

So

$$\begin{aligned} P^{-1}A_4P &= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

which is upper triangular.
So A_4 is triangularizable.

I also choose not to define a basis. A connection between the eigenvalues and the trace of a matrix can be made.

Claim. *Let N be a square matrix with eigenvalues $\lambda_1, \lambda_2, \ldots \lambda_k$. Then*

$$tr(N) = \sum_{i=1}^{k} \lambda_i$$

Proof. The proof of this claim is given in [4]. Note that this claim relies on the fact that we're working over $\mathbb{C}$. $\square$

For a deeper understanding of these topics, I recommend [1].

1.1.3 Fields

Fields are not *strictly* something that one would work with extensively in a first-level linear algebra course. For this reason, I suggest [3] and [7] for more reading. I will, however, give the definition of a field since I heavily rely on the concept of one throughout the paper. Since it will also be helpful to be reminded of the definition of a group, I will start there.

Definition 1.9. *A set G forms a group under a binary operation $\cdot$ if the following hold:*

(i) $\cdot$ *is associative*

(ii) *G has an identity element*

(iii) *G is closed under inverses*

Now we can define a field. I take this definition from [3].

Definition 1.10. *A field is a set F together with two commutative binary operations $+$ and $\cdot$ on F such that:*

(i) *F is a group under $+$*

(ii) *$(F - \{0\})$ is a group under $\cdot$*

(iii) *the distributive law holds. That is $\forall\ a, b, c \in F$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

2 Sets of Matrices

2.1 Nilpotent Matrices

The first set of matrices we will look at are nilpotent matrices.

Definition 2.1. *A matrix N is nilpotent if $\exists\ k \in \mathbb{N}$ s.t. $N^k = 0$.*

Definition 2.2. *If k is the least natural number s.t. $N^k = 0$, then we say k is the index of N.*

Let us look at an example of a nilpotent matrix.

Example 5. *Let*

$$A_5 = \begin{pmatrix} 2 & 2 & -2 \\ 5 & 1 & -3 \\ 1 & 5 & -3 \end{pmatrix}$$

Then

$$A_5^2 = \begin{pmatrix} 12 & -4 & -4 \\ 12 & -4 & -4 \\ 24 & -8 & -8 \end{pmatrix}$$

And

$$A_5^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So A_5 is a nilpotent matrix.

Notice that the most obvious example of nilpotent matrices are strictly upper or strictly lower triangular. In fact,

Proposition 2.1. *Every strictly upper or strictly lower triangular matrix is nilpotent.*

Proof. Suppose A is an $n \times n$ strictly upper triangular matrix. Let

$$A = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \alpha_{2,1} & 0 & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n-1} & 0 \end{pmatrix}$$

Let the entries of A^2 be $\beta_{i,j}$. Then

$$\beta_{i,j} = \sum_{k=1}^{n} \alpha_{i,k}\alpha_{k,j}$$

By looking at the entries of A, we can see that $\beta_{i,j} = 0$ when $j \geq i$. Consider $\beta_{i,i-1}$.

$$\begin{aligned}
\beta_{i,j-1} &= \sum_{k=1}^{n} \alpha_{i,k}\alpha_{k,i-1} \\
&= \alpha_{i,1}\alpha_{i,i-1} + \alpha_{i,2}\alpha_{2,i-1} + \ldots + \alpha_{i,n}\alpha_{n,i-1}
\end{aligned}$$

Each $\alpha_{i,n}$ or $\alpha_{n,i-1}$ will lie on or above the diagonal of A. So $\beta_{i,i-1} = 0$. Then we have

$$A^2 = \begin{pmatrix}
0 & 0 & \ldots & \ldots & 0 \\
0 & 0 & \ldots & \ldots & 0 \\
\beta_{2,1} & 0 & \ddots & & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\beta_{n,1} & \ldots & \beta_{n,n-2} & 0 & 0
\end{pmatrix}$$

Similarly, if the entries of A^3 are represented by $\gamma_{i,j}$ then we get

$$\gamma_{i,j} = \sum_{m=1}^{n} \alpha_{i,m}\beta_{m,j}$$

Again, $\gamma_{i,j} = 0$ when $j \geq i-2$, by what we saw above. So

$$A^3 = \begin{pmatrix}
0 & 0 & \ldots & \ldots & 0 & 0 \\
0 & 0 & \ldots & \ldots & 0 & 0 \\
\gamma_{3,1} & 0 & \ddots & & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots & 0 \\
\gamma_{n,1} & \ldots & \gamma_{n,n-3} & 0 & 0 & 0
\end{pmatrix}$$

Each time we raise A to a power, the next diagonal row under the main diagonal becomes zeroes. If we continue in this fashion, we can see that eventually we will obtain $A^n = 0$. I will wave my hands a bit here since we have more important things to prove. □

We will now consider the eigenvalues of nilpotent matrices.

Proposition 2.2. *Every nilpotent matrix has all of its eigenvalues equal to 0.*

Proof. Let A be a nilpotent matrix with index k, and let λ be an eigenvalue of A. Then, for some nonzero vector $\vec{v}$,

$$A\vec{v} = \lambda\vec{v}$$

Since $A^{k-1} \neq 0$, then we can multiply by A^{k-1} on each side, producing

$$\begin{aligned}
A^{k-1}A\vec{v} &= A^{k-1}\lambda\vec{v} \\
A^k\vec{v} &= \lambda A^{k-1}\vec{v} \\
0 &= \lambda A^{k-1}\vec{v}
\end{aligned}$$

So $\lambda = 0$, for any eigenvalue λ of A. $\square$

Since the trace of a square matrix over the complex numbers is the sum of its eigenvalues, then

Remark. *The trace of any nilpotent matrix over the complex numbers is 0.*

Now that we have some properties of nilpotent matrices, let us look at triangularization, the central topic of this paper.

Proposition 2.3. *Every nilpotent matrix can be triangularized.*

Proof. Suppose A is a nilpotent, $n \times n$ matrix. We will prove this claim by induction on the size of n. Suppose $n = 1$. Then A is triangular by definition. Suppose now that if A is a nilpotent, $(n-1) \times (n-1)$ matrix, then it is triangularizable. Let A be $n \times n$. Suppose $\vec{v_1}$ is an eigenvector of A. Then let the first column of P be $\vec{v_1}$. So,

$$P = \begin{pmatrix} | & & \\ \vec{v_1} & \cdots & \\ | & & \end{pmatrix}$$

We complete the vector $\vec{v}$ to a basis, and fill in the remaining columns of P with that basis. When we consider at $P^{-1}AP$, we'll get the eigenvalue corresponding to $\vec{v_1}$ as the first entry of the first column of the resulting matrix, with zeroes below. However, we recall that each eignvalue of A must be 0. Then

$$P^{-1}AP = \begin{pmatrix} 0 & \beta_1 & \dots & \beta_{n-1} \\ 0 & \alpha_1 & \dots & \alpha_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_3 & \dots & \alpha_4 \end{pmatrix}$$

Where each α_i, β_j is some complex number. Notice now that

$$\begin{pmatrix} \alpha_1 & \dots & \alpha_2 \\ \vdots & \ddots & \vdots \\ \alpha_3 & \dots & \alpha_4 \end{pmatrix}$$

is a $(n-1) \times (n-1)$ matrix. Since A is nilpotent, this matrix is also nilpotent (this will become more obvious in the discussion below). Then by the induction hypothesis, this matrix can be triangularized. So A can be triangularized. $\square$

In this proof, we can see the motivation behind our method in example 4 when we found an eigenvector and then completed that eigenvector to a basis.

We have seen a few properities of nilpotent matrices, which brings up the question: how can we tell if a matrix is nilpotent? First I will show that

Proposition 2.4. *Any two similar matrices share the same eigenvalues.*

Proof. Suppose you have two similar matrices, N and N_0. Then there exists some invertible matrix P such that

$$P^{-1}NP = N_0$$

Let N_0 have an eigenvector of $\vec{v}$ with corresponding eigenvalue λ. Then $N_0\,\vec{v} = \lambda\vec{v}$. Substituting in $P^{-1}NP$ for N_0, we get

$$P^{-1}NP\vec{v} = \lambda\vec{v}$$

So

$$NP\vec{v} = \lambda P\vec{v}$$

From this, we can see that N has an eigenvector of $P\vec{v}$ with corresponding eigenvalue λ. So N and N_0 have the same eigenvalues. □

Again since the trace of a matrix is the sum of its eigenvalues, it follows from proposition 2.4 that

Remark. *Any two similar matrices over the complex numbers have the same trace.*

Now we can prove the following proposition.

Proposition 2.5. *If a matrix N has all of its eigenvalues equal to 0, then N is nilpotent.*

Proof. Let N be an $n \times n$ matrix with all of its eigenvalues equal to 0. Then

$$det(N - \lambda I) = \lambda^n$$

By the Cayley-Hamilton Theorem, we get $N^n = 0$. So N is nilpotent. □

Notice that by proposition 2.2 and proposition 2.5, we see that

Proposition 2.6. *A matrix N is nilpotent if and only if all of its eigenvalues are equal to 0.*

This proof was partially adapted from the one presented at [6]. From 2.4 and 2.6, it is immediate that

Remark. *Any matrix which is similar to a nilpotent matrix is nilpotent.*

In fact, we can prove that

Proposition 2.7. *Any nilpotent matrix is similar to a strictly upper or lower triangular matrix.*

Proof. Let N be a nilpotent matrix. We know N can be triangularized, so there exists an invertible matrix P and a triangular matrix T such that

$$P^{-1}NP = T$$

Since N and T are similar matrices, they share the same eigenvalues by proposition 2.4. Then by proposition 2.6, both N and T have all of their eigenvalues equal to 0. Since T is triangular, its eigenvalues lie on its diagonal. So T must be a strictly upper or lower triangular matrix. □

Recall that a matrix determines a linear transformation. It is useful to note that

Proposition 2.8. *A nilpotent matrix will remain nilpotent with respect to any basis.*

Proof. Let A be a nilpotent matrix. Suppose you have a chance of basis matrix, B. Then consider $B^{-1}AB$. If

$$A^n = 0$$

Then

$$\begin{aligned}(B^{-1}AB)^n &= B^{-1}A^nB \\ &= 0\end{aligned}$$

So $B^{-1}AB$ is still nilpotent. □

For a discussion on change of basis matrices, see [1]. Suppose now that we want to consider a field other than the complex numbers, or perhaps want to consider any field in general.

Proposition 2.9. *A nilpotent matrix is triangularizable over any field.*

Proof. We already know that any nilpotent matrix can be triangularized. Suppose N is a matrix is nilpotent over a field F. If we consider N over another field, say G, then N would simply undergo a change of basis. And we've already seen that any nilpotent matrix will remain nilpotent with respect to any basis. □

Now that we have worked a bit with nilpotent matrices, let us consider a set of nilpotent matrices. Suppose we wanted to triangularize each matrix in this set. That is, suppose we want this set to be simultaneously triangularizable.

Definition 2.3. *A set of matrices* $(N_1, N_2, \ldots, N_k, \ldots)$ *can be simultaneously triangularized if there exists an invertible matrix* P *such that* $P^{-1}N_jP$ *is triangular, for any* j*.*

Let's specifically look at a semigroup of matrices. I call the theorem below "Levitzki's Theorem"after the mathematician who proved it (see section 5.2), although it is not commonly known by that name.

Theorem. (Levitzki's Theorem) *A semigroup of nilpotent matrices can be simultaneously triangularized.*

Definition 2.4. *A semigroup is a set of elements with a binary operation.*

A light discussion of this theorem can be found in [8]. We will not look at the proof of this theorem, though we will use the theorem in section 3. The proof can be found in [7].

2.2 Unipotent Matrices

The next type of matrices that we will look at are unipotent matrices. Unipotent matrices and nilpotent matrices are closely related, as can be seen in definition 2.5.

Definition 2.5. *A matrix N is unipotent if (N-I) is a nilpotent matrix, where I is the identity matrix.*

Let us look at an example of a unipotent matrix.

Example 6. *Let*

$$A_6 = \begin{pmatrix} 3 & 2 & -2 \\ 5 & 2 & -3 \\ 1 & 5 & -2 \end{pmatrix}$$

Then

$$A_6 - I = \begin{pmatrix} 2 & 2 & -2 \\ 5 & 1 & -3 \\ 1 & 5 & -3 \end{pmatrix}$$

Which is a nilpotent matrix by example 5. So A_6 is unipotent.

Proposition 2.10. *Each unipotent matrix has an inverse.*

Proof. Let N be a unipotent matrix. Then $N = I + N_0$, where N_0 is a unipotent matrix. Recall the power series

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

Then

$$\frac{1}{1+N_0} = 1 - N_0 + N_0^2 - N_0^3 + \cdots$$

Since N_0 is a nilpotent matrix, $N^k = 0$ for some k, so

$$\frac{1}{1+N_0} = 1 - N_0 + N_0^2 - N_0^3 + \cdots \pm N_0^{k-1}$$

Multiplying $(1 + N_0)$ over, we get

$$1 = (1 + N_0)(1 - N_0 + N_0^2 - N_0^3 + \cdots \pm N_0^{k-1})$$

So $I - N_0 + N_0^2 - N_0^3 + \cdots \pm N_0^{k-1}$ is the inverse of N $\square$

Note that in this proof, I use 1 and I interchangeably, as they are the same as far as our use with them goes. Since every unipotent matrix has an inverse, we can have a multiplicative group of unipotent matrices U such that $\forall\ u_1, u_2 \in U$, $u_1 u_2 \in U$. This arises in Lie Theory (see section 5.3). Unipotent matrices also arise in the theory of algebraic groups.

Just as we looked at simultaneous triangularization of a set of nilpotent matrices in section 2.1, we will now look at simultaneously triangularizing unipotent matrices. In particular, we'll look at a multiplicative group of unipotent matrices.

Before we move on, I would like to address the following question. Why do we need to stipulate that we are working in a group? That is because in general, the product of two unipotent matrices is not necessarily nilpotent.

Example 7. *Let*

$$A_0 = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$$

and

$$B_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Then

$$A_0 - I = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

and

$$B_0 - I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Let $A_0 - I = A$ *and* $B_0 - I = B$. *Note that*

$$A^2 = B^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

so both A *and* B *are nilpotent, which means* A_0 *and* B_0 *are unipotent matrices. Consider* A_0B_0. *Is* A_0B_0 *a unipotent matrix?*

$$\begin{aligned} A_0B_0 &= (I + A)(I + B) \\ &= I + A + B + AB \end{aligned}$$

If $A + B + AB$ *is nilpotent, then* A_0B_0 *is a unipotent matrix.*

$$A + B + AB = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -1 & -2 \end{pmatrix}$$

Since $A + B + AB$ *is a* 2×2 *matrix, then if it is nilpotent, it is nilpotent with index at most 2. Let* $A + B + AB = C$. *Clearly* $C \neq 0$. *And*

$$C^2 = \begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix}$$

So $C^2 \neq 0$. *Then* C *is not nilpotent. So* A_0B_0 *is not a unipotent matrix.*

3 Lie-Kolchin Theorem

The Lie-Kolchin Theorem is the mathematical child of two prominent mathematicians, Sophus Lie and Ellis Kolchin. You can read more about them, as well as a bit of background of the theorems, in sections 5.3 and 5.1 respectfully. The Lie-Kolchin Theorem is fundamental to the theory of algebraic groups.

Theorem. (Lie-Kolchin Theorem) *A multiplicative group of unipotent matrices can be simultaneously triangularized.*

We will look at a proof sketch of this theorem. We begin by looking at the nilpotent parts of these unipotent matrices.

Let U be a multiplicative group of unipotent matrices. Suppose we have $A_0, B_0 \in U$. Then let $A_0 - I = A$, $B_0 - I = B$ where A, B are nilpotent matrices. I will show:

Proposition 3.1. *In a multiplicative group of unipotent matrices, the product of two nilpotent parts of unipotent matrices is always a sum of nilpotent elements.*

That is, I will prove that AB is a sum of nilpotent elements.

Proof. Consider the product A_0B_0.

$$\begin{aligned} A_0B_0 &= (I + A)(I + B) \\ &= I + A + B + AB \end{aligned}$$

Since we are in a group of unipotent matrices, then $I + A + B + AB$ is also unipotent. So by definition, $A + B + AB$ is a nilpotent matrix. I pause here to define the circle product.

Definition 3.1. *The circle product of two elements, x_1 and x_1, is defined as:*

$$x_1 \circ x_2 \equiv x_1 + x_2 + x_1x_2$$

Then note that $A \circ B = A + B + AB$. So $A \circ B$ is a nilpotent element. Obvserve that by subtracting A and B on both sides of the equation, we obtain $AB = A \circ B - A - B$, which is a sum of nilpotent elements. $\square$

Now let us take a subalgebra, M, generated by the nilpotent parts of these unipotent elements in U.

Definition 3.2. *A subalgebra is a set closed under addition, scalar multiplication, and multiplication of its elements.*

Then $M = \{\Sigma\alpha_i A_i \ | A_i$ nilpotent parts of elements in U, $\alpha_i \in \mathbb{C}\}$. Suppose we want to take the multiplication of two elements in M. Observe that

$$(\Sigma\alpha_i A_i)(\Sigma\beta_j B_j) = \Sigma\alpha_i\beta_j A_i B_j$$

We know that A_iB_j sum of nilpotent parts of elements in U. So each element in M is a sum of nilpotent elements.

Next we will show that each element of M is nilpotent by considering the trace of each matrix. Recall that the trace of any nilpotent matrix is 0, and that for any two matrices T, T',

$$tr(T + T') = tr(T) + tr(T')$$

Then since every element of M is a sum of nilpotent elements, each element of M has trace 0. In fact, $\mathrm{tr}(A^k) = 0 \ \forall\ k \in \mathbb{N}$, for any $A \in M$, since any power of an element in M will also be in M (M is multiplicative). But I wanted each of these elements to be nilpotent.

Proposition 3.2. *Suppose N is an $n \times n$ matrix over $\mathbb{C}$ such that $tr(N^k) = 0$ for all $k \in \mathbb{N}$. Then N is nilpotent with index n.*

Proof. Let N be an $n \times n$ matrix and we'll induct on the size of n. If $n = 1$ and $\mathrm{tr}(N)$=0, then clearly $N = 0$. So N is nilpotent with index 1. Suppose this holds up to matrices of size $(n-1) \times (n-1)$. Let N be an $n \times n$ matrix with characteristic equation:

$$det(N - \lambda I) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1 \lambda + \alpha_0$$

Suppose that N has no eigenvalues equal to zero. If λ_1 are the eigenvalues of N, then

$$\alpha_0 = \prod_i \lambda_i$$

So if $\lambda_i \neq 0$ for all i, then $\alpha_0 \neq 0$. Then by the Cayley-Hamilton Theorem,

$$\alpha_n N^n + \alpha_{n-1}N^{n-1} + \cdots + \alpha_1 N + \alpha_0 I = 0$$

Taking the trace of each side of the equation, we get

$$tr(\alpha_n N^n + \alpha_{n-1}N^{n-1} + \cdots + \alpha_1 N + \alpha_0 I) = tr(0)$$

Since trace is linear, we can reduce this to

$$\alpha_n tr(N^n) + \alpha_{n-1} tr(N^{n-1}) + \cdots + \alpha_1 tr(N) + \alpha_0 tr(I) = 0$$

We know that $\mathrm{tr}(N^k) = 0$ for all $k \leq n$, so then we get

$$\alpha_0 n = 0$$

Which is a contradiction. So N has an eigenvalue of 0.

Now we are going to begin to triangularize N. Let P be the matrix that is used to triangularize N. We can assume the first column of P is the eigenvector of N corresponding to the eigenvalue of 0 (note that this is the same method I used in example 4). Again, we complete P by completing the basis of this eigenvector. Then if $P^{-1}NP = T$, where T is a triangular matrix, we have

$$T = \begin{pmatrix} 0 & 0 & \dots & 0 \\ \beta_1 & \gamma_1 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n-1} & \zeta & \dots & \gamma_{n-1} \end{pmatrix}$$

Recall that since N and T are similar matrices, they share the same eigenvalues by proposition 2.4 and thus they also have the same trace. Note that if N and T are similar,

$$P^{-1}N^kP = T^k$$

for all $k \in \mathbb{N}$. So $\mathrm{tr}(N^k) = \mathrm{tr}(T^k) = 0$ for all such k. Notice that we have an $(n-1) \times (n-1)$ matrix,

$$T_0 = \begin{pmatrix} \gamma_1 & & 0 \\ \vdots & \ddots & \\ \zeta & \dots & \gamma_{n-1} \end{pmatrix}$$

So if $\mathrm{tr}(T^k) = 0$ then $\mathrm{tr}(T_0^k) = 0$ as well. By our induction hypothesis, since T_0 is an $(n-1) \times (n-1)$ matrix, then T_0 is nilpotent. So T is nilpotent. □

Please note that this is a variation of the proof of Lemma 6.8.3 in [5]. I have chosen to take a slightly different approach in areas, while filling in gaps in others. We also could have used proposition 2.2 and used induction to show that

Proposition 3.3. *Suppose N is an $n \times n$ matrix over $\mathbb{C}$. Suppose $tr(N^k) = 0$ for all $k \in \mathbb{N}$. Then all of the eignevalues of N are 0.*

This would have given us the same result as above. If we look back at example 5, we can see that $\mathrm{tr}(A) = \mathrm{tr}(A^2) = 0$, exactly what we would expect since A is nilpotent.

Let's look at an example in characteristic $\neq 0$, where this theorem fails.

Example 8. *Suppose we're in a field of characteristic 2. Let*

$$A_8 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Then $A_8^k = I\ \forall k \in \mathbb{N}$, but $tr(A_8^k) = 0$.

Now by Levitzki's Theorem that we looked at in section 2.1, we know that M can be simultaneously triangularized. Since M was generated by the nilpotent parts of our unipotent elements in U, all we have to do is extract our results back to the unipotent elements and we have showed the theorem. How exactly can we do this?

Proof. Let N_0 be a unipotent matrix in our group, U. Then let $N_0 - I = N$, where N is a nilpotent matrix. We know by looking at our subalgebra, M, that each of these matrices can be simultaneously triangularized. Let P be the matrix that simultaneously triangularizes M. Then $P^{-1}NP = T$, where T is a triangular matrix. Observe:

$$P^{-1}NP = T$$
$$P^{-1}(N_0 - I)P = T$$
$$P^{-1}N_0P - P^{-1}IP = T$$
$$P^{-1}N_0P - I = T$$
$$P^{-1}N_0P = T + I$$

Since T is a triangular matrix, then $T + I$ is also a triangular matrix. So N_0 is similar to a triangular matrix, which means N_0 is triangularizable. Since N_0 was an arbitrary matrix in U, then we can conclude that every matrix in U is triangularizable by P. So our multiplicative group of unipotent matrices can be simultaneously triangularized. □

We could have concluded our result at an earlier stage if we had used Wedderburn's Theorem.

Theorem. (Wedderburn's Theorem) *If an algebra has a nilpotent basis, it is itself nilpotent.*

Once we know that each element of M is a sum of nilpotent elements we can apply Wedderburn's Theorem in the following way. M is a finite dimensional subalgebra, which means it has a finite basis. Each element of this basis is a sum of nilpotent elements so by taking a spanning set of these elements and then reducing to a basis, we have a nilpotent basis for M. Then we get that each element of M is nilpotent, so we have reduced our size of the Lie-Kolchin Theorem by quite a bit. By using Wedderburn's Theorem, we have avoided using traces of matrices, which means that we know our result holds over any field of characteristic 0 or of characteristic p, where p is prime. The proof of Wedderburn's Theorem can be found in [11]. I do not give it here because it requires advanced methods.

Now we can apply this method to another theorem that Kaplansky calls "Theorem H"on p. 137 of [7]. We are effectively improving Kaplansky's method of proof in a way that has not been noticed before.

Theorem. (Theorem H) *Let S be a multiplicative semigroup of matrices over a field F. Suppose each has the form $\lambda I + N$ for λ in F and N nilpotent. Then S can be simultaneous triangular form.*

That is, the elements of S can be simultaneously triangularized.

Proof. Consider a subalgebra, M, generated by the nilpotent elements N of S. Because S is a multiplicative semigroup, and is therefore closed under multiplication, then we know that for any two elements $\lambda_0 I + N_0$ and $\lambda_1 I + N_1$ in S,

$$(\lambda_0 I + N_0)(\lambda_1 I + N_1) = \lambda_0\lambda_1 I + N_0 + N_1 + N_0N_1$$

Therefore N_0N_1 is a sum of nipotent elements. Again, by what we did above, we know then that each element in M is a sum of nilpotent elements. By Wedderburn's Theorem, each element of M is nilpotent. Here we apply Levitzki's Theorem and we have that M can be simultaneously triangularized. Suppose P is the matrix that simultaneously triangularizes M. Consider $P^{-1}(\lambda I + N)P$. Then

$$\begin{aligned} P^{-1}(\lambda I + N)P &= P^{-1}(\lambda I)P + P^{-1}NP \\ &= \lambda I + T \\ &= T_0 \end{aligned}$$

where T, T_0 are triangular. Then S can be simultaneously triangularized. $\square$

Note that unlike the proof of the Lie-Kolchin Theorem given earlier, since we used Wedderburn's Theorem we know that Theorem H holds not only for characteristic 0, but also for characteristic p.

4 Outlook

The results that I have shown here hold over a field of characteristic 0, though as noted above it is not hard to see that the results hold in a field of characteristic p, where p is prime. In [8], Kaplansky notes that Levitzki's Theorem also works over a division ring. Note that

Definition 4.1. *A division ring is a field in which multiplication is not commutative.*

In [9], Mochizuki shows the following theorem:

Theorem. *Let G be a unipotent group of $n \times n$ matrices over a division ring Δ of characteristic 0 or prime p greater than $(n-1)(n-[\frac{n}{2}])$ where $[\frac{n}{2}]$ is the greatest integer less than or equal to $\frac{n}{2}$. Then G can be simultaneously triangularized.*

In [2], it is shown that:

Theorem. *A unipotent group of matrices over a division ring of characteristic 2 is nilpotent.*

But in general, the results of the Lie-Kolchin Theorem are not known over division rings of characteristic $\neq 0$. It is my hope that I can continue to examine these results in my graduate studies.

5 Historical Notes

While researching material for this course, I've come across many mathematicians who have significant contributions to the field. Along with these contributions, I've learned about the mathematicians themselves. Professor Lance Small has impressed upon me the importance of knowing a bit about each of these people, in an effort to remember that each has a story. This section is a collection of the historical background of each mathematician that I often came across during my studies this year. While it is not the most important section in this paper, I believe it was a necessary supplement. Most of the information in this section has come from Professor Small, although I make use of [10].

5.1 Ellis Kolchin

Photo courtesy of [10]

Ellis Kolchin was an American mathematician who proved the aforementioned Lie-Kolchin Theorem in 1948. Despite the fact that Kolchin served in the U.S. Army in World War II, he was able to publish papers while the war was going on. Kolchin spent his entire career at Columbia University, though he served as a visiting professor in several places, including the Institute for Advanced Study at Princeton University.

5.2 Jacob Levitzki

Photo courtesy of Lance Small

Jacob Levitzki was an Israeli mathematician. He was a doctoral student of Emmy Noether, and proved the theorem that I called “Levitzki’s Theorem”in 1931. Levitzki had a very famous student, Shimshon Amitsur. He died young, in his thirties, as a result of a heart condition. Some of his papers were delayed five or six years due to the breakout of World War II.

5.3 Sophus Lie

Photo courtesy of [10]

Sophus Lie was a Norwegian mathematician in the late 1800s. It is interesting to note the number of amazing mathematicians that Norway has produced, especially when considering its size of the country. Other prominent Norwegian mathematicians that are worth mentioning include Ludwig Sylow and Niels Abel. Lie is most famously known for his advances that led to the branch of mathematics now called Lie theory.

5.4 Joseph Wedderburn

Photo courtesy of [10]

Joseph Wedderburn was a Scottish mathematician who spent most of his professional career in the United States. Wedderburn famously proved that any finite division ring must be a field. In 1909 he began work at Princeton University, and was there until he entered World War I fighting for the British. He served with distinction, but suffered afterwards from what was most likely shell-shock or PTSD. The paper [11] was the last he ever published.

References

[1] Birkhoff, Garrett, and Saunders Mac Lane. *A Survey of Modern Algebra.* New York: Macmillan, 1965.

[2] Derakhshan, Jamshid, and Frank Wagner. Skew Linear Unipotent Groups. Bull. London Math. Soc. 38. 2006 (447-449)

[3] Dummit, David S., and Richard M. Foote. *Abstract Algebra.* 3rd ed. Hoboken, NJ: Wiley, 2004.

[4] Gilbert, Strang. *Linear Algebra and Its Applications.* 5th ed. Thomson, 2006.

[5] Herstein, I. N. *Topics in Algebra.* 4th ed. Wiley, 1975.

[6] jgade. *Nilpotent Matrix* (version 14). PlanetMath.org. Available at http://planetmath.org/NilpotentMatrix.html.

[7] Kaplansky, Irving. *Fields and Rings.* 2nd ed. Chicago: University of Chicago, 1972.

[8] Kaplansky, Irving. The Engel-Kolchin theorem revisited. In Kolchin, E. R., Hyman Bass, Phyllis J. Cassidy, and Jerald Kovacic. *Contributions to Algebra: A Collection of Papers Dedicated to Ellis Kolchin.* New York: Academic, 1977. (233-237)

[9] Mochizuki, H. Y. Unipotent matrix groups over division rings. Canad. Math. Bull. 21. 1978. (249-250)

[10] O'Connor, John J., and Edmund F. Robertson. *The MacTutor History of Mathematics archive.* Available at http://www-history.mcs.st-and.ac.uk/index.html

[11] Wedderburn, J. H. M. Note on Algebras. Ann. of Math, 38. 1937. (854-856)