Dimensions of Nilpotent Algebras

A Major Qualifying Project Report Submitted to The Faculty of Worcester Polytechnic Institute In partial fulfillment of the requirements for the Degree of Bachelor of Science

Written By

Kwabena A. Adwetewa-Badu

Approved by:

March 14, 2021

Abstract

In this thesis, we study algebras of nilpotent matrices. In the first chapter, we give complete proofs of some well-known results on the structure of a linear operator. In particular, we prove that every nilpotent linear transformation can be represented by a strictly upper triangular matrix. In Chapter 2, we generalise this result to algebras of nilpotent matrices. We prove the famous Lie-Kolchin theorem, which states that an algebra is nilpotent if and only if it is conjugate to an algebra of strictly upper triangular matrices.

In Chapter 3 we introduce a family of algebras constructed from graphs and directed graphs. We characterise the directed graphs which give nilpotent algebras. To be precise, a digraph generates an algebra of nilpotency class d if and only if it is acyclic with no paths of length $\geq d$. Finally, in Chapter 4, we give Jacobson's proof of a Theorem of Schur on the maximal dimension of a subalgebra of $M_n(k)$ of nilpotency class 2. We relate this result to problems in external graph theory, and give lower bounds on the dimension of subalgebras of nilpotency class d in $M_n(k)$ for every pair of integers d and n. We conclude the thesis with some open problems.

Introduction

A nilpotent matrix is a special case that, if multiplied by itself a certain number of times, equals the zero matrix. An algebra of nilpotent matrices is simply the operation of nilpotent, Addition, Multiplication, that allow for the result to still be nilpotent. In constructing algebras of nilpotent matrices, one might want to know of the limits to how small or how large the dimensions of such an algebra could be while still maintaining its properties. Using published resources, Mirzakhani and Erdös among others, as well as various literature in Linear Algebra and Graph Theory, this project will explore the upper and lower bounds on the dimensions of subspaces of nilpotent matrices in the matrix algebra $M_n(k)$. For integers n,t we would like to compute the dimension of the largest subspace of $M_n(k)$ which all matrices satisfy $M^t = 0$.

Acknowledgements

I would like to express my deep gratitude to Professor Ó Catháin for their guidance, encouragement, and infinite patience with this thesis. I would also like to thank Professor Weekes, and Guillermo Nunez Ponasso for their aid and resources, as well as the math department for their help in understanding the fundamentals that were the foundation of this thesis. Finally, I wish to thank my family for their support and backing throughout this endeavour.

CONTENTS

1. LINEAR ALGEBRA AND JORDAN CANONICAL FORM

In this chapter, we review eigenvalues, eigenvectors and generalized eigenvectors of linear transformations. This material is essential for understanding the Jordan Canonical Form Theorem. Eventually, we will want to discuss the structure of a nilpotent linear operator. Then in Chapter 2, we will discuss algebras of nilpotent matrices.

1.1. **Background.**

While we expect some familiarity with Linear Algebra, we will review all the notions required for the Jordan Canonical Form. While everything that we discuss holds over an arbitrary algebraically closed field, the reader may assume that all vector spaces are defined over $\mathbb{C}$ without issues. The material in this section is standard, and can be found in any Linear Algebra Textbook, for example *Linear Algebra and its Application*[3] and *Linear Algebra, Fourth Edition*[5]

1.1.1. *Subspaces.*

Definition 1. ([5], 6-7) A vector space V over a field F consists of a set on which two operations, addition and scalar multiplication, are defined so that for each pair of elements $x, y.$ in V there is a unique element $x + y$ in V, and for each element a in F and each element x in V there is a unique element ax in V, such that the following conditions hold:

(1) For all x, y in V, $x + y = y + x$ (commutativity of addition).
(2) For all x, y, z in V, $(x + y) + z = x + (y + z)$ (associativity of addition).
(3) There exists an element in V denoted by 0 such that $x + 0 = x$ for each x in V.
(4) For each element x in V there exists an element y in V such that $x + y = 0$.
(5) For each element x in V, $1x = x$.
(6) For each pair of elements a, b in F and each element x in V, $(ab)x — a(bx)$.
(7) For each clement a in F and each pair of elements x, y in V, $a(x + y) — ax + ay$.
(8) For each pair of elements a, b in F and each element x in V, $(a + b) — ax + bx$.

Definition 2 ([5], page 16-17)**.** A subset W of a vector space V over a field k is called a *subspace* of V if W is a vector space over k with the operations of addition and scalar multiplication defined on V. The zero subspace is also a subspace of V.

Definition 3. ([3],196) Let $v_1, \cdots, v_n$ be vectors in a vector space V. The Span of $v_1, \cdots, v_n$ is

$$Sp(v_1, \cdots, v_n) = \sum_{i=1}^{n} \alpha_i v_i | \alpha \in F$$

$Sp(v_1, \cdots, v_i)$ is always a subspace of V. If $Sp(u_1, \cdots, u_n) = Sp(v_1, \cdots, v_n)$, then $u_1, \cdots, u_n$ is a generating sset for $Sp(v_1, \cdots, v_n)$

Definition 4. ([3], 211) Let H be a subspace of a vector space V. A set of vectors $\mathbb{B} = b_1 \cdots b_n$ in V is a basis for H if:

(1) $\mathbb{B}$ is a linearly independent set, and
(2) the subspace spanned by $\mathbb{B}$ coincides with H; that is, H=Span ($b_1 \cdots b_n$)

Theorem 5 ([5], page 16-17)**.** *Let V be a vector space and W a subset of V. Then W is a subspace of V if and only if the following three conditions hold with the operations defined in V.*

(1) *$0 \in W$*
(2) *x+$y \in W$, whenever $x \in W$ and $y \in W$*
(3) *$\alpha x \in W$, whenever $\alpha \in k$ and $x \in W$*

1.1.2. *Linear transformations and matrices.*

Definition 6. Let V and W be Vector spaces. A function $T : V \longrightarrow W$ is linear if it holds to the following:

(1) : $T(v_1$+$v_2)$ = $T(v_1) + T(v_2)$, for all $v_1, v_2 \in V$
(2) : $T(\alpha v)$= $\alpha T(v)$, for any $\alpha \in \mathbb{R}$, for all $v_1, v_2 \in V$

Definition 7. If T be a linear transformation and B is a basis for V, the matrix $M$with respect to B is as shown below:

$$[M]_B^B = [(Mb_1)_B|(Mb_2)_B|\cdots|(Mb_n)_B]$$

with Mb_i being interpreted as the image of b_i under M.with that knowledge, $[M]_B^B$ can be written as shown below,

$$a_1b_1 + a_2b_2 \cdots a_nb_n = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}_B$$

this way that the matrices can be described in is in terms of change of basis, this way it could give future proving of proofs a more structured foundation onto which they could be attained.

1.1.3. *Eigenvalues and eigenvectors.*

Definition 8. Let T be a linear operator on a vector space V. An *eigenvector* of T is vector v such that

$$Tv = \lambda v$$

for some scalar λ. We refer to λ as the *eigenvalue* of T associated with v.

Suppose that we want to find the eigenvectors of a matrix M. To do this, We will be some initial definitions.

Definition 9. ([3], 105) Let A= $\begin{bmatrix} \text{a} & \text{b} \\ \text{c} & \text{d} \end{bmatrix}$. The *determinant* of A = ad - bc

Definition 10. ([5],209; [10], 396) Let M be a $n \times n$ matrix as shown as below:

$$M = \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix}$$

The $(n-1) \times (n-1)$ matrix $\overline{M}_{i,j}$ can be obtained by deleting the ith row and jth column as follows:

$$\overline{M}_{i,j} = \begin{bmatrix} m_{1,1} & \cdots & m_{1j-1} & \cdots & m_{1,j+1} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{i-1,1} & \cdots & m_{i-1,j-1} & \cdots & m_{i-1,j+1} & \cdots & m_{i-1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{i+1,1} & \cdots & m_{i+1,j-1} & \cdots & m_{i+1,j+1} & \cdots & m_{i+1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ m_{n,1} & \cdots & m_{n,j-1} & \cdots & m_{n,j+1} & \cdots & m_{n,n} \end{bmatrix}$$

Definition 11. ([5], 209-210) Let M be a $n \times n$ matrix, with $m \geq 2$. the determinant of M is as follows:

$$det(M) = \sum_{j=1}^{n} (-1)^{1+j} \times M_{1j} \times det(\overline{M}_{i,j}) .$$

Theorem 12. *([3],105) det $A \neq 0$ iff M has full rank.*

With these Defintions and theorems, We have now have the necessary info for the following Lemma.

Lemma 13. *The scalar λ is an eigenvalue of the matrix M if and only if* $\det(M - \lambda I) = 0$.

Proof. Assume that λ is an eigenvalue of M. Then there exists a non-zero vector v such that $Mv = \lambda v$. So $Mv - \lambda v = \mathbf{0}$. This can be rewritten as $(M - \lambda I)v = \mathbf{0}$, so the matrix $M - \lambda I$ has the non-zero vector v in its Null Space, so $M - \lambda I$ is non-invertible, and has determinant 0.

In the other direction, assume that $\det(M - \lambda I) = 0$. Then there exists a vector $v \in$ Null-space$(M - \lambda I)$. But then $(M - \lambda I)v = 0$, which is equivalent to $Mv = \lambda v$ and v is an eigenvector of M with eigenvalue λ, as required. □

We have shown that every eigenvalue of M is a root of the *characteristic polynomial* $\det(M - \lambda I)$. Furthermore, each eigenvector of M associated with the eigenvalue λ is a vector in the null-space of $M - \lambda I$. In fact, the eigenvectors of M with eigenvalue λ form a subspace of V.

Definition 14. Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. The eigenspace of T corresponding to λ, denoted E_λ is the subset of V defined by

$$E_\lambda = \{x \in V : (T - \lambda I)(x) = 0\}$$

Lemma 15. *The eigenvectors of M with eigenvalue λ along with the zero vector form a subspace of V.*

Proof. We will use Theorem 5 to show that the eigenvectors of M with eigenvalue λ form a subspace of V. First, we observe that $M\mathbf{0} = \lambda\mathbf{0} = \mathbf{0}$. Next, suppose that u, v are eigenvectors. Then

$$M(u + v) = Mu + Mv = \lambda u + \lambda v = \lambda(u + v) ,$$

so $u+v$ is also an eigenvector of V with eigenvalue λ. Finally, if u is an eigenvector and μ is a scalar, then $M(\mu u) = \mu(Mu)$ so μu is also an eigenvector.

We have verified all conditions of Theorem 5, so the eigenvectors of M with eigenvalue λ form a subspace of V as required. □

1.2. **Diagonalisable matrices.** Diagonal matrices are particularly easy to work with. They are precisely the matrices which admit a basis of eigenvectors. Unfortunately, not all matrices are diagonalisable. In this section, we focus on diagonalisable matrices, and later we will prove the Jordan Canonical Form Theorem which explains the general structure of a linear operator.

Definition 16. A matrix M is *diagonal* if $m_{ij} = 0$ whenever $i \neq j$.

Proposition 17. *$[M]_B^B$ is diagonal if and only if every basis vector $b \in B$ is an eigenvector of M.*

Proof. As this is an if and only if statement, we assume one half of the proposition and prove the other.
(1) If $[M]_B^B$ is diagonal , then every basis vector $b \in B$ is an eigenvector of M.
(2) if every basis vector $b \in B$ is an eigenvector of M , then $[M]_B^B$ is diagonal .

For (1), suppose that :

$$[M]_B^B = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Since that the basis b_i is all zeroes except for the ith row

$$[b_i]_B = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}_B$$

because of :

$$[M]_B^B * [b_i]_B = \begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix}_B ,\lambda_i * [b_i] = \begin{bmatrix} 0 \\ \vdots \\ \lambda_i \\ \vdots \\ 0 \end{bmatrix}_B ,[M]_b^b * [b_i]_b = \lambda_i * [b_i]$$

for every i, then every basis $b \in B$ is an eigenvector of M.

For (2), suppose that every vector in B is an eigenvector of M. Then $[Mb_i]$, same as $[b_i]_b$, is non-zero only in the the ith row. Constructing $[M]_b^b$ will produce a diagonal matrix. □

It is useful to be able to decide whether a matrix is diagonalisable.

Definition 18. A matrix M is *diagonalizable* if there exists an invertible matrix P and a diagonal matrix D such that $D = P^{-1} * M * P$

We finish this section with a demonstration of how to diagonalise a matrix A. First, we compute the characteristic polynomial of A, then we factor this polynomial to find the eigenvalues of A (this relies on Lemma 13). Finally, we construct the eigenvectors of A. Since they form a basis for the underlying vector space, we can represent A as a linear transformation with respect to this basis, and by Proposition 17 this matrix must be diagonal. Example:

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

$$A - \lambda * I = \begin{bmatrix} 2-\lambda & 1 & -2 \\ 1 & -\lambda & 0 \\ 0 & 1 & -\lambda \end{bmatrix}$$

$$\det(A - \lambda * I) = (2-\lambda)((-\lambda)^2 - 0) - 1((-\lambda) - 0) - 2(1-0) \Rightarrow (1-\lambda)(1+\lambda)(\lambda-2)$$

$$\lambda = -1, 1, 2$$

with these eigenvalues, we can find the null-space

$$nullspace(A - (-1)\lambda) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, nullspace(A - 1\lambda) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, nullspace(A - 2\lambda) = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$$

with that, we can construct a matrix P that has the eigenvectors of the matrix A as columns, which were found using the nullspace of $(A - \lambda I)$, as its columns:

$$P = \begin{bmatrix} 1 & 1 & 4 \\ -1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

the matrix becomes upper triangular and the diagonal contains the eigenvalues of the matrix A. each of the eigenvalues have instances equal to the multiplicity of the matrix's characteristic polynomial (-λ^3+2*λ^2+λ-2), which when factored, becomes (1-λ)(1+λ)(λ-2), each one having a multiplicity of one.

One way that the matrix can be described is in terms of change of basis, this way , it could give future proving of proofs a more structured foundation onto which they could be attained.

Theorem 19. *A $n \times n$ matrix M is diagonalizable if and only if M has n linearly independent eigenvectors*

Proof. Due to this being a if and only only if statements, we must show that if a $n \times n$ matrix M is diagonalizable, then M has n linearly independent eigenvectors, as well as show that if M has n linearly independent eigenvectors, then n-by-n matrix M is diagonalizable.

Suppose M is diagonalizable. We show that M has n linearly independent eigenvectors. we have to use the definition of a diagonalizable matrix, $D = P^{-1}MP$. Since M is diagonalizable, we can use the diagonal matrix from the definition, as show below:

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

With respect to this basis, the following set of vectors are all eigenvectors of M:

$$\left\{ \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} \cdots \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$$

since the matrix d has n linearly independent eigenvectors, we can conclude that the matrix M does as well.

Suppose M has n linearly independent eigenvectors, then n-by-n matrix M is diagonalizable, for this direction, the second part of proposition 17 will be used (if every basis vector $b \in B$ is an eigenvector of M , then $[M]_B^B$ is diagonal). the basis of M will be diagonal. □

Unfortunately, an $n \times n$ matrix M need not have n linearly independent eigenvectors. It may happen that M has generalized eigenvectors, which we define next.

1.3. **Generalised Eigenvectors and generalised eigenspaces.** There are limits to diagonalization. As stated in Theorem 19, if a $n \times n$ matrix M has n linear independent eigenvectors ,then M is diagonalizable. The limitations of diagonalization occur when a matrix is unable to abide by the rules of diagonalization due to the matrix having less then n linear independent eigenvectors. For example, there will be an attempt made to diagonalize the matrix A, as shown in the example below:

Example 20.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix}$$

$$A - \lambda * I = \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -2 & -5 & -4-\lambda \end{bmatrix}$$

$$\det(A - \lambda * I) = -\lambda^3 - 4 * \lambda^2 - 5 * \lambda - 2$$

$$\lambda = -1, -2$$

the eigenvalues of this matrix is -2 and -1 , with -1 having a multiplicity of 2. By creating eigenvectors from the two eigenvalues, we will produce the eigenvectors, as shown below:

$$nullspace(A - -1\lambda) = nullspace(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -5 & -4 \end{bmatrix} - \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix})$$

$$nullspace(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -2 & -5 & -3 \end{bmatrix})$$

$$[R_3 + 2R_1 \implies R_3] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -3 & -3 \end{bmatrix}$$

$$[R_3/3 \implies R_3] = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$[R_3 + R_1 \implies R_1] = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{bmatrix}$$

$$[R_3 + R_2 \implies R_3] = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$x_1 = x_3, x_2 = -x_3, x_3 = x_3$$

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$nullspace(A - -2\lambda) = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$$

Since there were only two real eigenvectors, it doesn't abide by the rule of the proof, therefore A isn't diagonalizable.

Definition 21. Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. The *generalized eigenspace* of T corresponding to λ, denoted K_λ is the subset of V defined by

$$K_\lambda = \{x \in V : (T - \lambda I)^p(x) = 0, p > 0, p \in \mathbb{Z}\}$$

Lemma 22. *For any T, The generalized eigenspace K_λ is a subspace of V*

Proof. In the same way that we proved that E_λ is a subspace of V, we must prove that K_λ is a subspace of V. We recall theorem 5 to see the requirements for the lemma. First , we will check that the zero vector exists in the generalized eigenspace, which does since $(T - \lambda * I)0$=0.

Next is proving the addition rule of the subspace theorem. Let $x+, y \in K_\lambda$. Suppose $(T-\lambda I)^a(x) = 0$ and $(T-\lambda I)^b(y) = 0$, where $a \geq b > 0$.In order to show that $x+y \in K_\lambda$ we can change $(T-\lambda I)^{(a+b)}(x+y)$ into $(T-\lambda I)^{(a)} * (T-\lambda I)^{(b)}(x+y)$, that way we could change $(T-\lambda I)^{(a+b)}(x+y)$ into $(T-\lambda I)^{(a)}(T-\lambda I)^{(b)}(x)$+ $(T-\lambda I)^{(a)}(T-\lambda I)^{(b)}(y)$ becomes 0 due to the definition of generalized eigenvectors.

The second part to prove is the scalar multiplication. Suppose that $(T-\lambda I)^{(a)}(x) = 0$. Then$(T-\lambda I)^{(a)}(c*x)$ will become $c*(T-\lambda I)^{(a)}(x)$, which becomes $c*0 = 0$ so it satisfies the subspace theorem. This ends the proof. □

We would like to understand the structure of a generalised eigenspace. First, we will construct a basis of K_λ using the linear transformation T.

Proposition 23. *Let $v \in K_\lambda$ be a generalised eigenvector, and suppose that $(T-\lambda I)^m v \neq \mathbf{0}$ but $(T-\lambda I)^{m+1}v = \mathbf{0}$. Then the set $\{(T-\lambda I)^k v \mid 0 \leq k \leq m\}$ is linearly independent and $(T-\lambda I)^m v$ is an eigenvector of T.*

Proof. Write $w_i = (T-\lambda I)^i v$ for $1 \leq i \leq m$. Suppose that

$$\alpha_0 w_0 + \alpha_1 + w_1 + \ldots + \alpha_m w_m = \mathbf{0}$$

for scalars α_i. Then applying $(T-\lambda I)^m$ to both sides of the equation, we get

$$\alpha_0 w_m = \mathbf{0}\,.$$

We conclude that $\alpha_0 = 0$. Next, we apply $(T-\lambda I)^{m-1}$ to both sides of the linear dependence and we get

$$\alpha_0 w_{m-1} + \alpha_1 w_m = \mathbf{0}\,.$$

But we already established that $\alpha_0 = 0$ so we see that $\alpha_1 = 0$ also. Continuing in this way, we find that all of the $\alpha_i = 0$ and so the vectors w_i are linearly independent by definition.

The second claim, that w_m is an eigenvector is immediate from the definition of an eigenvector. □

Using Proposition 23, we can construct a basis for any generalised eigenspace as follows: first we choose a vector v_1 in K_λ which is **not** in the image of $(T-\lambda I)$. Beginning with this vector, we construct a maximal linearly independent set of vectors $w_{1,0}, w_{1,1}, \ldots, w_{1,m}$ as in Proposition 23. If these vectors do not span K_λ then we repeat this process with a vector not contained in Sp$(w_{1,0}, w_{1,1}, \ldots, w_{1,m})$. Eventually, we get a basis of K_λ consisting of vectors $w_{i,j}$. Note that with respect to this basis, the matrix of $T-\lambda I$ is of a very special form: it is zero except possibly for some entries 1 immediately above the diagonal. We will consider such matrices further later in this chapter.

Next, we show that generalised eigenvectors with distinct eigenvalues are linearly independent. First we prove the result for ordinary eigenvectors.

Proposition 24. *Suppose that $v_1, \ldots, v_t$ are eigenvectors of T with distinct eigenvalues $\lambda_1, \ldots, \lambda_t$. Then the eigenvectors are linearly independent.*

Proof. Suppose that there were a linear dependence, say

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_t v_t = \mathbf{0}\,.$$

Consider what happens to this equation when we apply T^k. We obtain

$$\alpha_1\lambda_1^k v_1 + \alpha_2\lambda_2^k v_2 + \ldots + \alpha_t\lambda_t^k v_t = \mathbf{0}\,.$$

Since the λ_i are distinct, we obtain a system of linear equations of the form

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \cdots & \lambda_t \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \cdots & \lambda_t^2 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \cdots & \lambda_t^3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{t-1} & \lambda_2^{t-1} & \lambda_3^{t-1} & \cdots & \lambda_t^{t-1} \end{bmatrix} \times \begin{bmatrix} \alpha_1 v_1 \\ \alpha_2 v_2 \\ \alpha_3 v_3 \\ \vdots \\ \vdots \\ \alpha_t v_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$$

Which is a Vandermode Matrix [5], which is Invertible whenever all λ_i are distinct. The solution to this matrix only works if α_i, for all i, equals zero.

□

The next theorem shows that there are no linear dependences between the generalised eigenvectors of T.

Theorem 25. *Suppose that $v_1, \ldots, v_t$ are generalised eigenvectors of T with distinct eigenvalues $\lambda_1, \ldots, \lambda_t$. Then the generalised eigenvectors are linearly independent.*

Proof. By definition, a generalised eigenvector of height 1 is an (ordinary) eigenvector. Write e_j for the height of v_j , and observe that $T^{e_j - t}v_j$ is a generalised eigenvector of height t. In particular, w_j = $T^{e_j-1}v_j$ is an ordinary eigenvector, with eigenvalue λ. Next, observe that $(T-\lambda_i)v_j$ = $(\lambda_j - \lambda_i)v_j$. Set M = $\prod_{i\neq j}(T-\lambda_j)^{e_j-1}$,then

$$Mv_j = \prod_{i\neq j}(\lambda_i - \lambda_j)^{e_j-1}$$

which is an eigenvector of T with eigenvalue λ_j . We write γ_j for the scalar such that $Mv_j = \gamma_j w_j$. Now, suppose that

$$\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_k v_k = 0$$

is a linear dependence between the generalised eigenvectors v_i. Applying M to this equation, we obtain

$$\alpha_1\gamma_1 w_1 + \alpha_2\gamma_2 w_2 + \ldots + \alpha_k\gamma_k w_k =,$$

which is a linear equation between eigenvectors of T. By Proposition 16, the scalars α_i are all zero, so the generalised eigenvectors are linearly independent.

□

Theorem 26 (Cayley Hamilton, I)**.** *Let $T : V \to V$ be a linear operator over an algebraically closed field, and $G_1, \ldots, G_k$ the generalised eigenspaces of T. Then $T = \oplus_{i=1}^k G_i$.*

Proof. We showed in Theorem 23 that the generalised eigenspaces are disjoint. It will suffice to show that they span V. We will prove this by induction. The base case is dim$V = 1$, in which case the result holds trivially.

Suppose that the result holds for all vector spaces with $\dim V \leq t$ and let V be a vector space of dimension $t+1$. Since T is defined over an algebraically closed field, T has an eigenvector, say $Tv = \lambda v$. So the generalised eigenspace G_1 corresponding to the eigenvalue λ is nonempty. Let U be a complement for G_λ, which is invariant under T. By the induction hypothesis, and by Theorem 25, the space U has a direct sum decomposition into eigenspaces $U = G_2 \oplus G_3 \oplus \cdots \oplus G_k$, where G_i is a generalised eigenspace with eigenvalue λ_i. We need to show that each generalised eigenspace of U is also a generalised eigenspace of V. It will suffice to show that a generalised eigenvector not contained in G_λ is contained in U. Suppose that $w \in G_\lambda$ and $u \in U$, such that $w+u$ is a generalised eigenvector of V with eigenvalue μ, distinct from λ (since otherwise $w+u$ would be contained in G_λ, by definition). By Theorem 25, $(T-\mu I)^{t+1}(w+u) = \mathbf{0}$. Hence $(T-\mu I)^{t+1}w = \mathbf{0}$, and $w \in G_\lambda \cap G_\mu$. But Theorem 25 forces $w = \mathbf{0}$ so the generalised eigenvector belongs to U, and by the induction hypothesis is contained in one of the G_i. Hence, G is a direct sum of generalised eigenspaces. $\square$

As a corollary to the Cayley Hamilton theorem, we can be a little more explicit about the form of a matrix which is written with respect to a basis of generalised eigenvectors of T. Since the image of a generalised eigenvector $v \in G_\lambda$ is another vector in G_λ, we have that the matrix of T with respect to the basis of generalised eigenvectors is of the following form:

$$\begin{bmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & M_t \end{bmatrix}$$

where the matrices M_i are arbitrary for the moment.

Next, consider the action of T on a single generalised eigenspace. We observe that there exists an ordering of the generalised eigenvectors such that $Tv_i = \lambda v_i + \sum_{j>i} \alpha_j v_j$. This follows from the definition of the height of a generalised eigenvector. Hence, it can be shown that each matrix admits a basis with respect to which it is upper triangular. This result is sometimes known as Cauchy's thoerem. Shortly, we will prove the Jordan Canonical Form theorem, which makes this result more precise.

Definition 27. The *characteristic polynomial* of T is the polynomial $\chi(t) = \prod (t-\lambda_i)^{m_i}$ where the product ranges over the eigenvalues of T and m_i is dimension of the generalised eigenspace G_λ.

The minimal polynomial of M is the unique monic polynomial of smallest degree for which $p(M) = \mathbf{0}$.

We can give an alternative statement of the Cayley Hamilton theorem in terms of the characteristic and minimal polynomials. Observe that the polynomial $p_\lambda(x) = (x-\lambda)^b$ annihilates all generalised eigenvectors of height at most b.

Theorem 28 (Cayley-Hamilton, II)**.** *Let M be a matrix, and $\chi_M(t)$ the characteristic polynomial of M. Then $\chi_M(M) = \mathbf{0}$. That is, M satisfies its own characteristic polynomial.*

While logically equivalent to our statement, this one fundamentally obscures the main application of the Cayley-Hamilton theorem: for any $T \in \mathrm{Hom}(V, V)$ there exists a unique decomposition of V into generalised eigenspaces, G_i. Each G_i is T-invariant, and the decomposition $V = \oplus_{i=1}^{t} G_i$ is a direct sum. Hence, with respect to a basis of generalised eigenvectors, T can be expressed as a block-diagonal matrix,

where M_i encodes the action of T on the generalised eigenspace G_i. Applying Cauchy's theorem to each M_i, these matrices are upper triangular with fixed diagonal λ_i.

1.4. **Jordan Canonical Form.** In this section , we will be discussing the Jordan canonical form (JCF) and it's uses for finding the dimensions of a nilpotent matrix.

Definition 29. Let M be a $n \times n$ matrix, with n $> 0. M$ is a *Jordan block* if there is λs on the diagonal, with ones on the super-diagonal, and zeros elsewhere.

Definition 30. A matrix M is in Jordan canonical Form if it was a direct sum of Jordan blocks on the diagonal and zero matrices elsewhere.

The JCF, as show by the matrix in Fig.1 is a way of representing a matrix that is upper triangular and with the eigenvalues being on the diagonal, with multiple instances of a number being equal to the multiplicity of the eigenvalues and the super-diagonal would only have either zeros or ones, with the ones only existing only if the eigenvalues are the same.

$$\left(\begin{array}{cc|c|c|cc}
\lambda_1 & 1 & 0 & \cdots & 0 & 0 \\
0 & \lambda_1 & 0 & \cdots & 0 & 0 \\
\hline
0 & 0 & \lambda_2 & \cdots & 0 & 0 \\
\hline
\vdots & \vdots & \vdots & \ddots & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & \lambda_n & 0 \\
0 & 0 & 0 & 0 & 0 & \lambda_n
\end{array}\right)$$

Theorem 31. *Let G_λ be the generalised eigenspace of T with eigenvalue λ. There exist generalised eigenvectors $v_1, \ldots, v_d \in G_\lambda$ and integers $e_1, \ldots, e_d$ such that each chain*

$$C_i = \langle v_i, (T - \lambda I)v_i, \ldots, (T - \lambda I)^{e_i} v_i \rangle$$

is a T-invariant subspace of G_λ, and $G_\lambda = \oplus_{j=1}^{d} C_j$.

Proof. This proof is by induction on the dimension of G_λ. The base case holds trivially when G_λ has dimension 1: any non-zero vector is a basis, and there is no non-trivial condition to be satisfied. Suppose that all generalised eigenspaces of dimension $\leq r$ can be expressed as a direct sum of cyclic subspaces. For any $v \in G_\lambda$ the cyclic subspace $C_v = \langle (T - \lambda I)^t v \mid t \in \mathbb{N} \rangle$ is T-invariant (the argument is identical to the one given in Proposition

We write M for the restriction of $T - \lambda I$ to G_λ. For a vector $v_i = v_{i,0}$ we write $M^j v_{i,0} = v_{i,j}$. Recall that the *height* of $v_{i,0}$ is the least j such that $v_{i,j} = \mathbf{0}$.

Now, suppose that G_λ has dimension $r + 1$. Since G_λ contains an eigenvector, M is neither injective nor surjective on G_λ. Hence the range of M is a proper subspace U of G_λ, of dimension $\leq k$. Applying the inductive hypothesis to U, we obtain a basis

$$u_{1,0}, \ldots, u_{1,e_1-1},\ u_{2,0}, \ldots, u_{2,e_2-1},\ \ldots,\ u_{d,0}, \ldots, u_{d,e_d-1}\,,$$

for U, where e_i is the height of u_i. Every vector $u \in U$ is of the form Mv for some vector $v \in G_\lambda$ (not necessarily unique). For each $1 \leq i \leq d$ choose a vector $v_i \in G_\lambda$ such that $Mv_i = u_{i,0}$. In particular, $v_{i,j+1} = u_{i,j}$ for any non-negative integer j.

We will show that the vectors

$$v_{1,0}, \ldots, v_{1,e_1},\ v_{2,0}, \ldots, v_{2,e_2},\ \ldots,\ v_{d,0}, \ldots, v_{d,e_d}$$

are linearly independent. Suppose that

$$\alpha_{1,0}v_{1,0} + \ldots + \alpha_{1,e_1}v_{1,e_1} + \ldots + \alpha_{d,e_d}v_{e,e_d} = \mathbf{0}\,.$$

Applying M to both sides of this equation (noting carefully that $Mv_{i,e_i} = 0$),

$$\alpha_{1,0}u_{1,0} + \ldots + \alpha_{1,e_1-1}u_{1,e_1-1} + \ldots + \alpha_{d,e_d-1}u_{d,e_d-1} = \mathbf{0}\,.$$

But the $u_{i,j}$ are linearly independent by the induction hypothesis, so $\alpha_{i,j} = 0$ for all $1 \leq i \leq d$ and $1 \leq j \leq e_i - 1$. What remains is an equation

$$\alpha_{1,e_1}v_{1,e_1} + \ldots + \alpha_{1,e_d}v_{d,e_d} = \mathbf{0}\,,$$

or equivalently, since $v_{i,e_i} = u_{i,e_i-1}$,

$$\alpha_{1,e_1}u_{1,e_1-1} + \ldots + \alpha_{1,e_d}u_{d,e_d-1} = \mathbf{0}\,,$$

Again, by the inductive hypothesis these vectors are linearly independent and so all $\alpha_{i,j}$ are 0.

A careful inspection of the proof thus far shows that we have constructed a basis for a subspace consisting of vectors which belong to $R(M)$, or have a non-zero image in $R(M)$. By the Replacement theorem, we can extend the linearly independent set $v_{i,j}$ given above to a basis of V. Any additional vectors satisfy $x_i \notin R(N)$ and $x_i \in \text{Null}(N)$. These are precisely the eigenvectors which do not belong to any cycle of dimension greater than 1. Hence a basis for G_λ is given by

$$v_{1,0}, \ldots, v_{1,e_1},\ v_{2,0}, \ldots, v_{2,e_2},\ \ldots,\ v_{d,0}, \ldots, v_{d,e_d}\ x_1, \ldots, x_\ell\,.$$

This completes the proof. □

Theorem 31 gives a precise structure for the generalised eigenspace G_λ of any linear operator on a finite dimensional vector space: the sets of vectors $v_{i,0}, \ldots, v_{i,e_i}$ generate maximal cyclic subspaces, each of dimension greater than 1, while the x_i correspond to eigenvectors which are not the image of any generalised eigenvector under $(T - \lambda I)$. Consider the restriction of T to a cyclic subspace, C_i:

$$T\mid_{C_i} = \begin{bmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{bmatrix},$$

which is a matrix with λ on the diagonal, 1 above the diagonal and 0 elsewhere. The restriction of T to G_λ has λ along the diagonal, 0 or 1 along the super-diagonal and 0's elsewhere.

A cyclic subspace of dimension e_i gives $e_i - 1$ consecutive 1's in the super-diagonal, and a zero in the super-diagonal separates the bases of each cyclic subspace.

Theorem 32 (Jordan Canonical Form)**.** *Let $T : V \to V$ be a linear operator. Then V admits a direct sum decomposition $V = \oplus_{i=1}^{r} G_i$ where each G_i is a generalised eigenspace of T. Each G_i admits a direct sum decomposition into cyclic subspaces $G_i = \oplus_{j=1}^{t_i} C_{i,j}$. If $dim C_{i,j} = d$ then the restriction of T to $C_{i,j}$ is a $d \times d$ matrix with λ_i on the diagonal, 1 above the diagonal and all other entries 0.*

The decomposition of V into cyclic subspaces is unique up to the ordering of terms. Let B be a basis of V formed as a union of the bases of the $C_{i,j}$. The matrix $[T]_B$ is a Jordan Canonical Form of T, and is unique up to ordering of the blocks on the diagonal.

Proof. The decomposition of V into generalised eigenspaces is the Cayley-Hamilton theorem, Theorem 26. The decomposition of each generalised eigenspace into cyclic subspaces is Theorem 31. □

It will be useful for us to have multiple characterisations of diagonalisable matrices. Here is a theorem which provides four.

Theorem 33. *let there exist a matrix M. the following statements are equivalent:*
(1) the Jordan canonical form of M is a diagonal matrix
(2)M is diagonalizable
(3) M has no generalized eigenvectors
(4) for everyλ, the algebraic and geometric multiplicity are equal

Proof. ($1 \implies 2$): *If the Jordan canonical form of M is a diagonal matrix, then M is diagonalizable* bending the definition of diagonalizable matrices, the matrix M is diagonalizable if there exists an invertable matrix P and a diagonal matrix , which in this case is JCF(M), such that $JCF(M) = P^{-1} * M * P$. Due to definition 8, if the JCF of M is diagonal, so is the matrix M

($2 \implies 3$): *if M is diagonalizable , then M has no generalized eigenvectors* since M is diagonalizable, then M can be changed into a diagonal matrix. A diagonal matrix is one that has only entries of the diagonal. Due to that, the eigenvectors of that matrix will not be generalized eigenvectors.another way to prove this is to recall back to theorem 11 (A n-by-n matrix M is diagonalizable if and only if M has n linearly independent eiegenvectors). the proof can reworded as *A n-by-n matrix M is diagonalizable if and only if M has n linearly independent (real) eigenvectors and zero generalized eigenvectors* ,as Generalized eigenvectors are not linearly independent.

($3 \implies 4$): *if M has no generalized eigenvectors, then for everyλ, the algebraic and geometric multiplicity are equal* A n-by-n Matrix M that only has real (linearly independent) eigenvectors,due to theorem 11,will be diagonalizable , have n real eigenvectors, and it's diagonal form will appear as below:

$$\left(\begin{array}{c|c|c} \lambda_1 & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & \lambda_n \end{array}\right)$$

having a characteristic polynomial of $(\lambda_1) * (\lambda_2) * \cdots * (\lambda_n)$ for each λ, their algebraic and geometric multiplicity is each 1

($4 \implies 1$): *If for every λ, the algebraic and geometric multiplicity are equal, then the Jordan canonical form of M is a diagonal matrix* If both algebraic and geometric multiplicity are equal, then the matrix M will only have linear independent eigenvectors, and due to theorem 11, will be diagonalizable, and if it is, then it's Jordan form will as well. □

With this section, We will apply the results found to the next section, where we will discuss Nilpotent matrices.

1.5. **Nilpotent matrices.** In this section, we will be discussing the properties of Nilpotent matrices and it's uses for eventually finding the upper limits of it's dimensions.

Definition 34. A matrix M is nilpotent if there exists an integer k such that $M^k = \mathbf{0}$.

Example 35. Let A be the following matrix:

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

Then $A^2 = \mathbf{0}$ so A is an example of a nilpotent matrix.

Lemma 36. *Let A be a nilpotent matrix with entries in $\mathbb{C}$. Then every eigenvalue of M is equal to 0.*

Proof. Since A is nilpotent, there exists an integer k such that $A^k = 0$. Suppose that $Av = \lambda v$ for some vector v. Then $A^k v = \lambda^k v$ and also $A^k v = 0v$. In the complex numbers, the only solution to $\lambda^k = 0$ is $\lambda = 0$, thus every eigenvalue of A is equal to zero. □

We observe that the only nilpotent matrix which is also diagonalisable is the zero matrix.

Lemma 37. *A nilpotent diagonalisable matrix is the zero matrix.*

Proof. If A is diagonalisable then A admits a basis of eigenvectors. By Lemma 36, $Av = \mathbf{0}$ for every v in this basis. So $A = \mathbf{0}$. □

Hence, we can write down the Jordan Canonical form of a nilpotent matrix. By Theorem 32 it suffices to count the number of generalised eigenspaces of each height to understand the Jordan Canonical Form of a nilpotent matrix A. It can be represented as follows:

$$JCF(A) = \begin{bmatrix} M_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & M_n \end{bmatrix}$$

Where $M_1 \cdots M_n$ is a jordan block that is nilpotent.

2. ALGEBRAS OF NILPOTENT MATRICES

In this section, we study sets of nilpotent matrices closed under addition and multiplication. According to our definitions, an algebra need **not** contain an identity element.

2.1. Definitions and elementary properties.

Definition 38. ([6],245) A set $\mathcal{A} \subseteq M_n(\mathbb{C})$ is an *algebra* if

(1) $\mathcal{A}$ is closed under matrix addition.
(2) $\mathcal{A}$ is closed under matrix multiplication.
(3) $\mathcal{A}$ is closed under scalar multiplication by λI for any $\lambda \in \mathbb{C}$.

In other words, an algebra $\mathcal{A}$ carries the structure both of a ring and a vector space.

Definition 39. ([6],267) let A be an algebra and I a subset of A.We say I is an *ideal* if:

(1) $A+B \in I$ for all $A, B \in I$
(2) MA and $AM \in I$ for all $A \in I$ and $M \in A$
(3) $rA \in A$ for $\forall A \in I$ and $\forall r \in R$

Now that we know the preliminaries, we can go forward in discussing the definitions of the Theorem.

Definition 40. A subspace of matrices of $M_n(\mathbb{C})$ which is closed under matrix multiplication is called a *sub-algebra* of $M_n(\mathbb{C})$. A subalgebra is nilpotent if every matrix in the subalgebra is nilpotent.

Definition 41. Let M be a n-by-n matrix

$$M = \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix}$$

M is *Strictly Upper Triangular* if m_{ij} = 0 for every i $\geq$ j.

We can see the strictly upper triangular matrix in the example as follows:

$$M = \begin{bmatrix} 0 & m_{1,2} & \cdots & m_{1,n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & m_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{bmatrix}$$

Lemma 42. *Matrices that are strictly upper triangular are also nilpotent.*

Proof. Suppose that M is a strictly upper triangular matrix. Then $m_{ij} = 0$ whenever $i > j$. This Lemma will be proved using induction on k.

Induction hypothesis: Let m_{ij}^k be the (i, j) entry of M^k. Then $m_{ij} = 0$ whenever $i \geq j - k + 1$.

Base case: When $k = 1$, this is the definition of a strictly upper triangular matrix, which holds for M by hypothesis.

Induction step: Suppose that the induction hypothesis holds for M^k. Then the (i,j) entry of M^{k+1} is obtained as follows:

$$m_{ij}^{k+1} = \sum_{t=1}^{n} m_{it}^{k} m_{tj} \, .$$

By hypothesis, $m_{it}^k = 0$ for all $i \geq t - k + 1$ and $m_{tj} = 0$ for all $t > j$. Hence every term in the sum is zero when at least one equality is satisfied for each value of t. Substituting j for t in the first equality, we see that this occurs precisely when $i \geq j - k + 1$ as required. □

As the next example shows, in general, sums and products of nilpotent matrices need not be nilpotent.

Example 43. Let A and B be the following matrices:

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

A and B are nilpotent by themselves, however, the following sums and products of A and B are **not** nilpotent.

$$A + B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$(A+B)^{2k} = \begin{bmatrix} 1^{2k} & 0 \\ 0 & 1^{2k} \end{bmatrix}, (A+B)^{2k+1} = \begin{bmatrix} 0 & 1^{2k+1} \\ 1^{2k+1} & 0 \end{bmatrix}, (AB)^k = \begin{bmatrix} 1^k & 0 \\ 0 & 0 \end{bmatrix}$$

We record some further properties of algebras of nilpotent matrices.

Proposition 44. *If $\mathcal{A}$ is an algebra of nilpotent matrices, then there does not exist an idempotent element in $\mathcal{A}$. That is, there is no solution in $\mathcal{A}$ to the equation $A^2 = A$ apart from the zero matrix.*

Proof. Suppose that $\mathcal{A}$ contains an element such that $A^2 = A$ and A is non-zero. Then for any integer $k \geq 3$, we have that $A^k = A^{k-2}A^2 = A^{k-2}A = A^{k-1}$ hence by induction, we find that $A^k = A$ for all $k \geq 1$ and so A cannot be nilpotent. □

In fact, we will prove that Proposition 44 is the only obstruction to an algebra being nilpotent. To achieve this goal, we will develop some of the structure theory of matrix algebras, in particular the Lie-Kolchin Theorem.[8]

Lemma 45. *Let $\mathcal{A}$ be a subalgebra of $M_n(k)$. If $\mathcal{N}$ is a nilpotent ideal of $\mathcal{A}$ then $\mathcal{N}^t = 0$ for some $t \leq n^2$.*

Proof. Suppose that $\mathcal{N}^j = \mathcal{N}^{j+1}$, then $\mathcal{N}^j = \mathcal{N}^{j+t}$ for all $t \geq 0$, so $\mathcal{N}^j = \mathbf{0}$. Consider the sequence

$$\dim(\mathcal{N}), \dim(\mathcal{N}^2), \dim(\mathcal{N}^3), \ldots$$

where we take the dimension of $\mathcal{N}^j$ as a vector space in each case. We have shown that this sequence is strictly decreasing until it hits zero. The initial term is bounded by n^2, which yields the result. □

Of course Lemma 45 is not best possible. We will sharpen this bound to $n-1$ once we develop some stronger tools. Since $\mathcal{N}$ is a vector space, we can construct a basis $\mathcal{B}$ from which we obtain a descending chain of ideals which refines that in Lemma 45.

Proposition 46. *If $\mathcal{N}$ is a nilpotent left-ideal of $\mathcal{A}$, then there exists a basis $\mathcal{B} = \{N_1, N_2, \dots, N_d\}$ for $\mathcal{N}$ as a vector space such that $\mathcal{I}_m = \langle N_{m+1}, \dots, N_d\rangle$ is a left-ideal of $\mathcal{N}$ for each m.*

Proof. Suppose that $\mathcal{N}$ is nilpotent of class c, with basis $\mathcal{B}$. We will show that this basis can be reordered to produce the result. Let $N_1 N_2 \dots N_{c-1}$ be a non-zero product of length $n-1$ consisting of elements of $\mathcal{B}$ (not necessarily distinct). Such a product exists by hypothesis. Consider the decomposition

$$\mathcal{N} = \langle N_1\rangle \oplus \mathcal{I}_1$$

where $\mathcal{I}_1$ is spanned as a vector space by the set $\mathcal{B} - \{N_1\}$. A solution to the equation $XY = N_1$ with $X \in \mathcal{I}_1$ and $Y \in \mathcal{N}$ (or $X \in \mathcal{I}_1$ and $Y \in \mathcal{N}$) would contradict the assumption on the nilpotency class of $\mathcal{N}$, it follows that $\mathcal{I}_1$ is a two-sided ideal of $\mathcal{N}$.

Let N_t be the initial term in a non-zero product of maximal length in $\mathcal{I}_t$, decomposing

$$\mathcal{I}_t = \langle N_t\rangle \oplus \mathcal{I}_{t+1}$$

gives an ideal $\mathcal{I}_{t+1}$ which is two-sided in $\mathcal{I}_t$.

By induction, $NN_t \in \mathcal{I}_t$ for any $N \in \mathcal{N}$, hence $NN_t = \alpha N_t + \beta N_{t+1}$ for some $N_{t+1} \in \mathcal{I}_{t+1}$. If α is non-zero, then $N^k N = \alpha^k N + X$ for some $X \in \mathcal{I}_{t+1}$ and $\mathcal{N}$ is not nilpotent. So $\mathcal{I}_{t+1}$ is a left-ideal of $\mathcal{N}$ for all $1 \leq t \leq d$. □

In Proposition 46, each ideal $\mathcal{I}_{t+1}$ is 2-sided in $\mathcal{I}_t$, but not necessarily in $\mathcal{N}$ (in analogy with the construction of a composition series of a finite group in the Jordan-Hölder theorem, for example).

2.2. **The Lie-Kolchin theorem.** Proposition 46 can be extended to characterise nilpotent *subalgebras* in $M_n(k)$. This is essentially the Lie-Kolchin theorem for associative algebras. A similar argument characterises nilpotent and solvable objects in a number of other categories, including Lie Algebras and Algebraic Groups (where the result can be slightly more complicated). Our proof is entirely constructive, though the algorithm we describe is not particularly efficient. We begin with a necessary and sufficient condition for an algebra to be 'triangularisable'.

Proposition 47. *An algebra $\mathcal{A} \subseteq M_n(k)$ is conjugate to an algebra of upper-triangular matrices if and only if there exists a maximal chain of subspaces $0 = V_n \leq V_{n-1} \leq \dots \leq V_0 = V$ which is stabilised by $\mathcal{A}$.*

Proof. Suppose that $\mathcal{A}$ is conjugate to an algebra of upper-triangular matrices, without loss of generality we may assume the $\mathcal{A}$ is upper triangular. Writing e_i for the i^{th} standard basis vector, the subspaces $L_i = \langle e_i, e_{i+1}, \dots, e_n\rangle$ are all fixed subspaces of $\mathcal{A}$. Equivalently, $Me_i = \lambda_M e_i + x_M$ where $x_M \in L_{i+1}$ for any $M \in \mathcal{A}$. Hence, $e_i + L_{i+1}$ is a common eigenvector for $\mathcal{A}$ in the quotient space V/L_{i+1}. So e_i is a generalised eigenvector for $\mathcal{A}$. So V admits consisting of generalised eigenvectors for all of $\mathcal{A}$.

In the other direction, suppose that $\mathcal{A}$ preserves a maximal chain of subspaces, which we write $0 = V_n \leq V_{n-1} \leq \cdots \leq V_0 = V$. Taking v_i to be a generator for the complement of V_{i+1} in V_i, we find that $Mv_i = \lambda_M v_i + x_M$ for some $x_M \in V_{i+1}$. Writing the matrix for M with respect to this basis, we find that it is upper triangular, as required. $\square$

By elementary linear algebra, all eigenvalues of a nilpotent matrix are zero. The 0-eigenspace of a matrix M is precisely the null-space of M, for which a basis exists by the Rank-Nullity theorem. As a result, in the case of nilpotent matrices, many of these results do not require the base field to be closed. A similar result holds for solvable algebras, in which case an algebraically closed field is required to triangularise matrices.

Theorem 48 (Lie-Kolchin, cf. Theorem 26.1 [2])**.** *Suppose that $\mathcal{N}_0$ is a nilpotent subalgebra of $M_n(k)$. Then $\mathcal{N}_0$ is conjugate to a subalgebra of $\mathcal{N}$, the sub-algebra of strictly upper triangular matrices.*

Proof. Let $\mathcal{B}$ be a basis for $\mathcal{N}_0$ constructed as in Proposition 46. Suppose that W_{t+1} is the space of common eigenvectors for $\mathcal{I}_{t+1}$. Recall that $\mathcal{I}_t = \langle N_t, \mathcal{I}_{t+1} \rangle$. We claim that $N_t v \in W_{t+1}$ for any non-zero $v \in W_{t+1}$. This holds since $\mathcal{I}_{t+1}$ is two-sided in $\mathcal{I}_t$ and hence

$$\mathcal{I}_{t+1} N_t v = \mathcal{I}_{t+1} v = 0 \,.$$

Either $N_t v = 0$ in which case v is a common eigenvector of $\mathcal{I}_t$, or $N_t v$ is a non-zero vector in W_{t+1}. In the latter case, $N_t^2 v = 0$ since $N_t^2 \in \mathcal{I}_{t+1}$ so $N_t v$ is a common eigenvector of $\mathcal{I}_t$.

Since $\mathcal{I}_d$, the last non-zero ideal in the chain determined by $\mathbb{B}$ contains a single nilpotent matrix, W_0 is non-zero. The argument above shows inductively that W_{t+1} is non-empty if W_t is non-empty. Hence there exists a common eigenvector v_0 for $\mathcal{N}_0$.

To conclude the proof, apply this argument (again, from scratch) to the induced action of $\mathcal{N}_0$ on $V/\langle v_0 \rangle$. We obtain a vector v_1 such that $V_1 = \langle v_0, v_1 \rangle$ is a 2-dimensional subspace fixed by $\mathcal{N}_0$, and in which $Nv_1 = \alpha v_0$ for any $N \in \mathcal{N}_0$. If $\alpha = 0$ then v_1 is another eigenvector for $\mathcal{N}_0$, otherwise v_1 is a generalised eigenvector. Continuing this process, we construct the required chain of subspaces in V to which we apply Proposition 47. $\square$

Corollary 49. *A nilpotent subalgebra of $M_n(k)$ has nilpotency class at most $n - 1$.*

Proof. Suppose that M is a product of k terms from $\mathcal{N}$, the algebra of s.u.t. matrices. A computation verifies that $m_{ij} = 0$ whenever $i + k < j$ and the result follows. $\square$

Hence, every nilpotent subalgebra of a matrix algebra is conjugate to an algebra of strictly upper triangular matrices.

3. DIGRAPHS AND ASSOCIATED MATRICES

In the previous section, we saw that every nilpotent subalgebra of $M_n(\mathbb{C})$ is conjugate to an algebra of upper triangular matrices. In particular, the dimension of a nilpotent algebra of $M_n(\mathbb{C})$ is bounded above by $\binom{n}{2}$. We also saw that the nilpotency class of such an algebra is at most $n-1$. In this section, we explore lower bounds on the dimension of a nilpotent subalgebra of $M_n(\mathbb{C})$ of bounded nilpotentcy class. Our main tool will be the idea of using the edges of a directed acyclic graph to construct a nilpotent algebra.

3.1. **Graphs, digraphs and their adjacency matrices.** We begin with the definition of a finite graph.

Definition 50. Let V be a finite set, and let E be a subset of $V^{(2)}$, the set of unordered pairs of elements of V. The pair (V, E) is a *graph.*

Normally, we call the elements of V *vertices* and the elements of E *edges*. An important tool for working with graphs is the adjacency matrix, which we define next.

Definition 51. ([11], 14) Let $\Gamma = (V, E)$ be a graph. Label the rows and columns of the matrix A_Γ by the elements of V. We define $A_\Gamma(v_i, v_j) = 1$ if $\{v_i, v_j\} \in E$ and 0 otherwise. We can say that A_Γ is the adjacency matrix of Γ

In other words, the adjacency matrix encodes the edges of of Γ. Perhaps slightly less familiar than the definition of a graph is that of a digraph.

Definition 52. Let V be a finite set, and let E be a subset of $V^{[2]}$, the set of **ordered** pairs of elements of V. The pair (V, E) is a *directed graph.*

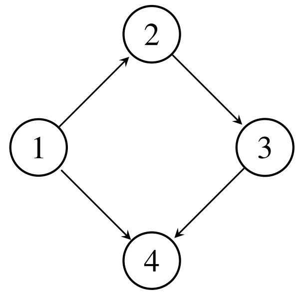
The definition of the adjacency matrix of a directed graph is identical to that for ordinary graphs. Note that while the adjacency matrix of an ordinary graph is always symmetric, any $(0, 1)$ matrix may be the adjacency matrix of a directed graph.

The key property that we will need later is that our graphs and digraphs contain no *cycles*. We define this property next.

Definition 53. A cycle in a graph is a sequence of edges $e_1 = (v_1, v_2)$, $e_2 = (v_2, v_3), \ldots, v_k = (v_k, v_1)$ in which each edge shares one vertex with the previous edge and one vertex with the next edge in the cycle. We consider an undirected edge to be a cycle of length 2. A directed cycle is defined similarly: the output of one edge in the cycle is equal to the input of the next edge. A graph is *acyclic* if it contains no directed cycles.

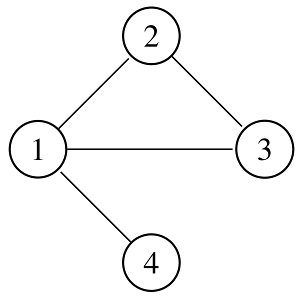
Example 54. We give an example of a directed graph which does not contain a directed cycle, and the corresponding adjacency matrix, which is nilpotent. For convenience, we include the vertex labels on the adjacency matrix.

Example 55. In this example, we give an undirected graph. Note that the adjacency matrix is symmetric and that the square of the matrix has non-negative terms on the diagonal. In fact these are the degrees of the vertices. More generally the diagonal entries of M^k count the number of walks of length k in the graph which start and finish at a given vertex. In particular, the adjacency matrix of an (undirected) graph with a positive number of edges is never nilpotent, because each edge is a cycle of length 2.



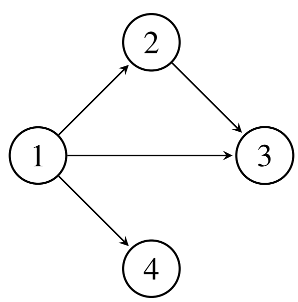
$$\Longleftrightarrow \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 0 & 1 \\ 2 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 1 \\ 4 & 0 & 0 & 0 & 0 \end{array}$$

FIGURE 1. An acyclic directed graph and it's adjacency matrix



$$\Longrightarrow \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 \\ 4 & 1 & 0 & 0 & 0 \end{array}$$

Example 56. Finally, we consider a directed version of the previous graph in which there are no directed cycles. In this case, we note that all directed paths have length at most 2, and that the third power of the adjacency matrix is zero. While this matrix is not S.U.T, it is conjugate to an S.U.T Matrix.



$$\Longrightarrow \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \end{array}$$

Now we formalise the observations made in the three examples given above. The next result relates paths of length k in a directed graph Γ entries in the kth power of the the adjacency matrix, $A(\Gamma)^k$.

Theorem 57. *The number of directed paths of length k in Γ between v_i and v_j is equal to the (i,j) entry of $A(\Gamma)^k$.*

Proof. Let Γ be an undirected graph, such that Γ_{ij}=Γ_{ji}, where Γ_{ij}= 1 if there is an edge between i and j and 0 otherwise, where $1 \leq i, j \leq n$.

$$\alpha - \Gamma = \begin{bmatrix} m_{1,1} & \cdots & m_{1,j} \\ \vdots & \ddots & \vdots \\ m_{i,1} & \cdots & m_{i,j} \end{bmatrix}$$

When $k = 1$, the result is trivially true.

Our induction hypothesis is that $A(\Gamma)^d_{ij}$ is the number of paths of length d between v_i and v_j for all pairs of vertices $v_i, v_j \in \Gamma$.

Then writing $A(\Gamma)^{d+1} = A(\Gamma)^d A(\Gamma)$ and applying the usual formula for matrix multiplication, we see that

$$A(\Gamma)_{ik} = \sum_{j=1}^{n} A(\Gamma)^d_{ij} A(\Gamma)_{jk}\,.$$

By the induction hypothesis, $A(\Gamma)^d_{ij}$ is the number of walks from v_i to v_j of length d. The term $A(\Gamma)_{jk}$ is 1 if there is an edge between v_j and v_k and 0 otherwise. So the product $A(\Gamma)^d_{ij} A(\Gamma)_{jk}$ counts the number of paths of length $d+1$ from v_i to v_k which pass through v_j at the final step. So summing over v_j gives the total number of paths of lenght $d+1$ from v_i to v_k as required. □

Suppose that Γ is a non-empty graph, with an edge between v_i and v_j. Then in every odd power of $A(\Gamma)$ the (i, j) entry is non-zero, and in every even power of $A(\Gamma)$ the (i, i) and (j, j) entries are non-zero. Hence no power of $A(\Gamma)$ is every equal to the zero matrix. We have the following result.

Proposition 58. *The adjacency matrix of an ordinary graph is never nilpotent.*

Proof. Since an ordinary graph isn't directed, the adjacency matrix of it will be symmetric. Symmetric matrices that aren't zero matrices contain the submatrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,which squares to a 2×2 identity matrix, and thus, when multiplied by itself any number of times, won't be a zero matrix □

On the other hand, it is possible to construct directed graphs for which the adjacency matrix is nilpotent. The next result characterises such directed graphs.

Theorem 59. *The following are equivalent.*

(1) *The directed graph Γ has no directed cycles.*
(2) *The adjacency matrix $A(\Gamma)$ is nilpotent.*
(3) *There exists an ordering on vertices such that for all directed edges $[e_i, e_j]$ we have $i < j$.*

Proof. **1 → 2:** If Γ has no directed cycles, then no directed path in Γ can visit the same vertex twice. If $|V| = n$ there can be no directed paths of length greater than or equal to $n - 1$. By Theorem 57 we know that $A(\Gamma)^{n-1} = \mathbf{0}$. Hence $A(\Gamma)$ is nilpotent.

2 → 3: If $A(\Gamma)$ is nilpotent, then $A(\Gamma)^k = \mathbf{0}$ for some k. Recall that the columns of $A(\Gamma)$ are labelled by vertices, and a column contains the zero vector if and only if there are no incoming edges at that vertex. Similarly, a column in $A(\Gamma)^k$ is zero if and only if there are no paths of length k terminating at that vertex. Let X_i be the set of vertices which label a column of zeros in $A(\Gamma)^i$ but not in $A(\Gamma)^{i-1}$. By construction, whenever $i > j$ there is no path from a vertex in X_i to a vertex in X_j. So any ordering of the vertices of Γ constructed by first labelling all vertices in X_1 then labelling all vertices in X_2 and so on will have the required property.

3 $\to$ 1: If $i < j$ for every directed edge, then there cannot be a sequence of edges which forms a closed cycle. $\square$

We can refine this result a little further, as shown below.

Corollary 60. *If the longest directed path in Γ has length $k-1$, and if Γ has no cycles, then $A(\Gamma)$ is nilpotent of class k.*

3.2. **Matrix algebras associated with directed graphs.** In this section we associate a nilpotent matrix algebra with an acyclic directed graph.

Definition 61. Let Γ be an acyclic directed graph on v vertices, and write M_{ij} for the $v \times v$ matrix which contains a 1 in the (i, j) entry and zeros elsewhere.

Define a set of matrices $\alpha(\Gamma) = \{M_{ij} \mid [i, j] \in E(\Gamma)\}$, and define

$$A_\Gamma = \langle \alpha(\Gamma) \rangle .$$

to be the algebra generated by $\alpha(\Gamma)$.

First we prove that $A(\Gamma)$ is really a matrix algebra (that is, that it is closed under matrix multiplication).

Proposition 62. *The algebra $A(\Gamma)$ is closed under matrix multiplication.*

Proof. We construct a basis for $A(\Gamma)$ as follows. Observe that $E_{ij}E_{kl} = \mathbf{0}$ unless $j = k$. Thus the product of the matrices associated with two directed edges is non-zero if and only if the end of the first edge is equal to the start of the second edge. Similarly, a product $E_{i_1j_1} \cdots E_{i_kj_k}$ is zero unless $j_t = i_{t+1}$ for all t. That is: unless the product corresponds to a path in the graph.

Since the graph is acyclic, all paths have finite lenght. In fact, a basis for $A(\Gamma)$ is given by the set of matrices

$$\{E_{ij} \mid \text{ there exists a path from } i \text{ to } j\} .$$

It is clear that the product of two matrices either is zero or corresponds to a path in the graph, and so the algebra is closed under multiplication. $\square$

Proposition 63. *The dimension of $A(\Gamma)$ is equal to the number of pairs (i, j) of vertices in Γ such that there exists a path from i to j in Γ.*

We will be particularly interested in graphs with the property that every pair of vertices joined by a path are already joined by an edge: such digraphs are called *transitive*.

Example 64. Suppose that Γ is the directed graph on $\{1, \dots, n\}$ such that (i, j) is an edge of Γ if and only if $i < j$. Then $A(\Gamma)$ is the algebra of all s.u.t. matrices.

Example 65. Suppose that $X = \{1, \dots, n\}$ and $Y = \{n + 1, \dots, 2n\}$. Define a digraph on $X \cup Y$ by adding all edges **from** X **to** Y. Then $A(\Gamma)$ is the algebra of matrices of the form

$$\begin{pmatrix} 0 & A \\ 0 & 0 \end{pmatrix}$$

It can be verified that this algebra has nilpotentcy class 2, all products in the algebra are zero.

Though it is less important for our main result, it is also possible to define a function from algebras to directed graphs. We define this function and give some of its properties in the remainder of this section.

Definition 66. Let $\mathcal{B}$ be an algebra of $v \times v$ matrices. Define a directed graph as follows:

$$\beta(\mathcal{B}) = (V, E)$$

where $V = \{1, 2, \ldots, v\}$ and $[i, j] \in E$ if and only if there exists a matrix B in $\mathcal{B}$ such that $B_{i,j} \neq 0$.

The result of applying the composite operation $\beta\alpha$ to a graph Γ, and the result of applying $\alpha\beta$ to an algebra $\mathcal{B}$ are described in the next result.

Theorem 67. *The following hold for compositions of α and β.*

(1) *The edge $[i, j]$ is in $\beta\alpha(\Gamma)$ if and only if there is a path between i and j in Γ.*
(2) *$M_{i,j} \in \alpha\beta(\mathcal{B})$ if and only if for some $B \in \mathcal{B}$ the entry $B_{i,j}$ is non-zero.*
(3) *For any graph Γ, it holds that $\alpha\beta\alpha(\Gamma) = \alpha(\Gamma)$.*
(4) *For any algebra $\mathcal{B}$, it holds that $\beta\alpha\beta(\mathcal{B}) = \beta(\mathcal{B})$.*

Proof. the proof is as follows:

(1) (a) If the edge $[i, j]$ is in $\beta\alpha(\Gamma)$, then there is a path between i and j in Γ: Using definition 53, we construct a set of matrices from the graph Γ. With, we use Defnition 58 to make a Graph from $\alpha(\Gamma)$, which would have the same edges as Γ
    (b) If there is a path between i and j in Γ, then the edge $[i, j]$ is in $\beta\alpha(\Gamma)$:
(2) (a) If $M_{i,j} \in \alpha\beta(\mathcal{B})$, then for some $B \in \mathcal{B}$ the entry $B_{i,j}$ is non-zero.
    (b) If for some $B \in \mathcal{B}$ the entry $B_{i,j}$ is non-zero, then $M_{i,j} \in \alpha\beta(\mathcal{B})$.
(3) Using Definitions 53 and 58, we can show that $\alpha(\Gamma)$ equals a set of matrices. Using this and definition 58, we can set $\beta\alpha(\Gamma)$ to be a graph, similar to Γ. With this, we use defintion 58 again, thus seeing that $\alpha(\Gamma)$= $\alpha\beta\alpha(\Gamma)$
(4) Similar to (3), we do the inverse. By using Definitions 53 and 58, we can show that $\beta(\mathcal{B})$ equals a set of matrices. Using this and definition 53, we can set $\alpha\beta(\mathcal{B})$ to be a set of matrices, similar to $\mathcal{B}$. With this, we use defintion 53 again, thus seeing that $\beta(\mathcal{B})$= $\beta\alpha\beta(\mathcal{B})$

□

4. Theorems of Schur, Jacobson, Gallai-Hasse-Roy-Vitaver and Applications

In light of Corollary 60, we can construct a nilpotent algebra of class k from a directed graph with no paths of length greater than $k-1$. It is clear from Definition 61 that the dimension of $\alpha(\Gamma)$ is equal to the number of directed edges in Γ. In this section we apply some theorems from graph theory to bound the number of edges in such a graph.

The Gallai-Hasse-Roy-Vitaver theorem describes the maximal number of edges in a graph with these properties. In this section we give a proof of this result. *The Gallai-Hasse-Roy-Vitaver theorem*

The following result details the relationship between digraphs and nilpotent algebras admitting a basis of $E_{i,j}$ matrices. (Not every nilpotent algebra need have such a basis - compare to Theorem 75.)

Theorem 68. *Let Γ be an acyclic digraph with no paths of length $\geq k$. Then the digraph algebra A_Γ has nilpotency class k. The dimension of A_Γ is the number of edges in Γ.*

The next theorem, discovered multiple times in graph theory gives a family of dense directed graphs having no paths of length $\geq k$.

Theorem 69 (Gallai-Hasse-Roy-Vitaver)**.** *Let Γ be an undirected graph. Over all orientations of the edges of G, the orientations with minimal longest paths come from colourings of G. In particular, the minimal longest path will have length equal to the chromatic number minus 1.*

Proof. Suppose that Γ has chromatic number k, and suppose that a k-colouring is given (with colours $\{1, 2, \ldots, k\}$). Define an orientation on Γ by directing edges toward the larger colour. By definition no edges join vertices with the same colour. So along any directed path, the colours are strictly increasing, so all paths have length at most $k-1$.

Conversely, let Γ be an acyclic directed graph. Label each vertex by the length of a **longest** path **beginning** at that vertex. Observe that each class under this labelling is an independent set: if there were an edge between two vertices in class i, then the source vertex has a path of length $i+1$ emanating from it. Hence this labelling is a proper colouring of Γ. □

So a lower bound on the dimension of an algebra of $n \times n$ matrices of nilpotency class k is obtained by finding the densest graph on n vertices with chromatic number $\leq k$. The latter is a well-studied problem in extremal graph theory, closely related to the so-called Turán problem.

Theorem 70 (Turán, Chapter 41 of [1])**.** *The densest graph on n vertices which does not contain a complete graph K_r as a subgroup is the complete multipartite graph on $r-1$ parts, where any two parts have size differing by at most 1. The number of edges in such a graph is equal to $\frac{r-2}{r-1}\binom{n}{2}$.*

While a graph containing K_r as a subgraph must have chromatic number at least $r+1$ the converse result is **not** true. So the densest graphs with given chromatic number are not known in general, though it is known that they are not asymptotically more dense than the Turán graphs, described above.

Lemma 71. *A k-colourable graph does not have a complete subgraph K_{k+1}.*

Proof. Such a subgraph would require $k+1$ colours. □

Theorem 72 (Erdös, [4])**.** *The number of edges in a graph with chromatic number $r+1$ is of the form*

$$(1+o(1))\,\frac{r-2}{r-1}\binom{n}{2}.$$

In this note we give an exposition of some results of Schur and Jacobsen on algebras of nilpotent matrices.

4.1. **The RREF basis of a nilpotent algebra.** The ideas in this section are drawn from Jacobsen, [7].

Definition 73. Let E_{ij} be the matrix with 1 in position (i,j) and zero elsewhere.

Proposition 74. *Provided $j>i$, the operation $M \mapsto (I+\alpha E_{ij})\,M\,(I-\alpha E_{ij})$ is a change of basis which sends s.u.t. matrices to s.u.t. matrices.*

Proof. First observe that $E_{ij}^2=\mathbf{0}$ so the inverse of $I+\alpha E_{ij}$ is $I-\alpha E_{ij}$. Hence the operation $M \mapsto SMS^{-1}$ is a change of basis. In the product, all three matrices are upper triangular, so the product is upper triangular. If M is assumed nilpotent then the product is nilpotent, all eigenvalues are zero and the matrix is in fact strictly upper triangular. □

We will write $S_{ij}(M) = (I+\alpha E_{ij})\,M\,(I-\alpha E_{ij})$ for this change of basis operation. Observe that the effect of S_{ij} is to **add** a multiple of row j to row i and to subtract the same multiple of column i from column j. The next result is Jacobsen's.

Theorem 75. *Let $\mathcal{N}$ be a $\mathbb{C}$-algebra of nilpotent matrices in s.u.t. form. There exists a conjugate algebra $\mathcal{N}'$ with a basis of matrices $E_{i,j}+X_{i,j}$. In analogy to the RREF, we call $E_{i,j}$ the pivot of the basis element. The matrix $X_{i,j}$ is zero in row i and at all pivots.*

Proof. Since $\mathcal{N}$ is an algebra of $n \times n$ matrices, it is a finite dimensional vector space. Let $B=\{U_1,U_2,\ldots,U_r\}$ be a basis for $\mathcal{N}$ where the U_i are ordered lexicographically. The process we describe below may be considered a modification of the usual algorithm for computing the RREF of a linear system.

Let (i_1,j_1) be the position of the first non-zero entry in U_1. Without loss of generality, we may assume that $U_1(i_1,j_1)=1$ and that $U_k(i_1,j_1)=0$ for $k>1$. By applying a sequence of operations $S_{i_1,j}$ for varying j we obtain a conjugate algebra $\mathcal{N}^{(1)}$ in which the first row of the image of U_1 is zero apart from position (i_1,j_1). Since the effect of $S_{i_1,j}$ is to add column i_i to column j the first rows of matrices $U_2,\ldots,U_r$ are unchanged (though later rows may be changed). Furthermore, all matrices remain s.u.t. Finally, observe that $\mathcal{N}=\langle U_1,\mathcal{N}_1\rangle$ where $\mathcal{N}_1$ is an algebra of dimension $r-1$ in which all matrices are 0 in position (i_1,j_1).

Proceeding with $\mathcal{N}_1$ as we did with $\mathcal{N}$ we obtain a basis $\langle V_1,\ldots,V_r\rangle$ where we have the analogue of the RREF property: each V_i has a leading 1 entry, the row in which this entry is contained has a unique non-zero entry, and all other V_i are zero at this position. □

It is well-known that the general classification problem for nilpotent algebras is intractable: it is known to be related to impossibly general problems in representation theory. Without further assumptions the result above is likely best possible. By enforcing some further assumptions it is possible to proceed further.

4.2. **Algebras of nilpotency class 2.** In this section we investigate algebras of nilpotency class 2.

Theorem 76. *Suppose that $\mathcal{N}$ is nilpotent of class 2. If $E_{i,j} + X_{i,j}$ belongs to the RREF basis of $\mathcal{N}$ then*

(1) *Every element of $\mathcal{N}$ is zero in row j.*
(2) *Every element of $\mathcal{N}$ is zero in column i.*

Proof. Let $M \in \mathcal{N}$. By hypothesis, $(E_{i,j} + X_{i,j})\, M = M\,(E_{i,j} + X_{i,j}) = \mathbf{0}$. On the other hand $X_{i,j}$ is zero in row i. So $(E_{i,j} + X_{i,j})\, M$ contains row j of M in the i^{th} row. Hence row j of M is zero.

The analogous operations on columns prove the second result. $\square$

As a corollary of Theorem 76 we obtain a famous theorem of Schur, which is most often stated for commutative algebras. Recall that an algebra $\mathcal{A}$ is *commutative* if $AB = BA$ for all $A, B \in \mathcal{A}$. This theorem was proved first by Schur, a later proofs were given by Jacobsen and Mirzakhani, [9, 7].

Proposition 77. *If $\mathcal{A}$ is nilpotent of class 2 then $\langle I_n \rangle \oplus \mathcal{A}$ is commutative.*

Proof. Let $a_1 I_n + A_1$ and $a_2 I_n + A_2$ be two elements of $\langle I_n \rangle \oplus \mathcal{A}$. It is easily verified that $\langle I_n \rangle \oplus \mathcal{A}$ is closed under addition. Since $\mathcal{A}$ is nilpotent of class 2, it follows that

$$(a_1 I_n + A_1)(a_2 I_n + A_2) = a_1 a_2 I_n + A_1 + A_2 \in \langle I_n \rangle \oplus \mathcal{A}\,,$$

and so the algbera is closed under multiplication. Finally,

$$(a_2 I_n + A_2)(a_1 I_n + A_1) = a_2 a_1 I_n + A_2 + A_1 = a_1 a_2 I_n + A_1 + A_2\,,$$

and so the algebra is commutative. $\square$

Corollary 78. *In dimension $2n$ the maximal dimension of an algebra of nilpotency class 2 is n^2. Equivalently, the maximal dimension of a commutative algebra of matrices in dimension $2n$ is n^2.*

Proof. By Theorem 76 the problem is reduced to a combinatorial one. From the set $\{(i,j) \mid 1 \leq i,j \leq 2n\}$ we must select the largest possible set D with the following properties:

- If $(i_0, j_0) \in D$ then $(j_0, x) \notin D$ for all $1 \leq x \leq 2n$.
- If $(i_0, j_0) \in D$ then $(x, i_0) \notin D$ for all $1 \leq x \leq 2n$.
- $i \geq j$ for all $(i,j) \in D$.

Equivalently, the entries in the first index must be disjoint from those in the second index. An optimal solution is to split the index set into two pieces, and then select all $E_{i,j}$ where i is drawn from one set and j from the other. A quick calculation shows that the unique optimal solution is to select $i \in \{n+1, \ldots, 2n\}$ and $j \in \{1, \ldots, n\}$, in which case an algebra of dimension n^2 is obtained. This completes the proof of the theorem. $\square$

A similar argument gives the maximal dimension of an algebra of nilpotency class 2 in the algebra of matrices of size $2n+1$, which is $n(n+1)$. We note the connection to digraph algebras (in which $E_{i,j}$ belongs to the algebra if and only if there is a directed path from vertex i to vertex j): the condition of Theorem 76 is *precisely* the condition that the graph algebra contains no paths of length 2. The maximal such graphs are then obvious from consideration of directed bipartite graphs.

Let $A(n,k)$ be the maximal dimension of an algebra of $n \times n$ matrices of nilpotency class k. Let $B(n,k)$ be the maximal number of edges in a graph on n vertices with chromatic number k and let $C(n,k)$ be the number of edges in a Turán graph. We have shown that

$$\frac{k-2}{k-1}\binom{n}{2} \leq C(n,k) \leq B(n,k) \leq A(n,k) \leq \binom{n}{2}.$$

We conclude this thesis with some questions and suggestions for further research.

(1) Are any of the quantities displayed above equal for large n?
(2) Are there any explicit examples of graphs demonstrating that $C(n,k) < B(n,k)$ for large values of n and k?
(3) Are there explicit examples of nilpotent algebras which deomnstrate that $B(n,k) < C(n,k)$ for large values of n and k?

AbstractTable of the values of $C(n, k)$, the number of edges in a Turan graph. Columns are labelled by the nilpotency class k and rows labelled by n

	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	2	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	5	6	6	6	6	6	6	6	6	6	6	6	6
5	6	8	9	10	10	10	10	10	10	10	10	10	10	10
6	9	12	13	14	15	15	15	15	15	15	15	15	15	15
7	12	16	18	19	20	21	21	21	21	21	21	21	21	21
8	16	21	24	25	26	27	28	28	28	28	28	28	28	28
9	20	27	30	32	33	34	35	36	36	36	36	36	36	36
10	25	33	37	40	41	42	43	44	45	45	45	45	45	45
11	30	40	45	48	50	51	52	53	54	55	55	55	55	55
12	36	48	54	57	60	61	63	64	64	65	66	66	66	66
13	42	56	63	67	70	72	73	75	76	76	77	78	78	78
14	49	65	73	78	81	84	85	87	88	89	89	90	91	91
15	56	75	84	90	93	96	98	100	101	102	103	103	104	105

The entry in row i and column j of the above table is the maximal number of edges in a Turan graph on n vertices with vertices divided into at most k groups. By Theorem 68, this quantity is a lower bound on the dimension of a subalgebra of nilpotency class k in $M_n(\mathbb{C})$.

We comment on some features of the table:

(1) Since every nilpotent algebra of class k is nilpotent of class $k+1$ the rows of the table are non-decreasing.
(2) Since every nilpotent algebra of dimension n can be embedded into the algebra of $(n+1) \times (n+1)$ matrices, the table is non-decreasing in columns. (Recall that we do not require a subalgebra to be unital.)
(3) The entries in the first column (labelled $k = 2$) are given by Schur's Theorem, which we proved as Corollary 78.
(4) By Lemma 49, the maximal nilpotency class of a subalgebra of $M_n(\mathbb{C})$ is $n - 1$. So each row of the table achieves its maximal value, which is $\binom{n}{2}$ on the diagonal, and all entries to the right of the diagonal in a given row are equal.

References

[1] Martin Aigner and Günter M. Ziegler. *Proofs from The Book*. Springer, Berlin, sixth edition, 2018. See corrected reprint of the 1998 original [MR1723092], Including illustrations by Karl H. Hofmann.

[2] Daniel Bump. *Lie groups*, volume 225 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[3] Judi J. McDonald David C. Lay, Steven R. Lay. *Linear Algebra and its Applications, Fifth Edition*. Pearson Education, Upper Saddle River, NJ, 1994.

[4] P. Erdös and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, 52:1087–1091, 1946.

[5] Stephen H. Friedberg, Arnold J. Insel, and Lawrence E. Spence. *Linear algebra*. Prentice Hall, Inc., Upper Saddle River, NJ, third edition, 1997.

[6] Joseph A. Gallian. *Contemporary Abstract Algebra Eight Edition*. Brooks/Cole: Cengage Learning, 20 Channel Center Street Boston, MA, 2012.

[7] N. Jacobson. Schur's theorems on commutative matrices. *Bull. Amer. Math. Soc.*, 50:431–436, 1944.

[8] Anthony W. Knapp. *Advanced Algebra, Second Edition*. Anthony W. Knapp, East Setauket, New York, 2016.

[9] M. Mirzakhani. A simple proof of a theorem of Schur. *Amer. Math. Monthly*, 105(3):260–262, 1998.

[10] J. Douglas Faires Richard L. Burden. *Numerical Analysis, Ninth Edition*. Cengage Learning, 20 Channel Center Street Boston, MA 02210 USA, 2010.

[11] Robin J. Wilson. *Introduction to Graph Theory, Fourth Edition*. Longman, Inc, Edinburgh Gate, Harlow, Essex CM20 2JE, England, 1996.