Schur's Theorem on Mutually Commuting Matrices

Sharif University of Technology

A. Abedini, B. Torabi

November 11, 2025

1 Introduction

In the year 1905, Issai Schur (1875 - 1941) wrote a paper [4]) titled "On the theory of commutating matrices" (original title in German: Zur Theorie der vertauschbaren Matrizen) in which he completely identified, up to isomorphism, the maximum number of linearly independent $n \times n$ matrices over $\mathbb{C}$ which pairwise commute with each other. Framed in modern language, the theorem states:

Theorem 1. *The dimension of a commutative $\mathbb{C}$-algebra of square matrices of size n is at most $\lfloor \frac{n^2}{4} \rfloor + 1$.*

Theorem 1 is the celebrated theorem of Schur; however during his proof, Schur actually characterized all such algebras of dimension $\lfloor \frac{n^2}{4} \rfloor + 1$ using the following two theorems:

Theorem 2. *If $n > 3$, every commutative $\mathbb{C}$-algebra of square matrices of order n whose dimension is $\lfloor \frac{n^2}{4} \rfloor + 1$ can be expressed as $E_n \oplus W$ where E_n is the algebra of all scalar matrices of order n and W is an algebra of nilpotent matrices with dimension $\lfloor \frac{n^2}{4} \rfloor$.*

We will prove these theorems in section 5. Before stating Schur's third theorem in the paper, we first need some nomenclature and notation:

We call W a **nil** algebra if every element of W is nilpotent; Schur's theorem 2 shows that it suffices to characterize all such nil algebras W. For that, we recall that $A, B \in M_n(\mathbb{C})$ are called **similar** if there exists an invertible $P \in M_n(\mathbb{C})$ such that $A = PBP^{-1}$. This property can be extended to algebras:

Defintion 3. *We call two $\mathbb{C}$-algebras $\mathcal{A}$ and $\mathcal{B}$ of $n \times n$ matrices* ***equivalent*** *or* ***conjugate*** *if there exists an invertible matrix $P \in M_n(\mathbb{C})$ such that the transformation $M \to PMP^{-1}$ for every $M \in \mathcal{A}$, defines an algebra isomorphism between $\mathcal{A}$ and $\mathcal{B}$.*

Now if we let $r = \lfloor \frac{n}{2} \rfloor$ and $s = \lceil \frac{n}{2} \rceil$, we define two specific $\mathbb{C}$-algebras for each $n \in \mathbb{N}$:

$$\mathcal{A}_n = \left\{ \begin{pmatrix} 0_{r\times r} & 0_{r\times s} \\ A & 0_{s\times s} \end{pmatrix} \middle| A \in M_{s\times r}(\mathbb{C}) \right\} \tag{1.1}$$

$$\mathcal{A}'_n = \left\{ \begin{pmatrix} 0_{s\times s} & 0_{s\times r} \\ A & 0_{r\times r} \end{pmatrix} \middle| A \in M_{r\times s}(\mathbb{C}) \right\} \tag{1.2}$$

By block matrix multiplication, we see that for every two matrices $A, B \in \mathcal{A}_n$ (or similarly for $\mathcal{A}'_n$ and $\mathcal{A}''_n$) we have $AB = BA = 0$ and that all these algebras share the same dimension of $\lfloor \frac{n^2}{4} \rfloor$ (due to the fact that $\lfloor \frac{n}{2} \rfloor \cdot \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$). We shall note in Lemma 7 that for $n > 1$, $\mathcal{A}_n$ and $\mathcal{A}'_n$ are equivalent $\mathbb{C}$-algebras if and only if n is even. Schur concludes his paper by finding a characterization for the nil algebras W described in Theorem 2:

Theorem 4. *A $\mathbb{C}$-algebra such as $\mathcal{W}$ of commuting nilpotent $n \times n$ matrices of dimension $\lfloor \frac{n^2}{4} \rfloor$ exists and:*

a. *If n is even, then $\mathcal{W}$ is equivalent to $\mathcal{A}_n$.*

b. *If n is odd and $n > 3$, then $\mathcal{W}$ is equivalent to $\mathcal{A}_n$ or $\mathcal{A}'_n$.*

c. *If $n = 3$, then $\mathcal{W}$ is equivalent to one of $\mathcal{A}_3$, $\mathcal{A}'_3$ or the algebra $\mathcal{A}''_3$ below:*

$$\mathcal{A}''_3 = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ b & a & 0 \end{pmatrix} \middle| a, b \in \mathbb{C} \right\}$$

In this paper, we revisit Schur's proof of this theorem and restate it using modern mathematical vocabulary; we also describe results that either Schur briefly mentions, or that are continuations of his train of thought.

2 Notation and Lemmas

In this section, we discuss some notation and several key lemmas which are helpful in proving Schur's theorems and at the same time are useful outside of this context.

Remark 5. We note that by "algebra" we are referring to associative algebras; however, we do not necessarily mean a **unital** algebra; in other words, our algebras don't necessarily have a multiplicative identity. In most cases, however, the aforementioned algebras are embedded in unital ones. For example, every algebra we discuss is a subalgebra of $M_n(\mathbb{F})$ for some field $\mathbb{F}$ and therefore we can use its multiplicative identity I_n in the appropriate situations.

Next, we know that a family of diagonalizable matrices commute if and only if they are simultaneously diagonalizable. We prove a generalization of this in the next lemma:

Lemma 6. *Let $M = \begin{pmatrix} X_{n_1 \times n_1} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & Y_{n_2 \times n_2} \end{pmatrix}$ be a complex matrix such that X and Y do not share any eigenvalues. If we name $n := n_1 + n_2$ then any $n \times n$ matrix N that commutes with M must necessarily be of the same block matrix form; in other words, $N = \begin{pmatrix} A_{n_1 \times n_1} & 0_{n_1 \times n_2} \\ 0_{n_2 \times n_1} & D_{n_2 \times n_2} \end{pmatrix}$.*

Proof. Let $N = \begin{pmatrix} A_{n_1 \times n_1} & B_{n_1 \times n_2} \\ C_{n_2 \times n_1} & D_{n_2 \times n_2} \end{pmatrix}$. We want to show that $B = 0$ and $C = 0$. For that, we set MN and NM equal to each other:

$$MN = NM \Rightarrow \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \Rightarrow \begin{pmatrix} XA & XB \\ YC & YD \end{pmatrix} = \begin{pmatrix} AX & BY \\ CX & DY \end{pmatrix}$$

So we have that $XB = BY$ and $YC = CX$. Now we see that by multiplying X on the left of the first equation we have that $X^2B = XBY = (BY)Y = BY^2$. By using induction we have that if $p(x) \in \mathbb{C}[x]$ is a polynomial, then $p(X)B = Bp(Y)$. Now let p be the characteristic polynomial of X; we know by the Cayley-Hamilton theorem that $p(X) = 0$. So we have that $0 = Bp(Y)$. Over $\mathbb{C}$ we know that $p(x)$ can be factored as $(x - \lambda_1)...(x - \lambda_{n_1})$ and therefore $p(Y) = (Y - \lambda_1 I)...(Y - \lambda_{n_1} I)$; since none of $\lambda_1, ..., \lambda_{n_1}$ are eigenvalues of Y, we see that each of the linear factors $Y - \lambda_i I$ is invertible and hence $p(Y)$ is also invertible. So we see that multiplying the equation $0 = Bp(Y)$ by $p(Y)^{-1}$ on the right gives us $B = 0$. A similar reasoning shows that $C = 0$. Therefore, $N = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}$ and this proves our lemma. ■

Now, we shift our focus to the statement promised in the introduction:

Lemma 7. *For $n > 1$, the algebras $\mathcal{A}_n$ and $\mathcal{A}'_n$ defined in (1.1) and (1.2) are equivalent $\mathbb{C}$-algebras if and only if n is even.*

Proof. First, if n is even, then $\lfloor \frac{n}{2} \rfloor = \lceil \frac{n}{2} \rceil$ and therefore $\mathcal{A}_n$ and $\mathcal{A}'_n$ coincide by definition, and are hence equivalent.

For the converse, suppose $n = 2k+1$ is odd. Suppose the contrary: that $\mathcal{A}_{2k+1}$ is equivalent to $\mathcal{A}'_{2k+1}$; so an invertible matrix P exists such that $P\mathcal{A}_{2k+1}P^{-1} = \mathcal{A}'_{2k+1}$. We define $E_{i,j}$ to be the $n \times n$ matrix where every entry is 0, except for the $(i,j)-$th entry which is equal to 1. Let P_j be the j-th column of P, and let P'_l be the l-th row of P^{-1}. By matrix multiplication we see:

$$PE_{i,1}P^{-1} = P_iP'_1 \in \mathcal{A}'_{2k+1} \tag{2.1}$$

If

$$P_i = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{n,i} \end{pmatrix}, \quad P'_1 = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \tag{2.2}$$

then we can deduce:

$$P_iP'_1 = \begin{pmatrix} a_{1,i}b_1 & a_{1,i}b_2 & \dots & a_{1,i}b_n \\ a_{2,i}b_1 & a_{2,i}b_2 & \dots & a_{2,i}b_n \\ \vdots & \vdots & & \\ a_{n,i}b_1 & a_{n,i}b_2 & \dots & a_{n,i}b_n \end{pmatrix} = \begin{pmatrix} 0_{(k+1)\times(k+1)} & 0_{(k+1)\times k} \\ * & 0_{k\times k} \end{pmatrix} \in \mathcal{A}'_{2k+1} \tag{2.3}$$

We know that at least one of the b_j's is non-zero (otherwise P can't be invertible). Using this fact, since the first $k+1$ rows of $P_iP'_1$ have to be zero, we deduce that $a_{i,j} = 0$ for all $j \in \{1, ..., k+1\}$. Since $i \in \{k+1, .., 2k+1\}$ is arbitrary, we see that P must be of the form:

$$P = \begin{pmatrix} A_{(k+1)\times k} & 0_{(k+1)\times(k+1)} \\ B_{k\times k} & C_{(k+1)\times k} \end{pmatrix} = \begin{pmatrix} *_{k\times k} & 0_{k\times 1} & 0_{k\times k} \\ *_{1\times k} & 0 & 0_{1\times k} \\ *_{k\times k} & *_{k\times 1} & *_{k\times k} \end{pmatrix} \tag{2.4}$$

However, we know that in a block upper-triangular matrix, the determinant is the product of the determinants of the block diagonals; in this case, we see that $\det(P) = 0$ which is in contradiction with the invertibility of P. This contradiction shows that for odd $n > 1$, $\mathcal{A}_n$ and $\mathcal{A}'_n$ cannot be equivalent. ■

In the next lemma we shall be using the following common result from linear algebra:

Proposition 8. *(Sylvester's rank inequality) Let $A, B \in M_n(\mathbb{F})$ be two square matrices. Then, we have the following inequality:*

$$\operatorname{rank}(AB) \geq \operatorname{rank}(A) + \operatorname{rank}(B) - n$$

Proof. Since $\ker(B) \subseteq \ker(AB)$, therefore we can restrict the domain of the corresponding linear transformation of B to $\ker(AB)$; we name this linear transformation T. Now, by using the rank-nullity theorem, and by noticing that $\operatorname{im}(T) \subseteq \ker(A)$ and $\ker(T) = \ker(B)$, we can conclude:

$$\begin{aligned} \operatorname{null}(AB) &= \operatorname{rank}(T) + \operatorname{null}(T) \leq \operatorname{null}(A) + \operatorname{null}(B) \\ \Rightarrow n - \operatorname{rank}(AB) &\leq 2n - (\operatorname{rank}(A) + \operatorname{rank}(B)) \Rightarrow \operatorname{rank}(AB) \geq \operatorname{rank}(A) + \operatorname{rank}(B) - n \end{aligned}$$

■

Using this proposition, we now find a neat characterization of $n \times n$ matrices whose square is zero:

Lemma 9. *If $M \in M_n(\mathbb{C})$ such that* $\text{rank}(M) = r$ *and $M^2 = 0$, then $n \geq 2r$ and M is similar to the matrix:*

$$M' = \begin{pmatrix} 0 & 0 \\ I_{r\times r} & 0 \end{pmatrix} \tag{2.5}$$

Proof. Since $M^2 = 0$, by Proposition 8 we can conclude:

$$\text{rank}(M^2) \geq \text{rank}(M) + \text{rank}(M) - n \quad \Rightarrow 0 \geq 2r - n \quad \Rightarrow n \geq 2r$$

Now we prove that M is similar to M'. We know that the matrix M has only one eigenvalue, which is equal to zero. By examining the Jordan normal form of M, we see it is equivalent to:

$$S = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & J_t \end{pmatrix} \tag{2.6}$$

where J_i is a Jordan block of the following form ($i = 1, 2, 3 \dots, t$):

$$J_i = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{l_i \times l_i} \tag{2.7}$$

Now notice that if $l_i > 2$, then $J_i^2 \neq 0$, so for all i we have $l_i \leq 2$. Without loss of generality we can assume that exactly $l_1, \dots, l_k$ ($k \leq t$) are equal to 2, while $l_{k+1}, ..., l_t$ are equal to 1. Because $\text{rank}(M) = r$ and each 2×2 Jordan block contributes exactly one to the rank of M, we must have that $k = r$. Hence M must be similar to S, which is of the form:

$$S = \left(\begin{array}{cccc|c} J_1 & 0_{2\times 2} & \dots & 0_{2\times 2} & 0_{2\times(n-2r)} \\ 0_{2\times 2} & J_2 & \dots & 0_{2\times 2} & 0_{2\times(n-2r)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0_{2\times 2} & 0_{2\times 2} & \dots & J_r & 0_{2\times(n-2r)} \\ \hline 0_{(n-2r)\times 2} & 0_{(n-2r)\times 2} & \dots & 0_{(n-2r)\times 2} & 0_{(n-2r)\times(n-2r)} \end{array}\right) \tag{2.8}$$

Now notice that the procedure above can be replicated for every matrix N such that $\text{rank}(N) = r$ and $N^2 = 0$, and we always obtain the matrix S as a result. In particular, the matrix M' as defined in (2.5) has both properties; so M' and S are similar. Since similarity between matrices of the same order is an equivalence relation, we see that M and M' are both similar to S and therefore similar to each other. ∎

Another result from linear algebra we shall use is the following (for a proof see [3, p. 223]):

Theorem 10. *(Jordan-Chevalley) Every matrix $A \in M_n(\mathbb{C})$ can be written uniquely as the sum of two matrices D and N such that N is nilpotent and D is diagonalizable, such that D has the same characteristic polynomial as A, and the matrices D and N commute.*

However, one theorem stands out above the others; it forms the backbone of the proofs that Schur brings, and is a remarkable theorem in its own right. This theorem was first proved by Élie Cartan (1869 - 1951) in his paper "Les groupes bilinéaires et les systèmes de nombres complexes" and bears his name [1]. Let $\mathcal{A}$ be a finite-dimensional nil $\mathbb{C}$-algebra. Cartan's Theorem states the following:

Theorem 11. *(Cartan) In every finite-dimensional nil $\mathbb{C}$-algebra $\mathcal{A}$, there exists a non-zero element $a \in \mathcal{A}$ such that $ax = 0$ and $xa = 0$ for every $x \in \mathcal{A}$.*

The proof here is based on Frobenius' proof of this theorem, appearing in his "Theorie der hyperkomplexen Grössen II" [2]. Note that as discussed in the beginning of this section we use I as the identity of our $\mathbb{C}$-algebra, since the algebras discussed in Schur's proofs are subalgebras of $M_n(\mathbb{C})$.

Lemma 12. *Let $\mathcal{A}$ be a finite-dimensional nil $\mathbb{C}$-algebra. Suppose $x_1, x_2, \ldots, x_n \in \mathcal{A}$ such that $x_1x_2...x_n \neq 0$. Then the n elements*

$$x_1,\ x_1x_2,\ x_1x_2x_3, \ldots,\ x_1x_2...x_n \tag{2.9}$$

are linearly independent.

Proof. Assume the contrary, that they are linearly dependent; so we can find $1 \leq l \leq n$ and coefficients $a, a_{l+1}, ..., a_n \in \mathbb{C}$ such that $a \neq 0$ and:

$$ax_1x_2...x_l + a_{l+1}x_1x_2...x_{l+1} + a_{l+2}x_1x_2...x_{l+2} \cdots + a_nx_1x_2...x_n = 0 \tag{2.10}$$

We name:

$$a_{l+1}x_{l+1} + a_{l+2}x_{l+1}x_{l+2} + \cdots + a_nx_{l+1}...x_n = -y \tag{2.11}$$

So by (2.10) we get:

$$x_1...x_l(aI - y) = 0 \tag{2.12}$$

Note that $aI - y$ is invertible because y is nilpotent and aI is invertible. So we can conclude that $x_1x_2...x_l = 0$ and therefore $x_1x_2...x_n = 0$ which contradicts the assumption. So the n elements in the statement of the lemma have to be linearly independent. ■

Corollary 13. *Let $\mathcal{A}$ be a finite-dimensional nil $\mathbb{C}$-algebra.Suppose $\dim_{\mathbb{C}} \mathcal{A} = n$; then, the product of any $n+1$ elements of $\mathcal{A}$ is zero.*

Proof. We apply Lemma 12: if $x_1, x_2, ..., x_{n+1} \in \mathcal{A}$ such that $x_1x_2...x_{n+1} \neq 0$, then the elements

$$x_1,\ x_1x_2,\ x_1x_2x_3, \ldots,\ x_1x_2...x_{n+1} \tag{2.13}$$

are linearly independent and that contradicts our assumption that $\dim_{\mathbb{C}} \mathcal{A} = n$ ■

Proof. (proof of Cartan's Theorem 11) Let $m \geq 2$ be the smallest positive integer such that the product of any m elements of $\mathcal{A}$ is zero; by Corollary 13, we know that such an m exists. Now since m is the smallest such number, we know that the product of some $m-1$ elements of $\mathcal{A}$ is not zero, such as $x_1, ..., x_{m-1}$. If we set $a := x_1...x_{m-1}$, we see that $ay = 0$ and $ya = 0$ for every $y \in \mathcal{A}$ because both ay and ya are the product of m elements of $\mathcal{A}$. So we have found a non-zero element $x \in \mathcal{A}$ that when multiplied by any element of $\mathcal{A}$ gives us zero; and hence, Cartan's Theorem is proved. ■

Corollary 14. *Every finite-dimensional nil $\mathbb{C}$-algebra is nilpotent.*

3 Proof with only one eigenvalue

We define v_n+1 to be the maximum dimension of a commutative $\mathbb{C}$-algebra of $n \times n$ matrices. First of all:

Proposition 15. *For all positive integers n we have $v_n \geq \lfloor \frac{n^2}{4} \rfloor$.*

Proof. It suffices to give an example of a commutative algebra of $n \times n$ matrices of dimension $\lfloor \frac{n^2}{4} \rfloor + 1$. In fact, $\mathcal{M} = E_n \oplus \mathcal{A}_n$ is such an algebra. To check commutativity, if we take two matrices in $\mathcal{M}$ such as $\alpha I + A$ and $\beta I + B$ where $\alpha, \beta \in \mathbb{C}$ and $A, B \in \mathcal{A}_n$, then we see that:

$$(\alpha I + A)(\beta I + B) = \alpha\beta I + \alpha B + \beta A + AB$$

$$(\beta I + B)(\alpha I + A) = \alpha\beta I + \alpha B + \beta A + BA$$

Both of the right hand sides are equal to $\alpha\beta I + \alpha B + \beta A$ since $AB = BA = 0$ for every $A, B \in \mathcal{A}_n$. So $\mathcal{M}$ is a commutative $\mathbb{C}$-algebra of dimension $\lfloor \frac{n^2}{4} \rfloor + 1$ of commuting $n \times n$ matrices; therefore, since $v_n + 1$ was defined to be the largest dimension of such an algebra, it must be at least that amount. ■

Remark 16. The reasoning of Proposition 15 also shows us that the maximum dimension of a commutative nil algebra of $n \times n$ matrices is also at least $\lfloor \frac{n^2}{4} \rfloor$, since the algebras $\mathcal{A}_n$ are examples of such algebras.

So we just need to prove the upper bound for v_n. We will first do this by putting a strong assumption on the matrices; afterwards we will use induction to relax this extra assumption.

Theorem 17. *Let $\mathcal{W}$ be a commutative nil $\mathbb{C}$-algebra of $n \times n$ matrices of maximum dimension. Then we have* $\dim_{\mathbb{C}}(\mathcal{W}) = \lfloor \frac{n^2}{4} \rfloor$.

Proof. We prove this statement by strong induction on n. For $n = 1$, there is nothing to prove. Now assume that for all $n' < n$, for any commutative nil $\mathbb{C}$-algebra of $n' \times n'$ matrices such as $\mathcal{W}'$ of maximum dimension, we have $\dim_{\mathbb{C}}(\mathcal{W}') = \lfloor \frac{n'^2}{4} \rfloor$. Let $\mathcal{W}$ be an algebra as described in the statement of the theorem, and let $m := \dim_{\mathbb{C}}(\mathcal{W})$. By invoking Cartan's Theorem 11, we have a non-zero matrix $M \in \mathcal{W}$ such that for every $X \in \mathcal{W}$ we have $MX = XM = 0$. We can also assume that $\operatorname{rank}(M)$ is the largest amount possible (this will be used in other parts of the proof). Substituting $X = M$ we get that $M^2 = 0$; therefore by Lemma 9, we know that by setting $r = \operatorname{rank}(M)$, we can write $n = 2r + s$ where $s \geq 0$, and that an invertible matrix $P \in M_n(\mathbb{C})$ exists such that:

$$PMP^{-1} = \begin{pmatrix} 0_{r\times r} & 0_{r\times s} & 0_{r\times r} \\ 0_{s\times r} & 0_{s\times s} & 0_{s\times r} \\ I_{r\times r} & 0_{r\times s} & 0_{r\times r} \end{pmatrix}$$

Then by naming $\mathcal{W}' = P\mathcal{W}P^{-1}$, if we let $X = \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{2,2} & X_{2,3} \\ X_{3,1} & X_{3,2} & X_{3,3} \end{pmatrix} \in \mathcal{W}'$, we see that:

$$(PMP^{-1})X = 0 = X(PMP^{-1}) \Rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ X_{1,1} & X_{1,2} & X_{1,3} \end{pmatrix} = 0 = \begin{pmatrix} X_{1,3} & 0 & 0 \\ X_{2,3} & 0 & 0 \\ X_{3,3} & 0 & 0 \end{pmatrix} \tag{3.1}$$

Therefore the block matrices $X_{1,1}, X_{1,2}, X_{1,3}, X_{2,3}, X_{3,3}$ must all be zero matrices of the appropriate sizes; hence, X is of the form:

$$X = \begin{pmatrix} 0_{r\times r} & 0_{r\times s} & 0_{r\times r} \\ A & B & 0_{s\times r} \\ C & D & 0_{r\times r} \end{pmatrix} \tag{3.2}$$

Since $\mathcal{W}$ is commutative, nil and of dimension m, so $\mathcal{W}'$ is too. This means that by Proposition 15 we can assume:

$$m \geq \lfloor \frac{n^2}{4} \rfloor = \lfloor \frac{(2r+s)^2}{4} \rfloor = r^2 + rs + \lfloor \frac{s^2}{4} \rfloor \tag{3.3}$$

Notice that X is a lower-triangular block matrix; therefore if $X^k = 0$ for some positive integer k, then since the middle block of X^k is B^k, the matrix B is therefore nilpotent. This ensures that the set $\mathcal{B}$ of all such B's in (3.2) forms a nil $\mathbb{C}$-algebra; let $b := \dim_{\mathbb{C}}(\mathcal{B})$. By our induction hypothesis, $b = \lfloor \frac{s^2}{4} \rfloor$. Also, the set of all matrices C forms a vector space $\mathcal{C}$ of dimension c; as $\mathcal{C} \subseteq M_r(\mathbb{C})$, we readily have that $c \leq r^2$.

Using bases for $\mathcal{B}$ and $\mathcal{C}$ which we call $\{B_1, ..., B_b\}$ and $\{C_1, ..., C_c\}$ respectively, we now construct a base of $\mathcal{W}'$. Let $X_1, ..., X_b$ be matrices of form (3.2) which are zero except in the B block, which contains the b elements of our basis for $\mathcal{B}$. Furthermore, let $X_{b+1}, ..., X_{b+c}$ be matrices of the same form such that they are everywhere zero except in the C block which contains the c matrices of our basis for $\mathcal{C}$. In essence, we construct a basis for a subspace of dimension $t := b + c \leq \lfloor \frac{s^2}{4} \rfloor + r^2$ in $\mathcal{W}$ as follows:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & \mathcal{B} & 0 \\ \mathcal{C} & 0 & 0 \end{pmatrix}$$

. We now know that we can extend this basis to a basis for a subspace of the same dimension in $\mathcal{W}'$, such as $\mathcal{X} = \{X_1, ..., X_m\}$. Without any loss of generality, let the matrices $X_{t+1}, ..., X_{t+g}$ span all the matrices in the A block of form (3.2), and that $X_{t+1}, ..., X_m$ span all the matrices in the D block of that form. To summarize: the first b elements of $\mathcal{X}$ span the B blocks, the next c elements span the C blocks, the next g elements span the A blocks and these g elements and the remaining $m - t - g$ elements span the D blocks. From (3.3) and the definition of t it can be deduced that $m - t \geq rs$.

Here, the magic happens. Let A_i be the A block of the base matrix X_{t+i}, for all $i \in \{1, ..., g\}$. We now construct a new matrix K by writing the A_i matrices together, so:

$$K_{s \times gr} = \big(A_1 | A_2 | \dots | A_g \big) \tag{3.4}$$

Suppose $\mathrm{rank}(K) = k$; so we know that $0 \leq k \leq s$. Let $l = s - k$. Assume, without loss of generality, that the k linearly independent rows of K are the $(l+1)$-th row down to the s-th row; so we can write the first l rows of K as linear combinations of the last k rows. This means that if we name the rows of K as $a_1, ..., a_l, b_1, ..., b_k$ then constants $[\lambda_{i,j}]$ for $1 \leq i \leq l, 1 \leq j \leq k$ exist such that $a_i = \sum_{j=1}^{k} \lambda_{i,j} b_j$. Let $T := [\lambda_{i,j}] \in M_{l \times k}(\mathbb{C})$; we see that:

$$\begin{pmatrix} a_1 \\ \hline \vdots \\ \hline a_l \end{pmatrix} = T \begin{pmatrix} b_1 \\ \hline \vdots \\ \hline b_k \end{pmatrix}$$

Now we introduce two matrices named S and Q:

$$S = \begin{pmatrix} I_{l \times l} & -T \\ 0 & I_{k \times k} \end{pmatrix} \in M_s(\mathbb{C}) \tag{3.5}$$

$$Q = \begin{pmatrix} I_{r \times r} & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & I_{r \times r} \end{pmatrix} \in M_n(\mathbb{C}) \tag{3.6}$$

Note that S is an invertible matrix, hence by (3.2) and block multiplication we have :

$$QXQ^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ SA & SBS^{-1} & 0 \\ C & DS^{-1} & 0 \end{pmatrix} \tag{3.7}$$

As the first l rows of SK are zero by our construction of the matrices T and S, and since the g matrices placed in K span any A block matrix in (3.2), therefore for any A matrix in that form, the first l rows of SA are also zero. So by naming $\mathcal{W}'' = Q\mathcal{W}'Q^{-1}$, the construction of S and Q shows that $\mathcal{W}'$ is equivalent to an algebra (namely, $\mathcal{W}''$) where the first l rows of the A matrices are identically zero. So we can assume without loss of generality that $\mathcal{W}'$ also has this property. As a consequence, we see that $g \leq rk$ as each A matrix can have at most rk non-zero elements.

Now by the commutativity of $\mathcal{W}'$, X_i commutes with every matrix

$$Y = \begin{pmatrix} 0 & 0 & 0 \\ A' & B' & 0 \\ C' & D' & 0 \end{pmatrix}$$

in our algebra; if we let $i > t + g$, then $X_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & D_i & 0 \end{pmatrix}$. Therefore we have:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ D_iA' & D_iB' & 0 \end{pmatrix} = X_iY = YX_i = \begin{pmatrix} 0 & 0 & 0 \\ A' & B' & 0 \\ C' & D' & 0 \end{pmatrix} \begin{pmatrix} 0_{r\times r} & 0_{r\times s} & 0_{r\times r} \\ 0_{s\times r} & 0_{s\times s} & 0_{s\times r} \\ 0_{r\times r} & D_i & 0_{r\times r} \end{pmatrix} = 0 \tag{3.8}$$

We obtain that $D_iA = 0$ for every A block matrix; especially, for the matrices $A_1, ..., A_g$ which spanned the A block of matrices in (3.2) and which were also placed in the matrix K. Therefore from (3.4), we deduce that $D_iK = 0$. Since the k last rows of K are linearly independent, we conclude that the k last columns of D_i must be zero. So D_i can have at most rl elements and therefore we have that $m - t - g \leq rl$. We now put all of our information together:

$$g \leq rk, m - t - g \leq rl \ \Rightarrow m - t \leq r(k + l) = rs \tag{3.9}$$

But from (3.3) and (3.9), we conclude that $m - t = rs$ and therefore all our inequalities are equalities and therefore:

$$g \ = rk, \ m - t - g \ = rl, \ b \ = \lfloor \frac{s^2}{4} \rfloor, \ c = r^2 \tag{3.10}$$

This proves that $m = r^2 + rs + \lfloor \frac{s^2}{4} \rfloor = \lfloor \frac{n^2}{4} \rfloor$. ■

Corollary 18. *If $\mathcal{M}$ is a commutative algebra of $n \times n$ matrices of maximum dimension such that every matrix in $\mathcal{M}$ has exactly one distinct eigenvalue, then* $\dim_{\mathbb{C}}(\mathcal{M}) = \lfloor \frac{n^2}{4} \rfloor + 1$.

Proof. We are assuming, without loss of generality, that $\mathcal{M}$ contains the scalar matrices E_n (otherwise we could add them to $\mathcal{M}$ and get a bigger algebra). Let $A' \in \mathcal{M}$ and let λ be its (only) eigenvalue. Therefore by Theorem 10 we know that $A = A' - \lambda I$ is a nilpotent matrix. Therefore, we can write:

$$\mathcal{M} = E_n \oplus \mathcal{W}$$

where $\mathcal{W}$ is the algebra of all the matrices A in the construction explained above. The result now follows from applying Theorem 17 to $\mathcal{W}$. ■

4 Describing nil commutative algebras

We now move on to characterize the nil commutative algebra $\mathcal{W}$ described. In doing so we shall prove Schur's Theorem 4 and get a better handle on the smaller cases of the theorems. Note that we retain all the notation and results we obtained in the previous section.

Proof. (Schur's Theorem 4) We have shown in Remark 16 that nil $\mathbb{C}$-algebras of $n \times n$ matrices of dimension $\lfloor \frac{n^2}{4} \rfloor$ exist; we now intend to prove that any such algebra is equivalent to $\mathcal{A}_n$, $\mathcal{A}'_n$ or $\mathcal{A}''_3$, depending on n. As mentioned, we retain all the notation and definitions from the proof of Theorem 17; in particular, we still assume that $\mathcal{W}$ is a commutative nil $\mathbb{C}$-algebra of maximum dimension (which we now know is $\lfloor \frac{n^2}{4} \rfloor$), that $\mathcal{W}'$ is the image of $\mathcal{W}$ under the function $M \mapsto PMP^{-1}$, and that $\mathcal{B}$ and $\mathcal{C}$ are the matrix algebras as previously discussed.

If $\lfloor \frac{s^2}{4} \rfloor > 0$, then since $\mathcal{B}$ is a nil algebra, by Cartan's Theorem 11, a matrix $B' \in \mathcal{B}$ exists such that $B'X = XB' = 0$ for every $X \in \mathcal{B}$. Now, we notice that if we take two matrices M_1 and M_2 that are spanned by the first t matrices of the basis $\mathcal{X}$, then they must only be non-zero in their B and C blocks. Now if we multiply them:

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_1 & 0 \\ C_1 & 0 & 0 \end{pmatrix},\ M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_2 & 0 \\ C_2 & 0 & 0 \end{pmatrix} \Rightarrow M_1M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_1B_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

This means that the vector space of dimension $t = b + c = \lfloor \frac{s^2}{4} \rfloor + r^2$ spanned by

$$\{X_1, ..., X_b, X_{b+1}, ..., X_t\}$$

is an algebra; we denote this algebra by $\mathcal{T}$. It is a commutative algebra, because $\mathcal{B}$ is a commutative algebra:

$$M_1M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_1B_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B_2B_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = M_2M_1$$

Therefore, by constructing the non-zero matrix $M' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B' & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{T}$ we see that $M'X = XM' = 0$ for every $X \in \mathcal{T}$; especially, for the basis $\{X_1, ..., X_t\}$. We claim that for $t < i \leq m$ we have that $M'X_i = X_iM' = 0$. Indeed, it is a matter of multiplication:

$$X_i = \begin{pmatrix} 0 & 0 & 0 \\ A_i & 0 & 0 \\ 0 & D_i & 0 \end{pmatrix} \xRightarrow{M' \in \mathcal{W}'} \begin{pmatrix} 0 & 0 & 0 \\ B'A_i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = M'X_i = X_iM' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & D_iB' & 0 \end{pmatrix}$$

By comparing both sides of the equations we see that $B'A_i = 0$ and $D_iB' = 0$; therefore, $M'X_i = 0 = X_iM'$. So we now have two matrices in $\mathcal{W}'$ that annihilate every element of it: the matrix $M'' := PMP^{-1}$ which we defined in the beginning of the proof of Theorem 17, and M'. We can easily see that any linear combination of M' and M'' also shares this property with them. Notice that Cartan's Theorem 11 tells us that $B' \neq 0$; this implies that $\operatorname{rank}(B') > 0$. Therefore if $\lfloor \frac{s^2}{4} \rfloor > 0$, we have

$$\operatorname{rank}(M'' + M') = \operatorname{rank} \begin{pmatrix} 0 & 0 & 0 \\ 0 & B' & 0 \\ I_r & 0 & 0 \end{pmatrix} = \operatorname{rank}(B') + r > r = \operatorname{rank}(M'') \tag{4.1}$$

However, we chose M'' in the proof of Theorem 17 such that $\operatorname{rank}(M'')$ was the maximum amount among the matrices that Cartan's Theorem describes; this is a contradiction. So $\lfloor \frac{s^2}{4} \rfloor = 0$, or equivalently, s is either 0 or 1. We examine these cases separately.

First if $s = 0$, then every $X \in \mathcal{W}'$ is of the form $\begin{pmatrix} 0 & 0 \\ C_{r\times r} & 0 \end{pmatrix}$ and therefore $\mathcal{W}'$ is equal (hence equivalent) to $\mathcal{A}_n = \mathcal{A}_{2r}$. For the other case of $s = 1$, we conclude $n = 2r + 1$. Notice that A_i (the A block of X_i) must be of the form:

$$\begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,r} \end{pmatrix} \tag{4.2}$$

and D_j (the D block of X_j) must be of the form:

$$D_j = \begin{pmatrix} d_{j,1} \\ d_{j,2} \\ \vdots \\ d_{j,r} \end{pmatrix} \tag{4.3}$$

We defined k to be the rank of the matrix K as defined in the proof of Theorem 17, and we defined $l = s - k$. Since $k + l = s = 1$, we have two other subcases: first assume that $k = 1$ and $l = 0$. By using the commutativity of X_i and X_j, we can see:

$$X_iX_j = X_jX_i \Rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ D_iA_j & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ D_jA_i & 0 & 0 \end{pmatrix} \Rightarrow D_iA_j = D_jA_i$$

$$\Rightarrow d_{i,\beta}a_{j,\alpha} = d_{i,\alpha}a_{j,\beta} \ (i, j \in \{c+1, ..., m\},\ \alpha, \beta \in \{1, ..., r\}) \tag{4.4}$$

Assume that $r > 1$ (equivalently, $n > 3$). Now we know that an invertible 2×2 matrix of the form $\begin{pmatrix} a_{i,\alpha} & a_{i,\beta} \\ a_{j,\alpha} & a_{j,\beta} \end{pmatrix}$ exists; because if not, every A_i will be a multiple of the other rows which is in contradiction with the linear independence of the A_i matrices. Without loss of generality, let this matrix be the following submatrix:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{4.5}$$

So for $\alpha = 1, 2, \ldots, r$ we have:

$$d_{1,\alpha}a_{21} = d_{2\alpha}a_{11}, \quad d_{1,\alpha}a_{22} = d_{2,\alpha}a_{12} \tag{4.6}$$

$$\Rightarrow \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} -d_{2,\alpha} \\ d_{1,\alpha} \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} -d_{2,\alpha} \\ d_{1,\alpha} \end{pmatrix} = 0$$

therefore, $d_{1,\alpha} = d_{2,\alpha} = 0$, and we therefore conclude that $D_1 = D_2 = 0$.

Thanks to (4.4) for every positive integer m, we know that $D_mA_1 = D_1A_m$, therefore for every m, $D_mA_1 = 0$; since we can't have every entry of A_1 be equal to zero, we must necessarily have that $D_m = 0$. This means that, like the $s = 0$ case, the form of every $X \in \mathcal{W}'$ collapses to a matrix in $\mathcal{A}_n$ and that ensures that $\mathcal{W}'$ is equivalent to $\mathcal{A}_n$. The other subcase of $k = 0$, $l = 1$ can be handled similarly to show that $\mathcal{W}'$ is equivalent to $\mathcal{A}_n^{'}$.

So at last, we need to look at the case $r = 1$ (equivalently $n = 3$) separately. For $n = 3$, if $k = 1, l = 0$ and D_1 is not zero (otherwise we would fall back into a case similar to the previous ones and $\mathcal{W}'$ would be equivalent to $\mathcal{A}_3$ or $\mathcal{A}_3'$), then the basis $\mathcal{X}$ must be of the form:

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ x & 0 & 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 0 & 0 \\ y & 0 & 0 \\ 0 & z & 0 \end{pmatrix} \tag{4.7}$$

where x, y, z are non-zero elements of $\mathbb{C}$; by defining $R := \begin{pmatrix} yz & 0 & 0 \\ 0 & z & 0 \\ 0 & 0 & 1 \end{pmatrix}$ we see that:

$$RX_1R^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ xy^{-1}z^{-1} & 0 & 0 \end{pmatrix}, \quad RX_2R^{-1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

As we can see, the algebra spanned by RX_1R^{-1} and RY_1R^{-1} together is exactly equal to $\mathcal{A}_3''$. Hence $\mathcal{W}'$ is equivalent to $\mathcal{A}_3''$ in this case.

Now, the proof is complete: the nilpotent algebra $\mathcal{W}$ was shown to be equivalent to $\mathcal{W}'$, and that in turn is equivalent to the algebras $\mathcal{A}_n$, $\mathcal{A}_n'$ and $\mathcal{A}_3''$ based on the different cases for n. ■

5 Proof of Theorem 1 and 2

We just proved Schur's Theorem 4 where we classified the nil commutative $\mathbb{C}$-algebras of $n \times n$ matrices with maximum dimension. Now, we return to Schur's first two theorems and finish the proofs off:

Proof. (of theorems 1 and 2) Suppose $\mathcal{V}$ is a commutative $\mathbb{C}$-algebra of matrices of size n with maximum dimension $m+1$. Because of Proposition 15 we can suppose that $m+1 \geq \lfloor \frac{n^2}{4} \rfloor + 1$. We use strong induction on n, knowing that the base case of $n = 1$ is trivial. If every matrix in $\mathcal{V}$ has only one distinct eigenvalue then by Corollary 18 we know that it has dimension at most $v_n + 1 = \lfloor \frac{n^2}{4} \rfloor + 1$. So suppose otherwise; that we can choose $A \in \mathcal{V}$ such that A has at least two distinct eigenvalues. Therefore, we can find square matrices G and H of sizes g and h and an invertible matrix $P \in M_n(\mathbb{C})$ such that $g + h = n$, G and H share no eigenvalues, and $PAP^{-1} = \begin{pmatrix} G & 0 \\ 0 & H \end{pmatrix}$. Now by Lemma 6 we see that since $\mathcal{V}$ is a commuting family of matrices, every matrix in $\mathcal{V}' := P\mathcal{V}P^{-1}$ is also of the same matrix block form. Therefore, by collecting all the $g \times g$ upper-left blocks in $\mathcal{V}'$ into a set $\mathcal{G}$ and doing the same for the lower-right $h \times h$ blocks and naming the resulting set $\mathcal{H}$, we can see that each of $\mathcal{G}$ and $\mathcal{H}$ are commutative $\mathbb{C}$-algebras of $g \times g$ and $h \times h$ matrices, respectively. By our induction hypothesis we know that $\dim_{\mathbb{C}}\mathcal{G} \leq \lfloor \frac{g^2}{4} \rfloor + 1$ and $\dim_{\mathbb{C}}\mathcal{H} \leq \lfloor \frac{h^2}{4} \rfloor + 1$. Since the dimension of $\mathcal{V}'$ is at most the sum of these two dimensions, putting this together with $g + h = n$ we see that:

$$\lfloor \frac{(g+h)^2}{4} \rfloor + 1 = \lfloor \frac{n^2}{4} \rfloor + 1 \leq m + 1 \leq \lfloor \frac{g^2}{4} \rfloor + 1 + \lfloor \frac{h^2}{4} \rfloor + 1 \tag{5.1}$$

Note that (5.1) is equivalent to $gh \leq 2$. If $n > 3$ we have that $gh \geq n - 1 > 2$ and we reach a contradiction. Therefore when $n > 3$, the maximum of v_n can only be reached when all matrices in $\mathcal{V}$ have a single distinct eigenvalue, which gives us the maximum of $v_n = \lfloor \frac{n^2}{4} \rfloor$ as Corollary 18 states. If $n = 2$ or $n = 3$, the inequalities in (5.1) all become equalities, and therefore the maximum of v_n can still only be $\lfloor \frac{n^2}{4} \rfloor$. Either way, our proof by induction is complete and we prove Theorems 1 and 2 in their entirety. ■

Theorem 1 is now complete; for $n \geq 4$ we also have a complete characterization of the $\mathbb{C}$-algebras that attain the upper bound of the theorem, thanks to Theorems 2 and 4. As mentioned before, $n = 1$ is a trivial case. So as our last result, we characterize the algebras attaining the maximum in question for $n = 2$ and 3. Theorems 2 and 4 still give us a glimpse of some possible cases, and the proof just given fills all the holes of our proof:

Proposition 19. (a) *For $n = 2$, every commutative $\mathbb{C}$-algebra of dimension $v_2 + 1 = 2$ is equivalent with one of the two following algebras (as before, E_n is the algebra of all scalar $n \times n$ matrices):*

$$E_2 \oplus \mathcal{A}_2 \quad \textbf{or} \quad \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \middle| a, b \in \mathbb{C} \right\}$$

(b) *For $n = 3$, every commutative $\mathbb{C}$-algebra of dimension $v_3 + 1 = 3$ is equivalent with either $E_3 \oplus \mathcal{W}$ where $\mathcal{W} \in \{\mathcal{A}_3, \mathcal{A}_3', \mathcal{A}_3''\}$ as described in Theorem 4, or is equivalent to one of the two following algebras:*

$$\left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \middle| a, b, c \in \mathbb{C} \right\} \quad \textbf{or} \quad \left\{ \begin{pmatrix} a & 0 & 0 \\ b & a & 0 \\ 0 & 0 & c \end{pmatrix} \middle| a, b, c \in \mathbb{C} \right\}$$

Proof. We retain the notation and definitions of the proof of Theorems 1 and 2. We name our algebra $\mathcal{M}$ in both cases. For (a), we know that $E_2 \subset \mathcal{M}$ (otherwise we reach a contradiction with the maximality of $\mathcal{M}$). So we can assume that I_2 (the 2×2 identity matrix) and another matrix $A \in \mathcal{M}$ form a basis for $\mathcal{M}$. By Theorem 10, we can uniquely write A as $A = N + D$ where N is nilpotent and D is a diagonal matrix with the same eigenvalues as A. Since $\operatorname{rank}(N) < 2$, we have two cases:

1) either $\operatorname{rank}(N) = 0$ and therefore $N = 0_{2\times 2}$, which means $A = D$ is a (nonzero) diagonal matrix; A and I_2 therefore are a basis for all 2×2 diagonal matrices, which is the second case of the statement of part (a).

2) or $\operatorname{rank}(N) = 1$ which means that N is a non-zero 2×2 nilpotent matrix. Let $N = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; since for a $k \times k$ nilpotent matrix T we necessarily know that $T^k = 0$, we can conclude that $N^2 = 0$ and therefore we know that $\operatorname{tr}(N) = a + d = 0$ and $\det(N) = ad - bc = 0$. So we conclude that $d = -a$ and $bc = -a^2$. Now if $a = 0$ then only one of b or c is non-zero and therefore N is either in $\mathcal{A}_2$ or is similar to a matrix in it by a permutation of rows and columns. If $a \neq 0$ then we can see the matrix $P = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$ gives us the relation $PNP^{-1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. In both cases, N is similar to the matrix $M := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$; therefore the algebra $P\mathcal{M}P^{-1}$ has a basis of I_2 and M, the algebra spanned by which is exactly $E_2 \oplus \mathcal{A}_2$, resulting in the first case of the statement of part (a).

For part (b) we see that by the proof of Theorems 1 and 2 that either $\mathcal{M}$ has only matrices of exactly one distinct eigenvalue, which results in it being equivalent to the three cases covered by Theorems 2 and 4, or it is similar to the matrix $\begin{pmatrix} G_{2\times 2} & 0 \\ 0 & c \end{pmatrix}$ for some $c \in \mathbb{C}$. Now by part (a) we see that G is either a diagonal matrix or is in the algebra $E_2 \oplus \mathcal{A}_2$; both of these cases result in the two remaining cases in the statement of part (b). ∎

6 Minimum dimension of maximal commutative subalgebras of $M_n(\mathbb{C})$

Let $\mathcal{M} \subseteq M_n(\mathbb{C})$ be a maximal commutative subalgebra of $M_n(\mathbb{C})$; maximal in the sense that if $\mathcal{M} \subseteq \mathcal{N} \subseteq M_n(\mathbb{C})$ and $\mathcal{N}$ is also a commutative subalgebra of $M_n(\mathbb{C})$ then necessarily $\mathcal{N} = \mathcal{M}$. By Schur's Theorem 1 we know that $\dim_{\mathbb{C}}(\mathcal{M}) \leq v_n + 1 = \lfloor \frac{n^2}{4} \rfloor + 1$. We can ask whether these maximal subalgebras necessarily attain the maximum dimension of $v_n + 1$. In this section we show that this is not true. To be more precise, let m_n be the minimum possible dimension of a maximal commutative subalgebra of $M_n(\mathbb{C})$. We show using two examples that $m_n \leq n$; we therefore conclude that for $n \geq 4$, the maximum possible dimension is not necessarily attained.

Remark 20. In this section if $\mathcal{A} \subseteq M_n(\mathbb{C})$ is a maximal commutative matrix algebra, by maximality we know that $I_n \in \mathcal{A}$; therefore all our algebras in this section are supposed to be unital. When viewed as a ring, whenever we say the algebra $\mathcal{A} \subseteq M_n(\mathbb{C})$ is generated by $\{A_1, ..., A_k\}$ we can see that I_n is also generated and is an element of $\mathcal{A}$. However when we talk about a linearly-independent basis of $\mathcal{A}$ we must take care to note that a linear combination of $A_1, ..., A_k$ might be equal to I_n. In order to avoid confusion, whenever we talk about such a linearly independent generating set we can safely assume that I_n is also in this set.

Example 21. We recursively construct a $\mathbb{C}$-algebra of commutative $n \times n$ matrices that is maximal of dimension n. For $n \leq 3$ we have the obvious choices from Proposition 19, since $n = v_n + 1$ for these values of n; we define $\mathcal{D}_n$ to be these maximal algebras. Now for $n \geq 4$ we claim that:

$$\mathcal{D}_n := \begin{pmatrix} \mathcal{D}_{n-2} & 0 \\ 0 & \mathcal{D}_2 \end{pmatrix} = \left\{ \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \middle| P_1 \in \mathcal{D}_{n-2},\ P_2 \in \mathcal{D}_2 \right\} \tag{6.1}$$

is a maximal commutative $\mathbb{C}$-algebra of dimension n. Note that $\dim_{\mathbb{C}}(\mathcal{D}_n) = \dim_{\mathbb{C}}(\mathcal{D}_{n-2}) + \dim_{\mathbb{C}}(\mathcal{D}_2) = (n-2) + 2 = n$. It is easy to see that $\mathcal{D}_n$ is a commutative subalgebra of $M_n(\mathbb{C})$. To show $\mathcal{D}_n$ is maximal, let $X \in M_n(\mathbb{C})$ be a matrix that commutes with every element of $\mathcal{D}_n$; take $Y = \begin{pmatrix} I_{n-2} & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{D}_n$. By matrix multiplication we have:

$$X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}, XY = YX \Rightarrow X_3 = 0, X_2 = 0 \tag{6.2}$$

Now by taking $Z = \begin{pmatrix} Z_1 & 0 \\ 0 & Z_2 \end{pmatrix}$ such that $Z_1 \in \mathcal{D}_{n-2}$ and $Z_2 \in \mathcal{D}_2$, since we want X to commute with Z we see that $X_1 Z_1 = Z_1 X_1$ and $X_2 Z_2 = Z_2 X_2$ and since $\mathcal{D}_{n-2}$ and $\mathcal{D}_2$ are maximal, we conclude that $X_1 \in \mathcal{D}_{n-2}$ and $X_2 \in \mathcal{D}_2$. So $X \in D_n$ and therefore $\mathcal{D}_n$ is maximal. So using induction we have proven that the recursively-defined algebras $\mathcal{D}_n$ are maximal commutative $\mathbb{C}$-algebras of $n \times n$ matrices of dimension n.

Example 22. As usual, for all $1 \leq i, j \leq n$ let $E_{i,j} \in M_n(\mathbb{C})$ be the matrix which has zero entries everywhere except for the (i,j)-th entry which is equal to 1. We know that $E_{i,j}E_{k,l} = \delta_{j,k}E_{i,l}$. Let $\mathcal{C}_n$ be the algebra generated by the identity matrix I_n and $E_{2,1}, ..., E_{n,1}$. We see that $E_{i,1}E_{j,1} = E_{j,1}E_{i,1} = 0$ for any $2 \leq i, j \leq n$; therefore, $\mathcal{C}_n$ is a commutative subalgebra of $M_n(\mathbb{C})$. Also we see that $\dim_{\mathbb{C}}(\mathcal{C}_n) = n$. We claim that $\mathcal{C}_n$ is also maximal for all positive integers n.

To prove our claim, let $B = [b_{i,j}] \in M_n(\mathbb{C})$ be a matrix of size n that commutes with every matrix in $\mathcal{C}_n$. As the matrices $E_{i,j}$ constitute a basis for $M_n(\mathbb{C})$, we can write $B = \sum_{1 \leq i,j \leq n} b_{i,j}E_{i,j}$. Since B obviously commutes with I_n, by fixing $k \in \{2, ..., n\}$, we need to ensure B commutes with $E_{k,1}$. We compute:

$$BE_{k,1} = \sum_{1 \leq i,j \leq n} b_{i,j}E_{i,j}E_{k,1} = \sum_{1 \leq i,j \leq n} b_{i,j}\delta_{j,k}E_{i,1} = \sum_{1 \leq i \leq n} b_{i,k}E_{i,1}$$

$$E_{k,1}B = \sum_{1 \leq i,j \leq n} b_{i,j}E_{k,1}E_{i,j} = \sum_{1 \leq i,j \leq n} b_{i,j}\delta_{1,i}E_{k,j} = \sum_{1 \leq j \leq n} b_{1,j}E_{k,j}$$

$$BE_{k,1} = E_{k,1}B \Rightarrow \sum_{1 \leq i \leq n} b_{i,k}E_{i,1} = \sum_{1 \leq j \leq n} b_{1,j}E_{k,j}$$

Since the matrices $E_{i,j}$ are a basis for $M_n(\mathbb{C})$, by comparing coefficients on both sides of the equation we see that $b_{k,k} = b_{1,1}$, and for every $i \neq k$ we have $b_{i,k} = 0$, and for every $j \geq 2$ we have $b_{1,j} = 0$. So, the first row of B is zero except for $b_{1,1}$, and the k-th column of B is zero except for $b_{k,k}$ which is equal to $b_{1,1}$. This is true for every $k \in \{2, ..., n\}$; therefore in every column other than the first, the only non-zero entries are the entries on the main diagonal, which are all equal to $b_{1,1}$. Therefore we can see that:

$$B = b_{1,1}I_n + b_{2,1}E_{2,1} + ... + b_{n,1}E_{n,1}$$

Therefore we can see that $B \in \mathcal{C}_n$ by definition; so $\mathcal{C}_n$ is a maximal commutative subalgebra of $M_n(\mathbb{C})$ of dimension n.

By Examples 21 and 22 we see that for every positive integer n we have $m_n \leq n$. Now we can answer when maximal commutative subalgebras of $M_n(\mathbb{C})$ with a dimension less than $v_n + 1$ exist. To complete its proof, we need to first prove a lemma:

Lemma 23. *Let $\mathcal{A}$ be a maximal commutative subalgebra of $M_n(\mathbb{C})$ which is generated, as an algebra, by the matrix I_n and $A \in M_n(\mathbb{C})$ such that A is not a scalar matrix. Then, we have that* $\dim_{\mathbb{C}}(\mathcal{A}) = n$.

Proof. By the Cayley-Hamilton theorem we know that A^n can be written as a linear combination of $I_n, A, A^2, ..., A^{n-1}$. Therefore since every matrix in $\mathcal{A}$ is a polynomial in A, the algebra $\mathcal{A}$ can have at most n linearly independent elements; this means that $\dim_{\mathbb{C}}(\mathcal{A}) \leq n$. We now need to show that $\dim_{\mathbb{C}}(\mathcal{A}) \geq n$.

Since $\mathbb{C}$ is an algebraically closed field, therefore we know we can find an invertible matrix $P \in M_n(\mathbb{C})$ such that $B := PAP^{-1}$ is upper-triangular. We name $\mathcal{B} := P\mathcal{A}P^{-1}$; it follows that $\mathcal{B}$ is a maximal commutative subalgebra of $M_n(\mathbb{C})$ and that it is generated as an algebra by B. We claim that there are at least n linearly independent matrices in $\mathcal{B}$.

If $C = [c_{i,j}] \in M_n(\mathbb{C})$ is an upper-triangular matrix, having B and C commute will ensure that C commutes with everything in $\mathcal{B}$. If $BC = CB$, then we have $\frac{n^2+n}{2}$ unknowns (the entries of C) and $\frac{n^2+n}{2}$ linear equations (by comparing every non-zero entry in BC and CB). However, the entries on the main diagonal of BC and CB are automatically equal; so we can essentially disregard these n equations, leaving us with $\frac{n^2-n}{2}$ non-trivial equations to be satisfied. We now formulate the situation as a set of linear equations $MX = 0$ where M is a $\frac{n^2-n}{2} \times \frac{n^2+n}{2}$ matrix with the coefficients of the linear equations in each row, and

$$X = \begin{pmatrix} c_{1,1} \\ c_{1,2} \\ \vdots \\ c_{1,n} \\ c_{2,2} \\ \vdots \\ c_{n,n} \end{pmatrix}$$

Let $W \subseteq \mathbb{C}^{\frac{n^2+n}{2}}$ be the solution space of this set of equations. We know that $\operatorname{rank}(M) \leq \min(\frac{n^2-n}{2}, \frac{n^2+n}{2}) = \frac{n^2-n}{2}$; therefore by the rank-nullity theorem we have that $t := \dim_{\mathbb{C}}(W) = \operatorname{null}(M) = \frac{n^2+n}{2} - \operatorname{rank}(M) \geq \frac{n^2+n}{2} - \frac{n^2-n}{2} = n$, so we know $t \geq n$. Let $X_1, ..., X_t \in \mathbb{C}^{\frac{n^2+n}{2}}$ be a basis for W. By placing the entries of each of the X_i's in an upper-triangular matrix in the correct order, we obtain t linearly independent matrices $C_1, ..., C_t$ that all commute with B. Since $\mathcal{B}$ was assumed to be maximal, we must necessarily have that $C_i \in \mathcal{B}$ for all $i \in \{1, ..., t\}$. This means that because of the independence of the C_i matrices, we must have $\dim_{\mathbb{C}}(\mathcal{B}) \geq t \geq n$; so the dimension of $\mathcal{B}$ is at least n and since we showed it must be at most n, we can conclude that $\dim_{\mathbb{C}}(\mathcal{B}) = \dim_{\mathbb{C}}(\mathcal{A}) = n$. ∎

Now we are ready to show that maximality and maximum dimension don't coincide for large enough n:

Proposition 24. *Let n be a positive integer. For $n \in \{2, 3\}$, we have that $m_n = v_n + 1$ and therefore every maximal commutative subalgebra of $M_n(\mathbb{C})$ necessarily has the maximum possible dimension; for $n \geq 4$, we can find maximal commutative subalgebras of $M_n(\mathbb{C})$ with dimension less than $v_n + 1$.*

Proof. Let $\mathcal{A}$ be a maximal commutative subalgebra of $M_n(\mathbb{C})$. For $n = 2$, we know that $\mathcal{A}$ contains I_n and at least one non-scalar matrix A (otherwise we could add it to $\mathcal{A}$ and generate a bigger commutative algebra). So by Schur's Theorem 1, we have $2 \leq \dim_{\mathbb{C}}(\mathcal{A}) \leq v_2 + 1 = 2$ and therefore $\mathcal{A}$ has dimension 2 which is the maximum possible amount.

For $n = 3$ suppose $\mathcal{A}$ needs at least t linearly independent elements other than I_n to generate it as an algebra. If $t \geq 2$, then we know that $\dim_{\mathbb{C}}(\mathcal{A}) \geq t + 1 \geq 3$; since by Schur's Theorem 1 we know $\dim_{\mathbb{C}}(\mathcal{A}) \leq v_3 + 1 = 3$, therefore $\dim_{\mathbb{C}}(\mathcal{A}) = 3$. If $t = 1$, Lemma 23 shows us that $\dim_{\mathbb{C}}(\mathcal{A}) = 3$. So in either case, for $n = 3$ we conclude that $\dim_{\mathbb{C}}(\mathcal{A})$ has to be the maximum value possible.

Now let $n \geq 4$. Note that Examples 21 and 22 give us maximal commutative subalgebras of $M_n(\mathbb{C})$ of dimension n. If $n = 2k$ is even we see that $n = 2k < k^2 + 1 = \lfloor \frac{n^2}{4} \rfloor + 1 = v_n + 1$ (since $n \geq 4$ means $k \geq 2$) and if $n = 2k + 1$ is odd (so $k \geq 2$), we can see that $n = 2k + 1 < k^2 + k + 1 = \lfloor \frac{n^2}{4} \rfloor + 1 = v_n + 1$; therefore in both cases, we have examples of maximal commutative subalgebras of $M_n(\mathbb{C})$ whose dimensions do not attain the upper bound of $v_n + 1$. ■

References

[1] E. Cartan. Les groupes bilinéaires et les systèmes de nombres complexes. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 12(2):B65–B99, 1898.

[2] G. Frobenius. Theorie der hyperkomplexen größen : Ii. In *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin : Jahrgang 1903 ; Erster Halbband Januar bis Juni*. 2024.

[3] K. Hoffman and R. Kunze. *Linear Algebra*. Prentice Hall, 1971.

[4] J. Schur. Zur theorie der vertauschbaren matrizen. *Journal für die reine und angewandte Mathematik*, 130:66–76, 1905.