On Commutativity and Lie nilpotency in Matrix Algebras

by

Mahlare Gerald Sehoana

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science (Mathematics) in the Faculty of Science at Stellenbosch University

Supervisor: Prof. L. van Wyk

December 2015

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: ....2015/09/30..........

Abstract

On Commutativity and Lie nilpotency in Matrix Algebras

M.G. Sehoana

Department of Mathematical Sciences,
Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MSc (Mathematics)

September 2015

In this thesis we first discuss the proof by Mirzakhani [9] of Schur's Theorem which gives the maximum number of linearly independent matrices in a commutative algebra of $n \times n$ matrices over a field F. An example illustrating the application of Schur's Theorem is given.

Secondly, we discuss the Cayley-Hamilton Theorem which asserts that any $n \times n$ matrix A satisfies its characteristic polynomial. A deduction of a Cayley-Hamilton trace identity for a 2×2 matrix A over a commutative ring from the Cayley-Hamilton Theorem is shown. We then discuss the Cayley-Hamilton trace identity for any matrix $A \in M_2(R)$ when
(i) R is commutative,
(ii) R is not necessarily commutative,
(iii) R is not necessarily commutative and $\mathrm{tr}(A) = 0$,
(iv) R is not necessarily commutative and satisfies the identity $[[x, y], [x, z]] = 0$.

Lastly, we discuss the matrix algebras $U_n^*(R)$, in particular the matrix algebras $U_3^*(R)$ and $U_4^*(R)$, in relation to polynomial identities $[[\ldots[[x_1, x_2], x_3], \ldots], x_n] = 0$, $[x, y][w, z] = 0$ and $[[x, y], [w, z]] = 0$.

Uittreksel

Oor Kommutatiwiteit en Lie nilpotensie in Matriksalgebras

("On Commutativity and Lie nilpotency in Matrix Algebras")

M.G. Sehoana

Departement Wiskundige Wetenskappe,
Universiteit Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MSc (Wiskunde)

September 2015

In hierdie tesis beskryf ons eerstens die bewys deur Mirzakhani [9] van Schur se Stelling wat die maksimum aantal lineêr onafhanklike matrikse in 'n kommutatiewe algebra van $n \times n$ matrikse oor 'n liggaam F gee. 'n Voorbeeld word gegee wat die toepassing van Schur se Stelling illustreer.

Tweedens bespreek ons die Cayley-Hamilton Stelling wat beweer dat elke $n \times n$ matriks A sy karakteristieke polinoom bevredig. 'n Afleiding van 'n Cayley-Hamilton spoor identiteit vir 'n 2×2 matriks A oor 'n kommutatiewe ring vanuit die Cayley-Hamilton Stelling word gegee. Ons bespreek dan die Cayley-Hamilton spoor identiteit vir enige matriks $A \in M_2(R)$ wanneer
(i) R kommutatief is,
(ii) R nie noodwendig kommutatief is nie,
(iii) R nie noodwendig kommutatief is nie en $\text{sp}(A) = 0$,
(iv) R nie noodwendig kommutatief is nie en die identiteit $[[x, y], [x, z]] = 0$ bevredig.

Laastens bespreek ons die matriksalgebras $U_n^*(R)$, in besonder die matriksalgebras $U_3^*(R)$ en $U_4^*(R)$, met betrekking tot die polinoom identiteite $[[\ldots[[x_1, x_2], x_3], \ldots], x_n] = 0$, $[x, y][w, z] = 0$ en $[[x, y], [w, z]] = 0$.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations:

◇ Prof Leon van Wyk for the guidance, direction and enlightenment throughout my studies. His patience and constructive comments kept my hopes alive. He brought back my love for Mathematics.

◇ Prof David Kubayi for getting me started on this project and for the support that, together with the Dean of FSA at UL, provided during his time as HOD.

◇ My family for their moral support and motivation.

◇ Department of Mathematical Sciences at Stellenbosch University for allowing me the opportunity to study for an MSc degree.

◇ Fellow students Sandile Mkhaliphi and Andry Rabenantoandro for assisting me with Latex.

◇ Colleagues in the school of Mathematical and Computer Sciences at the University of Limpopo for their encouragement.

◇ DST-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-Mass) for their financial support.

◇ Above all, I thank God our Lord, he made it possible.

Dedications

This thesis is dedicated to my grandparents.

Contents

Chapter 1

Introduction

This chapter is mainly a brief background and overview of the subsequent Chapters 2 to 5. Also in this chapter we discuss some of the concepts which we have used a number of times in this thesis. Where possible, we have supplied examples to substantiate claims made. Chapter 5 can be viewed as a continuation of Chapter 3. Though in Chapter 3 we deal with matrices over the field $\mathbb{R}$ of real numbers and in Chapter 5 we deal with matrices over an arbitrary ring R. Coincidentally, Chapters 2 and 4 deal with upper triangular matrices. In Chapter 2 we encounter upper triangular matrices which mutually commute whereas in Chapter 4 commutativity of the said matrices is not necessary.

A binary operation $*$ on a set A is said to be commutative if and only if $x * y = y * x$ for all $x, y \in A$. We recall that a ring R is an algebraic structure with two binary operations called addition and multiplication. One of the axioms of a ring R states that addition is commutative. However, in a ring R multiplication is not necessarily commutative. Thus, if multiplication is commutative such a ring is called a commutative ring. A nonempty subset B of a ring R is said to be a subring of R if B is itself a ring with respect to the operations of addition and multiplication in R. For a nonempty subset B of a ring R to be a subring it is sufficient that $ab \in B$ and $a - b \in B$ for all $a, b \in B$. The collection $M_n(R)$, where R is a ring, of all $n \times n$ matrices having elements of R as entries is a ring. For $n \geq 2$, $M_n(R)$ is not a commutative ring. But, for a commutative ring R, the ring

$$B = \left\{ \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \mid a, b \in R \right\}$$

is commutative and B is a subring of $M_2(R)$. For every ring R the trivial subring B=$\{0\}$ is commutative. Thus every ring has at least one subring which is commutative.

In Chapter 2 we shall encounter a maximal commutative subalgebra of $M_n(F)$ which we use to illustrate an application of Schur's Theorem. (A subring B of a ring R is said to be a maximal subring with respect to property Y if $B \neq R$ and there exists no subring C in R with the property Y such that $B \subset C \subset R$.) In fact, according to Schur's Theorem the number of linearly independent matrices in the subalgebra B above is $\lfloor 2^2/4 \rfloor + 1 = 2$. Moreover, we see that the number of elements in a basis for B is 2. A basis of a vector space V is a subset $W \subseteq V$ which is linearly independent and spans V. A basis of a vector space is a maximal linearly independent subset of that vector space.

The expression $\lfloor x \rfloor$ represents the greatest integer less than or equal to x. If x is an integer, we have $\lfloor x \rfloor = x$ and $\lfloor x + b \rfloor = x$ for $0 < b < 1$.

The characteristic polynomial of an $n \times n$ matrix A over a commutative ring R is defined to be

$$p(\lambda) = \det(A - \lambda I) = k_n \lambda^n + k_{n-1}\lambda^{n-1} + \ldots + k_1 \lambda + k_0,$$

$k_i \in R$. By the trace of a square matrix A, denoted $\operatorname{tr}(A)$, is meant the sum of the entries on the main diagonal of A, from the upper left to the lower right. The following properties of traces of any matrices $A, B \in M_n(R)$ are used in this thesis:
(i) $\operatorname{tr}(A \pm B) = \operatorname{tr}(A) \pm \operatorname{tr}(B)$,
(ii) $\operatorname{tr}(cA) = c \operatorname{tr}(A)$, where c is a constant.

The nontrivial coefficients of the characteristic polynomial of a matrix A can be expressed explicitly in terms of traces of powers of A, (see [15], [13]). In [13], it is shown that if $n = 4$, then

$$\begin{aligned} p(\lambda) &= \lambda^4 - T_1\lambda^3 + \frac{1}{2}\big(T_1^2 - T_2\big)\lambda^2 - \frac{1}{6}\big(T_1^3 - 3T_1T_2 + 2T_3\big)\lambda \\ &\quad + \frac{1}{24}\big(T_1^4 - 6T_1^2T_2 + 8T_1T_3 + 3T_2^2 - 6T_4\big) \\ &= \lambda^4 - \operatorname{tr}(A)\lambda^3 + \frac{1}{2}\big(\operatorname{tr}^2(A) - \operatorname{tr}(A^2)\big)\lambda^2 \\ &\quad - \frac{1}{6}\big(\operatorname{tr}^3(A) - 3\operatorname{tr}(A)\operatorname{tr}(A^2) + 2\operatorname{tr}(A^3)\big)\lambda \\ &\quad + \frac{1}{24}\big(\operatorname{tr}^4(A) - 6\operatorname{tr}^2(A)\operatorname{tr}(A^2) + 8\operatorname{tr}(A)\operatorname{tr}(A^3) + 3\operatorname{tr}^2(A^2) - 6\operatorname{tr}(A^4)\big), \end{aligned}$$

where $T_m = \operatorname{tr}(A^m)$. In this thesis we deal with the case when $n = 2$, that is,

$$p(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \frac{1}{2}\big(\operatorname{tr}^2(A) - \operatorname{tr}(A^2)\big)$$

and obviously replacing λ by A and introducing the identity matrix $I \in M_2(R)$ yields

$$p(A) = A^2 - \operatorname{tr}(A)A + \frac{1}{2}\big(\operatorname{tr}^2(A) - \operatorname{tr}(A^2)\big)I.$$

The equation just given leads to the Cayley-Hamilton trace identity for a 2×2 matrix which we discuss in Chapter 5 under various hypotheses with $\frac{1}{2} \in R$.

Definition 1. Let R be a ring (not necessarily commutative) and $a, b \in R$. An element of the form $[a, b] = ab - ba$ is called a commutator, more precisely, the commutator of a and b.

Definition 2. A ring R is called Lie nilpotent of index n $(n \geq 2)$ if R satisfies the identity

$$[[[\ldots[[x_1, x_2], x_3], \ldots], x_n], x_{n+1}] = 0$$

but not the identity

$$[[\ldots[[x_1, x_2], x_3], \ldots], x_n] = 0.$$

Definition 3. A ring R is called Lie nilpotent if it is Lie nilpotent of index n for some $n \geq 2$. (We note that a commutative ring may be called a "Lie nilpotent ring of index 1".)

The concepts of commutativity and Lie nilpotency are "somehow" related in the sense that one implies the other. Commutativity always implies Lie nilpotency and Lie nilpotency implies commutativity only if the concerned ring is of Lie nilpotent index 1. Thus, we can say commutativity is a stronger condition than Lie nilpotency. For example, in a commutative ring the identity $[x_1, x_2] = 0$ holds which in turn implies $[[x_1, x_2], x_3] = 0$, which in turn implies $[[[x_1, x_2], x_3], x_4] = 0$ and so on. We shall see in Proposition 33 that the ring $U_3^*(R)$ satisfies the hypothesis for Lie nilpotency but by Example 30, $U_3^*(R)$ is not commutative.

We further notice that in the algebras $U_n^*(R)$ we discuss in this thesis, R is required to be commutative in order to attain Lie nilpotency. This fact is observed in Theorem 44. It is shown in Example 45 that Lie nilpotency may not necessarily be attained in $U_n^*(R)$ if R is noncommutative. (In contrast, a subring $V_n^{**}(R) \subset U_n^*(R)$, with a noncommutative ring R, defined in Chapter 4 is Lie nilpotent of index 2, for all n. The subring $V_n^{**}(R)$ is actually a version of the subring $\mathfrak{F}_n$ given in Chapter 2 but with a noncommutative ring R.)

We see again the fundamental role of commutativity in Corollary 37 where an algebra $U_3^*(U_3^*(R))$, with noncommutative ring $U_3^*(R)$ and commutative R, satisfies the identity $[[x, y], [w, z]] = 0$ and does not satisfy either of the stronger identities $[[x, y], z] = 0$ and $[x, y][w, z] = 0$. On the other hand, the algebra $U_4^*(U_3^*(R))$ with commutative R does not satisfy $[[x, y], [w, z]] = 0$ but $[[[x, y], [w, z]], [[u, v], [r, s]]] = 0$.

For a commutative ring R, any product $[x_1, y_1][x_2, y_2]$ in $U_n^*(R)$, $n \leq 4$, is equal to zero and the discussion just after Remark 48 shows that for $n \geq 7$ a product $[x_1, y_1][x_2, y_2][x_3, y_3]$ may not necessarily be equal to zero. This leads to Theorems 52 (ii) and 53 which give a smallest value of k for which any product of the form

$$[x_1, y_1][x_2, y_2] \cdots [x_k, y_k] \tag{1.1}$$

in $U_n^*(R)$, R commutative, is equal to zero. This value of k depends on n. When R is not commutative, (1.1) may not necessarily be equal to zero for such values of k as shown in Remark 56 (i.e, there are matrices in $U_n^*(R)$ for which $[x_1, y_1][x_2, y_2] \cdots [x_k, y_k] \neq 0$).

Definition 4. Let R be an arbitrary ring. The matrix unit $E_{i,j}$ in $M_n(R)$ with $1 \leq i, j \leq n$, is defined to be the matrix with 1 in position (i, j) and zeros elsewhere.

Note that

$$E_{i,j}E_{k,l} = \begin{cases} E_{i,l} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Alternatively $E_{i,j}E_{k,l} = \delta_{j,k}E_{i,l}$, where

$$\delta_{j,k} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

The function $\delta_{j,k}$ is called the Kronecker delta. We observe that for $a, b \in R$,

$$(aE_{i,j})(bE_{k,l}) = (ab)E_{i,j}E_{k,l} = \begin{cases} (ab)E_{i,l} & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{1.2}$$

It should be noted that in all the examples that show that certain statements do not necessarily hold when R is noncommutative, we used specific matrices in the ring $U_n^*(R)$ because there may be matrices in $U_n^*(R)$ for which the statements hold. For example, if R is noncommutative, a product of the form (1.1) can be zero as in

$$[E_{1,2}, E_{1,3}][A_2, B_2] \cdots [A_k, B_k] = 0,$$

for all k and $E_{1,2}, E_{1,3}, A_k, B_k$ in $U_n^*(R)$. We have $[E_{1,2}, E_{1,3}] = 0$, according to Definition 1 and the observation that follows after Definition 4.

Lastly, at the end of Chapter 5 we discuss the proof of the fact that in a ring containing $\frac{1}{2}$ the identity $[[x, y], [x, z]] = 0$ implies the stronger identity $[[x, y], [w, z]] = 0$.

Chapter 2

Mirzakhani's Simple Proof of Schur's Theorem

For a field F, the ring $M_n(F)$ of $n \times n$ matrices with $n > 1$ and elements in F is noncommutative. But there are subalgebras $S \subset M_n(F)$ which are commutative. Theorem 10 in this chapter known as the Theorem of Schur deals with subalgebras of $M_n(F)$ of mutually commutative matrices. It gives the maximum number of linearly independent matrices in a maximal commutative subalgebra. Mirzakhani [9] provided a simpler proof, using induction, of this theorem which we dissect in this chapter. Mirzakhani [9] also provided an example of a commutative subalgebra satisfying the hypothesis of Theorem 10.

Definition 5. A set $X = \{x_1, x_2, \dots, x_n\}$ is said to be a spanning set of a vector space W over a field F if every $w \in W$ is a linear combination of the $x_i \in X$, i.e.,

$$w = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

where $a_i \in F$.

Definition 6. A set $V = \{v_1, v_2, \dots, v_r\}$ of non-zero vectors in a vector space over field F is said to be linearly independent whenever

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = 0$$

implies that $a_i = 0$ for all i, where $a_i \in F$.

Definition 7. Let $A = [A_1|A_2|\dots|A_n]$ be an $m \times n$ matrix over a field F, with the A_i's as columns of A. The rank of A is defined to be the maximum number of linearly independent vectors in the set $\{A_1, A_2, \dots, A_n\}$.

Definition 8. Let A be an $m \times n$ matrix over a field F. The null space of A is the set of all x in F^n such that $Ax = 0$. The dimension of the null space of A is called the nullity of A.

Remark 9. The nullity of a matrix A is the dimension of the solution space of $Ax = 0$, which is the same as the number of parameters in the general solution of $Ax = 0$ and which is the same as the number of free variables.

Theorem 10. *Let A be an $m \times n$ matrix. Then*

$$\text{rank}\, A + \text{nullity}\, A = n.$$

A submatrix of any given matrix A is a matrix which is obtained from A by removing any number of rows and/or columns. A matrix is said to be partitioned whenever it is divided into submatrices by drawing vertical and horizontal lines between its rows and columns.

Example 11. If

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix},$$

then a partitioning of A into submatrices B, C, D and E of order 3×2, $3\times(n-2)$, $(m-3)\times2$ and $(m-3)\times(n-2)$ respectively is

$$\left[\begin{array}{cc|ccc} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \hline a_{4,1} & a_{4,2} & a_{4,2} & \cdots & a_{4,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{array}\right] = \left[\begin{array}{c|c} B & C \\ \hline D & E \end{array}\right]$$

where B, C, D and E can be thought of as elements of A. We note that the partitioning of a matrix into submatrices is not unique.

Definition 12. A field F is said to be algebraically closed if every polynomial equation

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0$$

with coefficients in F has a solution in F.

The field $\mathbb{R}$ of real numbers is not algebraically closed as the equation $x^2 + 1 = 0$ does not have a solution in $\mathbb{R}$. But the field $\mathbb{C}$ of complex numbers is algebraically closed, (see[6]).

We state, without proof, the following theorem (which can be found in [4]) because of its significance in the proof of Theorem 14.

Theorem 13. *Let $\mathfrak{F}_n$ be a family of commuting matrices of order n over an algebraically closed field F. Then there exists a nonsingular matrix P of order n with entries in F such that $P^{-1}\mathfrak{F}_nP$ is a family of upper triangular matrices.*

Theorem 14. *The maximum number of mutually commuting linearly independent matrices of order n over a field F is $\lfloor n^2/4 \rfloor + 1$.*

Proof. The proof is by mathematical induction.

(i) For $n = 1$ the statement is obviously true.

(ii) Assume the theorem is true for $n-1$, i.e., a set consisting of mutually commuting matrices of order $(n-1)\times(n-1)$ has at most $\lfloor (n-1)^2/4 \rfloor + 1$ linearly independent matrices.

(iii) Let $\mathfrak{F}_n$ be a family of commuting matrices of order n over a field F. Suppose $\mathfrak{F}_n$ has more than $\lfloor n^2/4 \rfloor + 1$ linearly independent matrices. Assume, without loss of generality, that F is algebraically closed, then by Theorem 9 there exists a nonsingular matrix P with entries in F such that $P^{-1}\mathfrak{F}_n P$ is a family of upper triangular matrices. Also for any $T, U \in \mathfrak{F}_n$,

$$(P^{-1}UP)(P^{-1}TP) = P^{-1}U(PP^{-1})TP = P^{-1}(UT)P$$

and

$$(P^{-1}TP)(P^{-1}UP) = P^{-1}T(PP^{-1})UP = P^{-1}(TU)P.$$

But $UT = TU$ and thus $(P^{-1}UP)(P^{-1}TP) = (P^{-1}TP)(P^{-1}UP)$ showing that the matrices in $P^{-1}\mathfrak{F}_n P$ are mutually commuting. Let V be the vector space spanned by the set $P^{-1}\mathfrak{F}_n P$. Then V consists of upper triangular matrices which commute. Thus,

$$\boxed{P^{-1}\mathfrak{F}_\mathfrak{n} P \subseteq \operatorname{span}\{P^{-1}\mathfrak{F}_\mathfrak{n} P\} = V}$$

and $\dim V \geq \lfloor n^2/4 \rfloor + 2$. Since a subset of a linearly independent set is also linearly independent there exists a linearly independent subset $\{A_1, A_2, \ldots, A_q\}$ of a basis of V, $q = \lfloor n^2/4 \rfloor + 2$. Since each of the A_i's is an upper triangular matrix, the partitioning of each A_i into submatrices M_i, H_i and a zero matrix of order $(n-1)\times(n-1)$, $1 \times n$ and $(n-1)\times 1$ respectively, yields matrices of the form

$$A_i = \left[\begin{array}{c|c} \multicolumn{2}{c}{H_i} \\ \hline 0 & \\ 0 & M_i \\ \vdots & \\ 0 & \end{array}\right].$$

Now, partitioning the H_i further (where L_i, N_i and M_i are submatrices of order 1×1, $1\times(n-1)$ and $(n-1)\times(n-1)$ respectively) we have

$$A_iA_j = \left[\begin{array}{c|c} L_i & N_i \\ \hline 0 & M_i \end{array}\right]\left[\begin{array}{c|c} L_j & N_j \\ \hline 0 & M_j \end{array}\right] = \left[\begin{array}{c|c} L_iL_j & L_iN_j + N_iM_j \\ \hline 0 & M_iM_j \end{array}\right]$$

and

$$A_jA_i = \left[\begin{array}{c|c} L_j & N_j \\ \hline 0 & M_j \end{array}\right]\left[\begin{array}{c|c} L_i & N_i \\ \hline 0 & M_i \end{array}\right] = \left[\begin{array}{c|c} L_jL_i & L_jN_i + N_jM_i \\ \hline 0 & M_jM_i \end{array}\right].$$

Since $A_iA_j = A_jA_i$ it follows that $M_iM_j = M_jM_i$, i.e., the M_i's commute.

Let W be the vector space spanned by the set $\{M_1, M_2, \ldots, M_q\}$, so the elements of W are $(n-1) \times (n-1)$ matrices. Suppose $k = \dim W$. Then it follows from the induction hypothesis that $k \leq \lfloor (n-1)^2/4 \rfloor + 1$. Assume, without loss of generality, that W is spanned by the linearly independent set of matrices $\{M_1, M_2, \ldots, M_k\}$. Then for $i > k$ each M_i is expressible as a linear combination $M_i = \sum_{j=1}^{k} c_{i,j} M_j$ with scalars $c_{i,1}, \ldots, c_{i,k}$ in F. Now, for $i > k$, let $B_i = A_i - \sum_{j=1}^{k} c_{i,j} A_j$. Then for $i = k+1, \ldots, \lfloor n^2/4 \rfloor + 2$,

$$
\begin{aligned}
B_i &= A_i - \sum_{j=1}^{k} c_{i,j} A_j \\
&= \left[\begin{array}{c|c} \multicolumn{2}{c}{H_i} \\ \hline 0 & M_i \end{array}\right] - \left[\begin{array}{c|c} \multicolumn{2}{c}{\sum_{j=1}^{k} c_{i,j} H_j} \\ \hline 0 & \sum_{j=1}^{k} c_{i,j} M_j \end{array}\right] \\
&= \left[\begin{array}{c|c} \multicolumn{2}{c}{t_i} \\ \hline 0 & 0 \end{array}\right] \\
&= \left[\begin{array}{c} t_i \\ \hline 0 \end{array}\right]
\end{aligned}
$$

where $t_i = H_i - \sum_{j=1}^{k} c_{i,j} H_j$ is a $1 \times n$ matrix. Now consider

$$\sum_{i=k+1}^{q} d_i B_i = \sum_{i=k+1}^{q} d_i \left(A_i - \sum_{j=1}^{k} c_{i,j} A_j \right) = 0.$$

Since the A_i's are linearly independent it follows that the scalars $d_i = d_i c_{i,j} = 0$ and so the B_i's are linearly independent. Furthermore, if

$$\sum_{i=k+1}^{q} d_i t_i = 0$$

then

$$\sum_{i=k+1}^{q} d_i \left[\begin{array}{c} t_i \\ 0 \end{array} \right] = \sum_{i=k+1}^{q} d_i B_i = 0$$

and thus, the scalars $d_i = 0$, hence the set $\{t_{k+1}, \ldots, t_q\}$ is linearly independent over F.

Now consider an alternative partitioning of each of $A_1, A_2, \ldots, A_q$, $q = \lfloor n^2/4 \rfloor + 2$ into submatrices M_i', H_i' and a zero matrix of order $(n-1) \times (n-1)$, $n \times 1$ and $1 \times (n-1)$ respectively, as follows:

$$A_i = \left[\begin{array}{c|c} M_i' & \\ & H_i' \\ \hline 0 \;\; 0 \;\; \cdots \;\; 0 & \end{array}\right]$$

for all i. Again, since $A_iA_j = A_jA_i$ it follows as before that $M'_iM'_j = M'_jM'_i$, i.e., the M'_i's commute. Let W' be the vector space spanned by the set $\{M'_1, M'_2, \dots, M'_q\}$, so the elements of W' are $(n-1)\times(n-1)$ matrices. Suppose $s = \dim W'$. Then from the induction hypothesis it follows that $s \le \lfloor (n-1)^2/4 \rfloor + 1$. Assume, again without loss of generality, that W' is spanned by the linearly independent set of matrices $\{M'_1, M'_2, \dots, M'_s\}$. Then for $i > s$ each M'_i is expressible as a linear combination $M'_i = \sum_{j=1}^{s} c'_{i,j} M'_j$ with scalars $c'_{i,1}, \dots, c'_{i,s}$ in F. Now, for $i > s$, let $B'_i = A_i - \sum_{j=1}^{s} c'_{i,j} A_j$. A similar argument as before shows that

$$B'_i = \left[\begin{array}{c|c} 0 & t'_i \end{array}\right]$$

where $t'_i = H'_i - \sum_{j=1}^{s} c'_{i,j} H'_j$ is an $n \times 1$ matrix, $i = s+1, \dots, \lfloor n^2/4 \rfloor + 2$. The B_i's are linear combinations of the A_i's, so the B_i's are in V. Similarly the B'_j's are in V and hence $B_iB'_j = B'_jB_i$. From

$$B_iB'_j = \begin{bmatrix} t_i \\ 0 \end{bmatrix} \begin{bmatrix} 0 & t'_j \end{bmatrix} = \begin{bmatrix} 0 & t_it'_j \\ 0 & 0 \end{bmatrix}$$

and

$$B'_jB_i = \begin{bmatrix} 0 & t'_j \end{bmatrix} \begin{bmatrix} t_i \\ 0 \end{bmatrix} = 0$$

it follows that $t_it'_j = 0$ for every $i = k+1, \dots, q$ and $j = s+1, \dots, q$, where $q = \lfloor n^2/4 \rfloor + 2$.

Let

$$A = \begin{bmatrix} t_{k+1} \\ t_{k+2} \\ \vdots \\ t_q \end{bmatrix} = \begin{bmatrix} a_{k+1,1} & a_{k+1,2} & \cdots & a_{k+1,n-1} & a_{k+1,n} \\ a_{k+2,1} & a_{k+2,2} & \cdots & a_{k+2,n-1} & a_{k+2,n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{q,1} & a_{q,2} & \cdots & a_{q,n-1} & a_{q,n} \end{bmatrix}$$

where $q = \lfloor n^2/4 \rfloor + 2$ and $t_i = (a_{i,1}, a_{i,2}, \dots, a_{i,n})$ for $i = k+1, \dots, q$. Since the t_i's are linearly independent it follows that

$$\operatorname{rank} A \ge \lfloor n^2/4 \rfloor + 2 - (k+1) + 1 = \lfloor n^2/4 \rfloor - k + 2.$$

And $t_it'_j = 0$ implies $At'_j = 0$ which implies that t'_j is in the null space of A for all $j = s+1, \dots, \lfloor n^2/4 \rfloor + 2$. Furthermore, the t'_i's are linearly independent, thus,

$$\operatorname{nullity} A \ge \lfloor n^2/4 \rfloor + 2 - (s+1) + 1 = \lfloor n^2/4 \rfloor - s + 2.$$

Now

$$\begin{aligned} n &= \operatorname{rank} A + \operatorname{nullity} A \\ &\ge \left(\left\lfloor \frac{n^2}{4} \right\rfloor - k + 2\right) + \left(\left\lfloor \frac{n^2}{4} \right\rfloor - s + 2\right) \\ &= 2\left\lfloor \frac{n^2}{4} \right\rfloor + 4 - (k+s). \end{aligned}$$

But $k+s \leq (\lfloor\frac{(n-1)^2}{4}\rfloor + 1 + \lfloor\frac{(n-1)^2}{4}\rfloor + 1)$, so

$$\begin{aligned} n &\geq 2\left\lfloor\frac{n^2}{4}\right\rfloor + 4 - \left(\left\lfloor\frac{(n-1)^2}{4}\right\rfloor + 1 + \left\lfloor\frac{(n-1)^2}{4}\right\rfloor + 1\right) \\ &= 2\left(\left\lfloor\frac{n^2}{4}\right\rfloor - \left\lfloor\frac{(n-1)^2}{4}\right\rfloor + 1\right). \end{aligned}$$

If n is even, say $n = 2m,\ m \in \mathbb{Z}^+$, then

$$\begin{aligned} n &\geq 2\left(\left\lfloor\frac{(2m)^2}{4}\right\rfloor - \left\lfloor\frac{(2m-1)^2}{4}\right\rfloor + 1\right) \\ &= 2\left(\left\lfloor m^2\right\rfloor - \left\lfloor(m^2-m)+\frac{1}{4}\right\rfloor + 1\right) \\ &= 2(m^2 - m^2 + m + 1) \\ &= 2m+2 \\ &= n+2. \end{aligned}$$

If n is odd, say $n = 2m+1,\ m \in \mathbb{Z}^+$, then

$$\begin{aligned} n &\geq 2\left(\left\lfloor\frac{(2m+1)^2}{4}\right\rfloor - \left\lfloor\frac{(2m)^2}{4}\right\rfloor + 1\right) \\ &= 2\left(\left\lfloor(m^2+m)+\frac{1}{4}\right\rfloor - \left\lfloor m^2\right\rfloor + 1\right) \\ &= 2(m^2 + m - m^2 + 1) \\ &= 2m+2 \\ &= n+1. \end{aligned}$$

We have thus arrived at a contradiction in both cases. Thus, the assumption that $\mathfrak{F}_n$ has more than $\lfloor n^2/4\rfloor + 1$ linearly independent matrices leads to a contradiction. Therefore $\mathfrak{F}_n$ has at most $\lfloor n^2/4\rfloor + 1$ linearly independent matrices. □

The following commutative subalgebra of $M_n(F)$ given in [9] is an example of a subalgebra containing $\lfloor n^2/4\rfloor + 1$ linearly independent matrices. The example is true for all positive integer n. This shows that the upper bound $\lfloor n^2/4\rfloor + 1$ cannot be lowered. We look at the case when $n = 5$ for illustrative purposes.

Example 15. The subalgebra

$$\mathfrak{F}_{\mathfrak{n}} = \{aI + a_{i,j}E_{i,j} \,|\, 1 \leq i \leq \lfloor n/2\rfloor, \lfloor n/2\rfloor + 1 \leq j \leq n,\ a, a_{i,j} \in F\}$$

of $M_n(F)$, F a field, is a commutative subalgebra which has $\lfloor n^2/4\rfloor+1$ linearly independent matrices.

We show that $\mathfrak{F}_n$ is indeed commutative. Since, by definition of $\mathfrak{F}_n$, i is never equal to j, it follows from 1.2 that $E_{i,j}E_{i,j} = 0$ for all i and j. For elements $aI + a_{i,j}E_{i,j}$ and

$bI + b_{i,j}E_{i,j}$ in $\mathfrak{F}_n$ we have

$$\begin{aligned}(aI + a_{i,j}E_{i,j})(bI + b_{i,j}E_{i,j}) &= abI + ab_{i,j}E_{i,j} + a_{i,j}bE_{i,j} + a_{i,j}b_{i,j}E_{i,j}E_{i,j} \\ &= abI + ab_{i,j}E_{i,j} + a_{i,j}bE_{i,j}\end{aligned}$$

and

$$\begin{aligned}(bI + b_{i,j}E_{i,j})(aI + a_{i,j}E_{i,j}) &= baI + ba_{i,j}E_{i,j} + b_{i,j}aE_{i,j} + b_{i,j}a_{i,j}E_{i,j}E_{i,j} \\ &= baI + ba_{i,j}E_{i,j} + b_{i,j}aE_{i,j}.\end{aligned}$$

Since F is a field, it then follows that elements of $\mathfrak{F}_n$ commute.

We notice that every element $aI + a_{i,j}E_{i,j}$ of $\mathfrak{F}_\mathfrak{n}$ can be expressed as a linear combination

$$\begin{aligned}aI + a_{i,j}E_{i,j} &= aI + a_{1,\lfloor n/2\rfloor+1}E_{1,\lfloor n/2\rfloor+1} + \ldots + a_{1,n}E_{1,n} \\ &\quad + a_{2,\lfloor n/2\rfloor+1}E_{2,\lfloor n/2\rfloor+1} + \ldots + a_{2,n}E_{2,n} \\ &\quad \vdots \\ &\quad + a_{\lfloor n/2\rfloor,\lfloor n/2\rfloor+1}E_{\lfloor n/2\rfloor,\lfloor n/2\rfloor+1} + \ldots + a_{\lfloor n/2\rfloor,n}E_{\lfloor n/2\rfloor,n}\end{aligned}$$

and also the set

$$W = \{I, E_{i,j} | \, 1 \le i \le \lfloor n/2\rfloor, \lfloor n/2\rfloor + 1 \le j \le n\}$$

is linearly independent. Thus, W is a basis for $\mathfrak{F}_n$. The number of elements in W is

$$\left\lfloor \frac{n}{2} \right\rfloor \left(n - \left\lfloor \frac{n}{2} \right\rfloor \right) + 1.$$

Now, if n is even, say $n = 2m$, $m \in \mathbb{Z}^+$, then

$$\begin{aligned}\left\lfloor \frac{n}{2} \right\rfloor \left(n - \left\lfloor \frac{n}{2} \right\rfloor \right) + 1 &= \left\lfloor \frac{2m}{2} \right\rfloor \left(2m - \left\lfloor \frac{2m}{2} \right\rfloor \right) + 1 \\ &= \lfloor m \rfloor (2m - \lfloor m \rfloor) + 1 \\ &= m(2m - m) + 1 \\ &= m^2 + 1 \\ &= \left\lfloor \frac{(2m)^2}{4} \right\rfloor + 1 \\ &= \left\lfloor \frac{n^2}{4} \right\rfloor + 1.\end{aligned}$$

If n is odd, say $n = 2m+1$, $m \in \mathbb{Z}^+$, then

$$\begin{aligned}
\left\lfloor \frac{n}{2} \right\rfloor \left(n - \left\lfloor \frac{n}{2} \right\rfloor \right) + 1 &= \left\lfloor \frac{2m+1}{2} \right\rfloor \left(2m+1 - \left\lfloor \frac{2m+1}{2} \right\rfloor \right) + 1 \\
&= \left\lfloor m + \frac{1}{2} \right\rfloor \left(2m+1 - \left\lfloor m + \frac{1}{2} \right\rfloor \right) + 1 \\
&= m(2m+1-m) + 1 \\
&= m^2 + m + 1 \\
&= \left\lfloor m^2 + m + \frac{1}{4} \right\rfloor + 1 \\
&= \left\lfloor \frac{(2m+1)^2}{4} \right\rfloor + 1 \\
&= \left\lfloor \frac{n^2}{4} \right\rfloor + 1.
\end{aligned}$$

Example 16. In particular, if $n = 5$ and F is a field, then

$$\begin{aligned}
\mathfrak{F}_5 &= \{aI + a_{i,j}E_{i,j} | \, 1 \leq i \leq \lfloor 5/2 \rfloor, \ \lfloor 5/2 \rfloor + 1 \leq j \leq 5, \ a, a_{i,j} \in F\} \\
&= \{aI + a_{i,j}E_{i,j} | \, 1 \leq i \leq 2, \ 3 \leq j \leq 5, \ a, a_{i,j} \in F\}.
\end{aligned}$$

Let $A, B \in \mathfrak{F}_5$ such that

$$A = \left[\begin{array}{cc|ccc} a & 0 & a_{1,3} & a_{1,4} & a_{1,5} \\ 0 & a & a_{2,3} & a_{2,4} & a_{2,5} \\ \hline 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & a \end{array}\right] = aI + A_{1,2}$$

and

$$B = \left[\begin{array}{cc|ccc} b & 0 & b_{1,3} & b_{1,4} & b_{1,5} \\ 0 & b & b_{2,3} & b_{2,4} & b_{2,5} \\ \hline 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & b & 0 \\ 0 & 0 & 0 & 0 & b \end{array}\right] = bI + B_{1,2}$$

where

$$A_{1,2} = \left[\begin{array}{cc|ccc} 0 & 0 & a_{1,3} & a_{1,4} & a_{1,5} \\ 0 & 0 & a_{2,3} & a_{2,4} & a_{2,5} \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \text{ and } B_{1,2} = \left[\begin{array}{cc|ccc} 0 & 0 & b_{1,3} & b_{1,4} & b_{1,5} \\ 0 & 0 & b_{2,3} & b_{2,4} & b_{2,5} \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

We see that $A_{1,2}B_{1,2} = B_{1,2}A_{1,2} = 0$ and thus,

$$\begin{aligned}
AB &= (aI + A_{1,2})(bI + B_{1,2}) \\
&= abI + aB_{1,2} + bA_{1,2}
\end{aligned}$$

and

$$\begin{aligned} BA &= (bI + B_{1,2})(aI + A_{1,2}) \\ &= baI + bA_{1,2} + aB_{1,2}. \end{aligned}$$

Therefore $AB = BA$ indicating that the elements of $\mathfrak{F}_5$ are indeed mutually commuting. It is clear that

$$\{I_5, E_{1,3}, E_{1,4}, E_{1,5}, E_{2,3}, E_{2,4}, E_{2,5}\}$$

is a linearly independent set of matrices in $\mathfrak{F}_5$. Moreover, note that this set has 7 elements and $\lfloor 5^2/4 \rfloor + 1 = 7$.

We observe that in the subalgebra $\mathfrak{F}_\mathfrak{n}$ in Example 15 the number $\lfloor n^2/4 \rfloor + 1$ is also equal to the number of elements in the basis W of $\mathfrak{F}_\mathfrak{n}$.

Chapter 3

Cayley-Hamilton Theorem

In this chapter we give a detailed proof of the Cayley-Hamilton Theorem for an arbitrary matrix in $M_n(\mathbb{R})$, where $\mathbb{R}$ denotes the field of real numbers. We further discuss in brief the trace identity for a 2×2 matrix over $\mathbb{R}$ which arises from the Cayley-Hamilton Theorem.

Consider an $n \times n$ matrix $A \in M_n(\mathbb{R})$ given by

$$A = (a_{i,j}) = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix}. \tag{3.1}$$

The following definitions are given in relation to the general $n \times n$ matrix $A \in M_n(\mathbb{R})$ in Equation 3.1.

Definition 17. (Nering E.D. [10])

The determinant of the matrix $A = (a_{i,j})$ is defined to be the scalar $\det(A) = |a_{i,j}|$ computed according to the rule

$$\det(A) = |a_{i,j}| = \sum_{\pi} (\operatorname{sgn} \pi) a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)},$$

where the sum is taken over all permutations of the elements of the set $S = \{1, \ldots, n\}$.

We note the following with regard to the above definition:
(i) "sgn π" means the "the sign of π".
(ii) A permutation π of a set S is defined to be a one to one mapping of S onto itself.
(iii) The element which the permutation π associates with i is denoted by $\pi(i)$.
(iv) $\operatorname{sgn} \pi = +1$ if π is an even permutation.
(v) $\operatorname{sgn} \pi = -1$ if π is an odd permutation.

Example 18. Consider a 3×3 matrix $B \in M_3(\mathbb{R})$ given by

$$B = \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix}. \tag{3.2}$$

The determinant of B is

$$\begin{aligned} \det(B) &= \sum_{\pi} (\operatorname{sgn} \pi) b_{1,\pi(1)} b_{2,\pi(2)} b_{3,\pi(3)} \\ &= b_{1,1}b_{2,2}b_{3,3} + b_{1,2}b_{2,3}b_{3,1} + b_{1,3}b_{2,1}b_{3,2} - b_{1,2}b_{2,1}b_{3,3} - b_{1,3}b_{2,2}b_{3,1} - b_{1,1}b_{2,3}b_{3,2} \\ &= b_{1,1}(b_{2,2}b_{3,3} - b_{2,3}b_{3,2}) + b_{1,2}(b_{2,3}b_{3,1} - b_{2,1}b_{3,3}) + b_{1,3}(b_{2,1}b_{3,2} - b_{2,2}b_{3,1}). \end{aligned}$$

Thus,

$$\det(B) = b_{1,1}B_{1,1} + b_{1,2}B_{1,2} + b_{1,3}B_{1,3} \tag{3.3}$$

where each $B_{i,j}$ is a determinant of a matrix obtained from B by deleting the ith row and jth column of B. That is,

$$B_{1,1} = \begin{vmatrix} b_{2,2} & b_{2,3} \\ b_{3,2} & b_{3,3} \end{vmatrix}, \quad B_{1,2} = -\begin{vmatrix} b_{2,1} & b_{2,3} \\ b_{3,1} & b_{3,3} \end{vmatrix}, \quad B_{1,3} = \begin{vmatrix} b_{2,1} & b_{2,2} \\ b_{3,1} & b_{3,2} \end{vmatrix}.$$

Remark 19. The determinant of the $n \times n$ matrix A in (3.1) can be written in the form

$$\det(A) = a_{i,j}A_{i,j} + (\text{terms which do not contain } a_{i,j} \text{ as a factor}).$$

Example 18 and Remark 19 lead us to the following definition, (see Jain and Gunawardena [5]).

Definition 20. The cofactor of any entry $a_{p,q}$ of an $n \times n$ matrix $A = (a_{i,j})$ is defined to be

$$A_{p,q} = (-1)^{p+q} \cdot \det(\text{the } (n-1) \times (n-1) \text{ matrix obtained by deleting the } pth \text{ row and } qth \text{ column}).$$

In Remark 19 the scalar $A_{i,j}$ is the cofactor of the entry $a_{i,j}$ and in Example 18, $B_{1,1}$, $B_{1,2}$, $B_{1,3}$ are cofactors of $b_{1,1}, b_{1,2}$, $b_{1,3}$, respectively.

Example 21. Using the matrix $B \in M_3(\mathbb{R})$ in Example 18, we have

$$\begin{aligned} & b_{2,1}B_{1,1} + b_{2,2}B_{1,2} + b_{2,3}B_{1,3} & (3.4) \\ = \; & b_{2,1}(b_{2,2}b_{3,3} - b_{2,3}b_{3,2}) + b_{2,2}(b_{2,3}b_{3,1} - b_{2,1}b_{3,3}) + b_{2,3}(b_{2,1}b_{3,2} - b_{2,2}b_{3,1}). & (3.5) \end{aligned}$$

Multiplying and grouping on the right hand side of Equation 3.5 gives

$$\begin{aligned} & b_{2,1}B_{1,1} + b_{2,2}B_{1,2} + b_{2,3}B_{1,3} \\ = \; & b_{2,1}b_{2,2}b_{3,3} - b_{2,2}b_{2,1}b_{3,3} + b_{2,3}b_{2,1}b_{3,2} - b_{2,1}b_{2,3}b_{3,2} + b_{2,2}b_{2,3}b_{3,1} - b_{2,3}b_{2,2}b_{3,1} \\ = \; & 0. \end{aligned}$$

Similarly, $b_{3,1}B_{1,1} + b_{3,2}B_{1,2} + b_{3,3}B_{1,3} = 0$.

Remark 22. In general

$$\sum_{j=1}^{n} a_{i,j}A_{k,j} = a_{i,1}A_{k,1} + a_{i,2}A_{k,2} + \ldots + a_{i,n}A_{k,n} = 0$$

whenever $i \neq k$.

Definition 23. The transpose of the matrix $A = (a_{i,j})$ is the matrix A^T whose element $a_{i,j}$ appearing in row i and column j is the element $a_{j,i}$ appearing in row j and column i of A.

Again in relation to (3.1) and Definition 20, the matrix

$$(A_{i,j}) = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ A_{n,1} & A_{n,2} & \cdots & A_{n,n} \end{bmatrix}$$

whose entries are cofactors of all entries of an $n \times n$ matrix $A = (a_{i,j})$ is called a cofactor matrix.

Definition 24. The transpose

$$(A_{j,i}) = \begin{bmatrix} A_{1,1} & A_{2,1} & \cdots & A_{n,1} \\ A_{1,2} & A_{2,2} & \cdots & A_{n,2} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ A_{1,n} & A_{2,n} & \cdots & A_{n,n} \end{bmatrix}$$

of the cofactor matrix $(A_{i,j})$ is called the adjoint of matrix A and it is denoted by $\operatorname{adj} A$.

Lemma 25. *The product of any matrix $A \in M_n(\mathbb{R})$ and its adjoint $\operatorname{adj} A$ is equal to $(\det(A))I$, i.e.,*

$$A(\operatorname{adj} A) = (\det(A))I$$

where I is the identity matrix in $M_n(\mathbb{R})$.

Lemma 26. *Let $A_0 + A_1\lambda + \ldots + A_m\lambda^m = 0$ for all $|\lambda|$ sufficiently large, where $A_i \in M_n(\mathbb{R})$ for all i and where $\mathbb{R}$ denotes the field of real numbers. Then each $A_i = 0$.*

Proof. Multiplying

$$A_0 + A_1\lambda + \ldots + A_m\lambda^m = 0 \tag{3.6}$$

by $\frac{1}{\lambda^m}$ yields

$$A_0\frac{1}{\lambda^m} + A_1\frac{1}{\lambda^{m-1}} + \ldots + A_{m-1}\frac{1}{\lambda} + A_m = 0. \tag{3.7}$$

Letting $|\lambda| \to \infty$ in (3.7) yields $A_m = 0$ and thus (3.7) becomes

$$A_0 \frac{1}{\lambda^m} + A_1 \frac{1}{\lambda^{m-1}} + \ldots + A_{m-2} \frac{1}{\lambda^2} + A_{m-1} \frac{1}{\lambda^1} = 0. \tag{3.8}$$

Multiplying (3.8) by λ yields

$$A_0 \frac{1}{\lambda^{m-1}} + A_1 \frac{1}{\lambda^{m-2}} + \ldots + A_{m-2} \frac{1}{\lambda} + A_{m-1} = 0. \tag{3.9}$$

Letting $|\lambda| \to \infty$ in (3.9) yields $A_{m-1} = 0$ and thus (3.9) becomes

$$A_0 \frac{1}{\lambda^{m-1}} + A_1 \frac{1}{\lambda^{m-2}} + \ldots + A_{m-2} \frac{1}{\lambda^1} = 0. \tag{3.10}$$

Continuing to multiply by λ and letting $\lambda \to \infty$ results eventually in

$$\frac{A_0}{\lambda} + A_1 = 0. \tag{3.11}$$

Finally, again letting $\lambda \to \infty$ gives $A_1 = 0$ and so (3.6) implies that $A_0 = 0$. Hence all the A_i's are zero. □

Corollary 27. *Let A_i and B_i be $n \times n$ matrices over $\mathbb{R}$ and suppose that*

$$A_0 + A_1\lambda + \ldots + A_m\lambda^m = B_0 + B_1\lambda + \ldots + B_m\lambda^m$$

for all $|\lambda|$ large enough. Then $A_i = B_i$ for all i.

Proof. Since

$$A_0 - B_0 + (A_1 - B_1)\lambda + \ldots + (A_m - B_m)\lambda^m = 0,$$

Lemma 26 implies that $A_i - B_i = 0$ for all i. Hence $A_i = B_i$ for all i. □

The following theorem is the well known Cayley-Hamilton Theorem and it states that any $n \times n$ matrix A is a solution of its characteristic polynomial $p(\lambda)$.

Theorem 28. *Let A be an $n \times n$ matrix over $\mathbb{R}$ and let $p(\lambda) = \det(A - \lambda I)$ be the characteristic polynomial of A. Then $p(A) = 0$.*

Proof. Consider a matrix $A \in M_n(\mathbb{R})$ such that $A = (a_{i,j})$. Then

$$A - \lambda I = \begin{bmatrix} a_{1,1} - \lambda & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - \lambda & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} - \lambda \end{bmatrix}$$

and the determinant of $A - \lambda I$ is a polynomial in λ of degree n, say,

$$\det(A - \lambda I) = k_n\lambda^n + k_{n-1}\lambda^{n-1} + \ldots + k_1\lambda + k_0, \tag{3.12}$$

for some $k_0, k_1, \ldots, k_n \in \mathbb{R}$. Now the cofactor

$$A_{1,1} = \begin{vmatrix} a_{2,2} - \lambda & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots \\ a_{n,2} & \cdots & a_{n,n} - \lambda \end{vmatrix}$$

is a polynomial in λ of degree $n-1$. In fact each cofactor $A_{i,j}$ is a polynomial in λ of at most $(n-1)$th degree. Hence the entries of the adjoint matrix $\operatorname{adj}(A-\lambda I)$ are polynomials in λ of at most $(n-1)$th degree. Thus,

$$\operatorname{adj}(A - \lambda I) = B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \ldots + B_1\lambda + B_0$$

where each B_i is an $n \times n$ matrix with entries from $\mathbb{R}$. Now

$$\begin{aligned} & (A - \lambda I) \operatorname{adj}(A - \lambda I) \\ = \; & (A - \lambda I)(B_{n-1}\lambda^{n-1} + B_{n-2}\lambda^{n-2} + \ldots + B_1\lambda + B_0) \\ = \; & AB_{n-1}\lambda^{n-1} - IB_{n-1}\lambda^{n} + AB_{n-2}\lambda^{n-2} - IB_{n-2}\lambda^{n-1} + \ldots + AB_1\lambda - IB_1\lambda^2 + AB_0 - IB_0\lambda, \end{aligned}$$

i.e.,

$$(A - \lambda I) \operatorname{adj}(A - \lambda I) = C_n\lambda^n + C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \ldots + C_1\lambda + C_0 \tag{3.13}$$

where

$$\begin{aligned} C_n &= -IB_{n-1} \\ C_{n-1} &= AB_{n-1} - IB_{n-2} \\ C_{n-2} &= AB_{n-2} - IB_{n-3} \\ &\vdots \\ C_2 &= AB_2 - IB_1 \\ C_1 &= AB_1 - IB_0 \\ C_0 &= AB_0. \end{aligned}$$

Multiplying the above equations from the left by

$$A^n, \; A^{n-1}, \; A^{n-2}, \; \ldots, \; A^2, \; A, \; I$$

respectively, yields

$$\begin{aligned} A^nC_n &= A^n(-IB_{n-1}) \\ A^{n-1}C_{n-1} &= A^{n-1}(AB_{n-1} - IB_{n-2}) \\ A^{n-2}C_{n-2} &= A^{n-2}(AB_{n-2} - IB_{n-3}) \\ &\vdots \\ A^2C_2 &= A^2(AB_2 - IB_1) \\ AC_1 &= A(AB_1 - IB_0) \\ IC_0 &= IAB_0. \end{aligned}$$

Adding the above equations yields

$$\begin{aligned} & A^nC_n + A^{n-1}C_{n-1} + A^{n-2}C_{n-2} + \ldots + A^2C_2 + AC_1 + C_0 \\ = & -A^nB_{n-1} + A^nB_{n-1} - A^{n-1}B_{n-2} + A^{n-1}B_{n-2} - A^{n-2}B_{n-3} + \ldots + A^3B_2 \\ & -A^2B_1 + A^2B_1 - AB_0 + AB_0 \\ = & \; 0. \end{aligned}$$

That is,

$$A^nC_n + A^{n-1}C_{n-1} + A^{n-2}C_{n-2} + \ldots + A^2C_2 + AC_1 + C_0 = 0. \tag{3.14}$$

Now, it follows from Lemma 25 and (3.12) that

$$(A - \lambda I)\text{adj}(A - \lambda I) = I(\det(A - \lambda I)) = I(k_n\lambda^n + k_{n-1}\lambda^{n-1} + \ldots + k_1\lambda + k_0). \tag{3.15}$$

Substituting (3.13) into (3.15) gives

$$C_n\lambda^n + C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \ldots + C_1\lambda + C_0 = I(k_n\lambda^n + k_{n-1}\lambda^{n-1} + \ldots + k_1\lambda + k_0).$$

By Corollary 27 we have

$$C_n = Ik_n,\ C_{n-1} = Ik_{n-1},\ C_{n-2} = Ik_{n-2},\ \cdots,\ C_2 = Ik_2,\ C_1 = Ik_1,\ C_0 = Ik_0.$$

Substituting the C_i by Ik_i in (3.14) we get

$$k_nA^n + k_{n-1}A^{n-1} + k_{n-2}A^{n-2} + \ldots + k_2A^2 + k_1A + k_0I \ = \ 0.$$

□

Consider a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ over $\mathbb{R}$ with characteristic polynomial

$$\begin{aligned} p(\lambda) &= \lambda^2 - (a+d)\lambda + ad - bc \\ &= \lambda^2 - \text{tr}(A)\lambda + \det(A). \end{aligned}$$

The Cayley-Hamilton Theorem asserts that

$$p(A) = A^2 - \text{tr}(A)A + I\det(A) = 0, \tag{3.16}$$

where I is the identity matrix in $M_2(\mathbb{R})$. Since $\text{tr}(I) = 2$, taking the trace of (3.16), it follows that

$$\text{tr}(A^2) - \text{tr}^2(A) + 2\det(A) = 0. \tag{3.17}$$

Thus,

$$\det(A) = \frac{1}{2}\big(\text{tr}^2(A) - \text{tr}(A^2)\big). \tag{3.18}$$

Substituting (3.18) into (3.16) gives

$$A^2 - \text{tr}(A)A + \frac{1}{2}\big(\text{tr}^2(A) - \text{tr}(A^2)\big)I = 0. \tag{3.19}$$

The identity in (3.19) is known as the Cayley-Hamilton trace identity for 2×2 matrices. We shall see in Chapter 5 that (3.19) is not necessarily zero when we deal with matrices over a noncommutative ring.

We note that, although we have dealt with the Cayley-Hamilton Theorem for square matrices over the field of real numbers, there exist several proofs of this theorem for matrices over the field $\mathbb{C}$ of complex numbers and arbitrary commutative rings. For instance, Zhang [14] provided a proof for square matrices over the field $\mathbb{C}$ of complex numbers and Straubing [12] provided a proof for a square matrix over an arbitrary commutative ring. Furthermore, to every linear transformation of a vector space of dimension n over a field $\mathbb{F}$ there corresponds an $n \times n$ matrix over $\mathbb{F}$ and conversely, to every such matrix there corresponds a linear transformation of the vector space. Accordingly, the Cayley-Hamilton Theorem hold for linear transformations and Hungerford [3] provides a proof for linear transformations.

Chapter 4

Commutativity and Lie nilpotency in the Matrix Algebra $U_n^*(R)$

In this chapter we dissect the paper [8] by Meyer, et al. We explore the ring $U_n^*(R)$, in particular the rings $U_3^*(R)$ and $U_4^*(R)$, in relation to the polynomial identities $[[x, y], z] = 0$, $[x, y][u, v] = 0$ and $[[x, y], [u, v]] = 0$. We discuss both cases when a ring R is commutative and when R is not commutative. It is shown in Theorem 44 that $U_n^*(R)$ is Lie nilpotent of index $n - 1$. A ring satisfying an identity $[[x, y], [u, v]] = 0$ is called Lie solvable of index 2. In [8], Meyer, et al, give an example of a Lie solvable ring of index 2 which we discuss in Corollary 37.

4.1 The ring $U_3^*(R)$

Definition 29. Let R be an arbitrary ring (not necessarily commutative) with unity 1, and let

$$U_n^*(R) = \left\{ \begin{bmatrix} a & a_{1,2} & \cdots & a_{1,n} \\ 0 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a \end{bmatrix} : a, a_{i,j} \in R \right\}$$

be the subring of $M_n(R)$ of upper triangular matrices with equal diagonal entries.

Example 30. Consider the elements $E_{1,2}, E_{2,3}$ in $U_3^*(R)$, i.e.,

$$E_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad E_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then $E_{1,2}E_{2,3} = E_{1,3} \neq 0 = E_{2,3}E_{1,2}$.

The preceding example shows that $U_n^*(R)$ is not commutative if $n \geq 3$. However, $U_2^*(R)$ is commutative if R is commutative because

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}\begin{bmatrix} c & d \\ 0 & c \end{bmatrix} = \begin{bmatrix} ac & ad+bc \\ 0 & ac \end{bmatrix}$$

and

$$\begin{bmatrix} c & d \\ 0 & c \end{bmatrix}\begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \begin{bmatrix} ca & cb+da \\ 0 & ca \end{bmatrix}.$$

We note that for any element

$$A = \begin{bmatrix} a & a_{1,2} & \cdots & a_{1,n} \\ 0 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a \end{bmatrix}$$

in $U_n^*(R)$, we have

$$\begin{aligned} A &= \begin{bmatrix} a & 0 & \cdots & 0 \\ 0 & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a \end{bmatrix} + \begin{bmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= a\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & 0 \end{bmatrix} \\ &= aI_n + X, \end{aligned} \tag{4.1}$$

i.e., A is the sum of a scalar matrix aI and a strictly upper triangular matrix X.

Furthermore, for any strictly upper triangular 3×3 matrices

$$X = \begin{bmatrix} 0 & x_1 & x_2 \\ 0 & 0 & x_3 \\ 0 & 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & y_1 & y_2 \\ 0 & 0 & y_3 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 0 & z_1 & z_2 \\ 0 & 0 & z_3 \\ 0 & 0 & 0 \end{bmatrix},$$

we have

$$\begin{aligned} XYZ &= \begin{bmatrix} 0 & x_1 & x_2 \\ 0 & 0 & x_3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & y_1 & y_2 \\ 0 & 0 & y_3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & z_1 & z_2 \\ 0 & 0 & z_3 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & x_1y_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & z_1 & z_2 \\ 0 & 0 & z_3 \\ 0 & 0 & 0 \end{bmatrix} \\ &= 0. \end{aligned}$$

Remark 31.

(i) More generally, the product of n strictly upper triangular $n \times n$ matrices is zero.

(ii) $E_{1,n}X = XE_{1,n} = 0$ for any strictly upper triangular $n \times n$ matrix X.

(iii) $XY = tE_{1,3}$, for some t, for any strictly upper triangular 3×3 matrices X, Y.

(iv) In general, the product $X_1X_2 \dots X_{n-1} = tE_{1,n}$, for some t, for any strictly upper triangular $n \times n$ matrices $X_1, X_2, \dots, X_{n-1}$.

If R is commutative, then $Xa = aX$ for every $X \in M_n(R)$ and every $a \in R$. Hence

$$X(aI) = (Xa)I = a(XI) = aX$$

and

$$(aI)X = a(IX) = aX.$$

Thus aI commutes with every $X \in M_n(R)$ (for a commutative ring R).

Proposition 32. *If R is commutative, then $[aI + X, bI + Y] = [X, Y]$ for all $a, b \in R$ and all $X, Y \in M_n(R)$.*

Proof.

$$\begin{aligned} & [aI + X, bI + Y] \\ = \; & (aI + X)(bI + Y) - (bI + Y)(aI + X) \\ = \; & aI(bI + Y) + X(bI + Y) - [bI(aI + X) + Y(aI + X)] \\ = \; & (aI)(bI) + (aI)Y + X(bI) + XY - (bI)(aI) - (bI)X - Y(aI) - YX \\ = \; & (aI)(bI) - (bI)(aI) + (aI)Y - Y(aI) + X(bI) - (bI)X + XY - YX, \end{aligned}$$

and so the observation following Remark 31 shows that

$$[aI + X, bI + Y] = XY - YX = [X, Y].$$

□

We note that a ring R is commutative if and only if $[x_1, x_2] = 0$ for all $x_1, x_2 \in R$. Since $[E_{1,2}, E_{2,3}] = E_{1,3} \neq 0$, we conclude that $U_3^*(R)$ does not satisfy the identity $[x_1, x_2] = 0$, i.e., $U_3^*(R)$ is not commutative. However, $U_3^*(R)$ does satisfy the following (albeit weaker) condition if R is commutative.

Proposition 33. *If R is a commutative ring, then $U_3^*(R)$ satisfies the polynomial identity*

$$[[x_1, x_2], x_3] = 0,$$

i.e., $[[A, B], C] = 0$ for all matrices A, B, C in $U_3^*(R)$.

Proof. Consider any matrices A, B, C in $U_3^*(R)$ with

$$A = \begin{bmatrix} a & x_1 & x_2 \\ 0 & a & x_3 \\ 0 & 0 & a \end{bmatrix}, B = \begin{bmatrix} b & y_1 & y_2 \\ 0 & b & y_3 \\ 0 & 0 & b \end{bmatrix}, C = \begin{bmatrix} c & u_1 & u_2 \\ 0 & c & u_3 \\ 0 & 0 & c \end{bmatrix}.$$

By (4.1), $A = aI + X,\ B = bI + Y,\ C = cI + U$ for suitable $a, b, c \in R$ and strictly upper triangular matrices X, Y and U. Thus, by Remark 31 (ii) and (iii),

$$\begin{aligned} [[A, B], C] &= [[X, Y], U] \\ &= [tE_{1,3}, U] \\ &= (tE_{1,3})U - U(tE_{1,3}) \\ &= 0. \end{aligned}$$

□

Proposition 34. *If R is a commutative ring, then $U_3^*(R)$ satisfies the polynomial identity*

$$[x, y][z, w] = 0,$$

i.e., $[A, B][C, D] = 0$ for all matrices A, B, C, D in $U_3^*(R)$.

Proof. Consider any matrices A, B, C, D in $U_3^*(R)$ with

$$A = \begin{bmatrix} a & x_1 & x_2 \\ 0 & a & x_3 \\ 0 & 0 & a \end{bmatrix}, B = \begin{bmatrix} b & y_1 & y_2 \\ 0 & b & y_3 \\ 0 & 0 & b \end{bmatrix}, C = \begin{bmatrix} c & u_1 & u_2 \\ 0 & c & u_3 \\ 0 & 0 & c \end{bmatrix}, D = \begin{bmatrix} d & v_1 & v_2 \\ 0 & d & v_3 \\ 0 & 0 & d \end{bmatrix}.$$

By (4.1), $A = aI + X,\ B = bI + Y,\ C = cI + U,\ D = dI + V$ for suitable $a, b, c, d \in R$ and strictly upper triangular matrices X, Y, U and V. Thus, by Proposition 32 and Remark 31 (iii),

$$\begin{aligned} [A, B] &= [aI + X, bI + Y] \\ &= [X, Y] \\ &= XY - YX \\ &= sE_{1,3} - tE_{1,3} \\ &= (s - t)E_{1,3} \ \text{ for some } s, t \in R. \end{aligned}$$

Similar calculations yield

$$[C, D] = uE_{1,3} \ \text{ for some } u \in R.$$

Now, by (1.2),

$$\begin{aligned} [A, B][C, D] &= ((s - t)E_{1,3})(uE_{1,3}) \\ &= 0. \end{aligned}$$

□

We conclude from Proposition 33 that if R is commutative, then $U_3^*(R)$ is Lie nilpotent of index 2, i.e., $U_3^*(R)$ satisfies the polynomial identity $[[x_1, x_2], x_3] = 0$ but not the identity $[x_1, x_2] = 0$. In Theorem 44 we will prove that if R is commutative, then $U_n^*(R)$ is Lie nilpotent of index $n - 1$.

The example below shows that if the ring R is not commutative, then Proposition 33 and Proposition 34 do not hold.

Example 35. Let R be a noncommutative ring. Then $[r, s] \neq 0$ for some $r, s \in R$, and so for $X = rI$, $Y = sI$, $W = sE_{1,2}$, $U = E_{1,2}$, $V = Z = E_{2,3}$ in $U_3^*(R)$, we have

$$\begin{aligned}
[X, Y][U, V] &= [rI, sI][E_{1,2}, E_{2,3}] \\
&= ((rI)(sI) - (sI)(rI))(E_{1,2}E_{2,3} - E_{2,3}E_{1,2}) \\
&= (rsI - srI)(E_{1,3} - 0) \\
&= ((rs - sr)I)E_{1,3} \\
&= [r, s]E_{1,3} \\
&\neq 0
\end{aligned}$$

and

$$\begin{aligned}
[[X, W], Z] &= [[rI, sE_{1,2}], E_{2,3}] \\
&= [(rs - sr)E_{1,2}, E_{2,3}] \\
&= (rs - sr)E_{1,2}E_{2,3} - E_{2,3}(rs - sr)E_{1,2} \\
&= (rs - sr)E_{1,3} - (rs - sr)E_{2,3}E_{1,2} \\
&= (rs - sr)E_{1,3} - (rs - sr)0 \\
&= [r, s]E_{1,3} \\
&\neq 0.
\end{aligned}$$

The following theorem, found in [8], shows that if the underlying ring S satisfies the identities $[x, y][u, v] = 0$ and $[[x, y], z] = 0$ then the matrix ring $U_3^*(S)$ is Lie solvable of index 2.

Theorem 36. *If S satisfies $[x, y][u, v] = 0$ and $[[x, y], z] = 0$ then $U_3^*(S)$ satisfies*

$$[[X, Y], [U, V]] = 0.$$

Proof. For

$$X = \begin{bmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} e & f & g \\ 0 & e & h \\ 0 & 0 & e \end{bmatrix}$$

in $U_3^*(S)$ we have

$$\begin{aligned}
[X,Y] &= XY - YX \\
&= \begin{bmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{bmatrix}\begin{bmatrix} e & f & g \\ 0 & e & h \\ 0 & 0 & e \end{bmatrix} - \begin{bmatrix} e & f & g \\ 0 & e & h \\ 0 & 0 & e \end{bmatrix}\begin{bmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{bmatrix} \\
&= \begin{bmatrix} ae & af+be & ag+bh+ce \\ 0 & ae & ah+de \\ 0 & 0 & ae \end{bmatrix} - \begin{bmatrix} ea & eb+fa & ec+fd+ga \\ 0 & ea & ed+ha \\ 0 & 0 & ea \end{bmatrix} \\
&= \begin{bmatrix} ae-ea & af+be-(eb+fa) & ag+bh+ce-(ec+fd+ga) \\ 0 & ae-ea & ah+de-(ed+ha) \\ 0 & 0 & ae-ea \end{bmatrix} \\
&= \begin{bmatrix} ae-ea & af-fa+be-eb & ag-ga+bh-fd+ce-ec \\ 0 & ae-ea & ah-ha+de-ed \\ 0 & 0 & ae-ea \end{bmatrix} \\
&= \begin{bmatrix} [a,e] & [a,f]+[b,e] & [a,g]+[c,e]+(bh-fd) \\ 0 & [a,e] & [a,h]+[d,e] \\ 0 & 0 & [a,e] \end{bmatrix} \\
&= \begin{bmatrix} [a,e] & 0 & 0 \\ 0 & [a,e] & 0 \\ 0 & 0 & [a,e] \end{bmatrix} + \begin{bmatrix} 0 & [a,f]+[b,e] & [a,g]+[c,e] \\ 0 & 0 & [a,h]+[d,e] \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & (bh-fd) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&= [a,e]\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & [a,f]+[b,e] & [a,g]+[c,e] \\ 0 & 0 & [a,h]+[d,e] \\ 0 & 0 & 0 \end{bmatrix} + (bh-fd)\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Therefore

$$[X,Y] = [a,e]I + C + \alpha E_{1,3} \tag{4.2}$$

where $\alpha = bh - fd$ and $C = \begin{bmatrix} 0 & [a,f]+[b,e] & [a,g]+[c,e] \\ 0 & 0 & [a,h]+[d,e] \\ 0 & 0 & 0 \end{bmatrix}$.

The entries of the strictly upper triangular matrix C are in the additive subgroup $[S,S]$ of S generated by all commutators. If

$$U = \begin{bmatrix} a' & b' & c' \\ 0 & a' & d' \\ 0 & 0 & a' \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} e' & f' & g' \\ 0 & e' & h' \\ 0 & 0 & e' \end{bmatrix},$$

it follows from (4.2) that

$$[U,V] = [a',e']I + C' + \alpha' E_{1,3}$$

where $\alpha' = b'h' - f'd'$, $C' = \begin{bmatrix} 0 & [a',f']+[b',e'] & [a',g']+[c',e'] \\ 0 & 0 & [a',h']+[d',e'] \\ 0 & 0 & 0 \end{bmatrix}$
and the entries of the strictly upper triangular matrix C' are in the additive subgroup $[S,S]$ of S generated by all commutators.

By hypothesis the ring S satisfies $[[a,e],s]=0$ for every $s \in S$. That is, $[a,e]s = s[a,e]$ for every $s \in S$ and hence $[a,e]I$ is central in $U_3^*(S)$. Moreover, it is central also in $M_3(S)$. Thus

$$\begin{aligned}
& [[X,Y],[U,V]] \\
= \; & [[a,e]I + C + \alpha E_{1,3}, [a',e']I + C' + \alpha' E_{1,3}] \\
= \; & [a,e]I([a',e']I + C' + \alpha' E_{1,3}) - ([a',e']I + C' + \alpha' E_{1,3})[a,e]I \\
& +(C + \alpha E_{1,3})([a',e']I + C' + \alpha' E_{1,3}) - ([a',e']I + C' + \alpha' E_{1,3})(C + \alpha E_{1,3}) \\
= \; & (C + \alpha E_{1,3})[a',e']I - [a',e']I(C + \alpha E_{1,3}) + (C + \alpha E_{1,3})(C' + \alpha' E_{1,3}) \\
& -(C' + \alpha' E_{1,3})(C + \alpha E_{1,3}) \\
= \; & (C + \alpha E_{1,3})(C' + \alpha' E_{1,3}) - (C' + \alpha' E_{1,3})(C + \alpha E_{1,3}) \\
= \; & CC' + CE_{1,3}\alpha' + \alpha E_{1,3}C' + (\alpha\alpha')E_{1,3}E_{1,3} - C'C - C'E_{1,3}\alpha - \alpha' E_{1,3}C - (\alpha'\alpha)E_{1,3}E_{1,3} \\
= \; & CC' + CE_{1,3}\alpha' + \alpha E_{1,3}C' - C'C - C'E_{1,3}\alpha - \alpha' E_{1,3}C.
\end{aligned}$$

We have $C, C' \in M_3([S,S])$ and by hypothesis $[x,y][u,v] = 0$, thus $CC' = C'C = 0$. Furthermore, C and C' are strictly upper triangular matrices, hence by Remark 31 (ii) we have $CE_{1,3} = E_{1,3}C' = C'E_{1,3} = E_{1,3}C = 0$. Therefore $[[X,Y],[U,V]] = 0$. □

Let R be a ring with unity 1. By Definition 29,

$$U_3^*(U_3^*(R)) = \left\{ \begin{bmatrix} W & X & Y \\ 0 & W & Z \\ 0 & 0 & W \end{bmatrix} \mid W, X, Y, Z \in U_3^*(R) \right\}.$$

Corollary 37. *If R is commutative, then the algebra $U_3^*(U_3^*(R))$ satisfies the polynomial identity $[[x,y],[u,v]] = 0$ but neither $[x,y][u,v] = 0$ nor $[[x,y],z] = 0$, i.e.,*
(i) $[[W,X],[Y,Z]] = 0$ for all matrices W, X, Y and Z in $U_3^*(U_3^*(R))$ whereas,
(ii) $[W,X][Y,Z] = 0$ and $[[W,X],Y] = 0$ do not necessarily hold.

Proof. Let R be a commutative ring. Then, by Proposition 33 and Proposition 34, we have $[[A,B],C] = 0$ and $[A,B][C,D] = 0$ for every A, B, C, D in $U_3^*(R)$. Thus, by Theorem 36, $U_3^*(U_3^*(R))$ satisfies $[[W,X],[Y,Z]] = 0$. Example 35 shows that if R is not commutative, then $U_3^*(R)$ does not satisfy the polynomial identities $[A,B][C,D] = 0$ and $[[A,B],C] = 0$. Since Example 30 shows $U_3^*(R)$ is not commutative, it follows that $U_3^*(U_3^*(R))$ does not satisfy $[W,X][Y,Z] = 0$ and $[[W,X],Y] = 0$. □

It follows from Corollary 37 (i) that the algebra $U_3^*(U_3^*(R))$ is Lie solvable of index 2. If R is commutative, then the ring $U_2^*(R)$ is also commutative and hence $U_3^*(U_2^*(R))$ is an example of an algebra which satisfies the polynomial identities $[X,Y][U,V] = 0$ (and hence $[[X,Y],[U,V]] = 0$) and $[[X,Y],Z] = 0$.

Below we give another example of a ring which satisfies the conclusion of Theorem 36. In this example, unlike in Corollary 37, R is a noncommutative ring. We use a ring $V_n^{**}(R)$

analogous to the ring $\mathfrak{F}_n$ given in Chapter 2. Elements of $V_n^{**}(R)$ are matrices over a noncommutative ring R whereas in $\mathfrak{F}_n$ a field was considered.

Let R be a noncommutative ring with unity 1 and let

$$V_n^{**}(R) = \{aI + a_{i,j}E_{i,j} | 1 \leq i \leq \lfloor n/2 \rfloor, \ \lfloor n/2 \rfloor + 1 \leq j \leq n, \ a, a_{i,j} \in R\}.$$

By definition of $V_n^{**}(R)$, i is never equal to j, so from (1.2) we have $E_{i,j}E_{i,j} = 0$ for all i and j. For elements $X = aI + a_{i,j}E_{i,j}$ and $Y = b + b_{i,j}E_{i,j}$ in $V_n^{**}(R)$ we have

$$XY = (aI + a_{i,j}E_{i,j})(bI + b_{i,j}E_{i,j}) = abI + ab_{i,j}E_{i,j} + a_{i,j}bE_{i,j} \tag{4.3}$$

and

$$YX = (bI + b_{i,j}E_{i,j})(aI + a_{i,j}E_{i,j}) = baI + ba_{i,j}E_{i,j} + b_{i,j}aE_{i,j}. \tag{4.4}$$

Since R is not commutative, we see that in general $XY \neq YX$. Thus, for all n, $V_n^{**}(R)$ is also not commutative. Furthermore,

$$\begin{aligned} [X, Y] &= [aI + a_{i,j}E_{i,j}, bI + b_{i,j}E_{i,j}] \\ &= abI + ab_{i,j}E_{i,j} + a_{i,j}bE_{i,j} - baI - ba_{i,j}E_{i,j} - b_{i,j}aE_{i,j} \\ &= [a, b]I + [a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j}. \end{aligned}$$

Similar calculations show that

$$[W, Z] = [c, d]I + [c, d_{i,j}]E_{i,j} + [c_{i,j}, d]E_{i,j},$$

for any $W, Z \in V_n^{**}(R)$.
Now, if R satisfies the identities $[[x, y], z] = 0$ and $[x, y][w, z] = 0$, then

$$\begin{aligned} [X, Y][W, Z] &= ([a, b]I + [a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j})([c, d]I + [c, d_{i,j}]E_{i,j} + [c_{i,j}, d]E_{i,j}) \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} & [[X, Y], Z] \\ = \ & ([a, b]I + [a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j})(dI + d_{i,j}E_{i,j}) \\ & -(dI + d_{i,j}E_{i,j})([a, b]I + [a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j}) \\ = \ & [a, b]I(dI + d_{i,j}E_{i,j}) - (dI + d_{i,j}E_{i,j})[a, b]I + ([a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j})(dI + d_{i,j}E_{i,j}) \\ & -(dI + d_{i,j}E_{i,j})([a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j}) \\ = \ & ([a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j})(dI + d_{i,j}E_{i,j}) - (dI + d_{i,j}E_{i,j})([a, b_{i,j}]E_{i,j} + [a_{i,j}, b]E_{i,j}) \\ = \ & [[a, b_{i,j}], d]E_{i,j} + [[a_{i,j}, b], d]E_{i,j} \\ = \ & 0. \end{aligned}$$

We conclude from Theorem 36 and Example 35 that if R satisfies the identities $[[x, y], z] = 0$ and $[x, y][w, z] = 0$, then $U_3^*(V_n^{**}(R))$ satisfies the polynomial identity $[[x, y], [w, z]] = 0$ but neither $[[x, y], z] = 0$ nor $[x, y][w, z] = 0$.

The following theorem and Remark 39 will be used in the proof of Proposition 40, which will in turn be used in the proof of Proposition 41.

Theorem 38. *The product of two upper triangular matrices over any ring R is also an upper triangular matrix.*

Proof. Consider any two upper triangular matrices $X = (x_{i,j})$ and $Y = (y_{j,k})$ each of order $n \times n$ with entries from R. We have that $x_{i,j} = 0$ when $i > j$ and $y_{j,k} = 0$ when $j > k$. We shall show that the matrix $XY = (\alpha_{i,k})$ of order $n \times n$ with $\alpha_{i,k} = \sum_{j=1}^{n} x_{i,j}y_{j,k}$ is also an upper triangular matrix. To this end, let $i > k$. Then

$$\begin{aligned} \alpha_{i,k} &= \sum_{j=1}^{n} x_{i,j}y_{j,k} \\ &= \sum_{j=1}^{i-1} x_{i,j}y_{j,k} + \sum_{j=i}^{n} x_{i,j}y_{j,k} \\ &= \sum_{j=1}^{i-1} 0 \cdot y_{j,k} + \sum_{j=i}^{n} x_{i,j} \cdot 0 \\ &= 0, \end{aligned}$$

since $x_{i,j} = 0$ when $i > j$ and $y_{j,k} = 0$ when $j \geq i > k$. This shows that XY is an upper triangular matrix. □

Remark 39. The difference of two upper triangular matrices with equal corresponding principal diagonal entries is a strictly upper triangular matrix.

Proposition 40. *Let R be a commutative ring and X_1, X_2 be matrices in $U_n^*(R)$. Then the principal diagonal and the first diagonal above the principal diagonal of $[X_1, X_2]$ are zero.*

Proof. Suppose R is a commutative ring and let $X_1 = (a_{i,j})$ and $X_2 = (b_{j,k})$ be matrices in $U_n^*(R)$. Then $a_{i,j} = 0$ whenever $i > j$ and $b_{j,k} = 0$ whenever $j > k$. Also $a_{p,p} = a_{q,q}$ and $b_{p,p} = b_{q,q}$ for all p and q. By Theorem 38 above,

$$X_1X_2 = (\alpha_{i,k}) = \left(\sum_{j=1}^{n} a_{i,j}b_{j,k}\right) \quad \text{with} \quad \alpha_{i,k} = 0 \quad \text{whenever} \quad i > k$$

and

$$X_2X_1 = (\beta_{i,k}) = \left(\sum_{j=1}^{n} b_{i,j}a_{j,k}\right) \quad \text{with} \quad \beta_{i,k} = 0 \quad \text{whenever} \quad i > k.$$

When $i = k$, then $\alpha_{i,i} = \sum_{j=1}^{n} a_{i,j}b_{j,i} = a_{i,i}b_{i,i}$ and $\beta_{i,i} = \sum_{j=1}^{n} b_{i,j}a_{j,i} = b_{i,i}a_{i,i}$ since $a_{i,j} = 0$ whenever $i > j$ and $b_{j,i} = 0$ whenever $j > i$. Since R is commutative, it follows that $\alpha_{i,i} = a_{i,i}b_{i,i} = b_{i,i}a_{i,i} = \beta_{i,i}$. By Remark 39, $X_1X_2 - X_2X_1$ is a strictly upper triangular matrix. Furthermore, we note that for the entries

$$\alpha_{i,i+1} = \sum_{j=1}^{n} a_{i,j}b_{j,i+1} \text{ and } \beta_{i,i+1} = \sum_{j=1}^{n} b_{i,j}a_{j,i+1}$$

we have

$$\begin{aligned}
&\alpha_{i,i+1} - \beta_{i,i+1} \\
= &\sum_{j=1}^{n} a_{i,j}b_{j,i+1} - \sum_{j=1}^{n} b_{i,j}a_{j,i+1} \\
= &\sum_{j=1}^{n}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) \\
= &\sum_{j=1}^{i-1}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) + \sum_{j=i}^{i+1}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) + \sum_{j=i+2}^{n}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) \\
= &\sum_{j=1}^{i-1}(0 \cdot b_{j,i+1} - 0 \cdot a_{j,i+1}) + \sum_{j=i}^{i+1}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) + \sum_{j=i+2}^{n}(a_{i,j} \cdot 0 - b_{i,j} \cdot 0) \\
= &\sum_{j=i}^{i+1}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1})
\end{aligned}$$

since $a_{i,j} = b_{i,j} = 0$ when $i > j$. But

$$\begin{aligned}
&\sum_{j=i}^{i+1}(a_{i,j}b_{j,i+1} - b_{i,j}a_{j,i+1}) \\
= &\; a_{i,i}b_{i,i+1} - b_{i,i}a_{i,i+1} + a_{i,i+1}b_{i+1,i+1} - b_{i,i+1}a_{i+1,i+1} \\
= &\; (a_{i,i}b_{i,i+1} - b_{i,i+1}a_{i+1,i+1}) + (a_{i,i+1}b_{i+1,i+1} - b_{i,i}a_{i,i+1}) \\
= &\; 0,
\end{aligned}$$

since R is commutative and $a_{i,i} = a_{i+1,i+1}$ and $b_{i,i} = b_{i+1,i+1}$ for all i. Thus $\alpha_{i,i+1} - \beta_{i,i+1} = 0$ for all $i \in \mathbb{Z}$ such that $1 \leq i \leq n-1$. We also know that $[X_1, X_2]$ is a strictly upper triangular matrix. Thus the principal diagonal and the first diagonal above the principal diagonal of $[X_1, X_2]$ are zero. □

Proposition 41. *If R is a commutative ring, then the ring $U_n^*(R)$ satisfies the identity*

$$[[[...[[X_1, X_2], X_3], ...], X_{n-1}], X_n] = 0.$$

Proof. We shall show that

$$[[[...[[X_1, X_2], X_3], ...], X_{n-1}], X_n] = 0$$

for all $X_1, X_2, ..., X_n \in U_n^*(R)$, with R a commutative ring. By Proposition 40, the principal diagonal and the first diagonal above the principal diagonal of $[X_1, X_2]$ are zero.

Assume that for some fixed k such that $2 \leq k \leq n-1$, the principal diagonal and the jth diagonal above the principal diagonal of

$$P = [[...[[X_1, X_2], X_3], ...], X_k] = (r_{i,j})$$

are zero for $j = 1, 2, ..., k-1$. Thus, $r_{p,q} = 0$ when $p > q$ or $|p-q| < k$. Now, for $X_{k+1} = (s_{i,j}) \in U_n^*(R)$, we have

$$X_{k+1}P = (\eta_{i,k}) = \left(\sum_{p=1}^{n} s_{i,p}r_{p,k}\right)$$

such that the entries of the jth diagonal above the principal diagonal are

$$\begin{aligned}
\eta_{i,i+j} &= \sum_{p=1}^{n} s_{i,p}r_{p,i+j} \\
&= \sum_{p=1}^{i-1} s_{i,p}r_{p,i+j} + \sum_{p=i}^{n} s_{i,p}r_{p,i+j} \\
&= \sum_{p=1}^{i-1} 0 \cdot r_{p,i+j} + \sum_{p=i}^{n} s_{i,p} \cdot 0 \\
&= 0.
\end{aligned}$$

Also, the entries of the jth diagonal above the principal diagonal of

$$PX_{k+1} = (\zeta_{i,k}) = \left(\sum_{p=1}^{n} r_{i,p}s_{p,k}\right)$$

are

$$\begin{aligned}
\zeta_{i,i+j} &= \sum_{p=1}^{n} r_{i,p}s_{p,i+j} \\
&= \sum_{p=1}^{i+j} r_{i,p}s_{p,i+j} + \sum_{p=i+j+1}^{n} r_{i,p}s_{p,i+j} \\
&= \sum_{p=1}^{i+j} 0 \cdot s_{p,i+j} + \sum_{p=i+j+1}^{n} r_{i,p} \cdot 0 \\
&= 0.
\end{aligned}$$

Thus, the principal diagonal and the jth diagonal above the principal diagonals of $X_{k+1}P$ and PX_{k+1} are zero for $j = 1, 2, ..., k-1$ and hence the same is true for $[P, X_{k+1}]$. Since $|i-(i+j)| < k$ we have $r_{i,i+j} = 0$ for $j = 1, 2, ..., k-1$ and $i = 1, 2, ..., n-1$, hence it follows that the entries of the kth diagonal above the principal diagonal of PX_{k+1} are

$$\begin{aligned}\zeta_{i,i+k} &= \sum_{p=1}^{n} r_{i,p}s_{p,i+k}\\ &= \sum_{p=1}^{i+(k-1)} r_{i,p}s_{p,i+k} + r_{i,i+k}s_{i+k,i+k} + \sum_{p=i+(k+1)}^{n} r_{i,p}s_{p,i+k}\\ &= \sum_{p=1}^{i+(k-1)} 0\cdot s_{p,i+k} + r_{i,i+k}s_{i+k,i+k} + \sum_{p=i+(k+1)}^{n} r_{i,p}\cdot 0\\ &= r_{i,i+k}s_{i+k,i+k}.\end{aligned}$$

Also, $|p-(i+k)| = |(i+t)-(i+k)| = |t-k| < k$ for $t = 1, 2, \ldots, n-i$ and so $r_{i+t,i+k} = 0$. Hence, it follows that the entries of the kth diagonal above the principal diagonal of $X_{k+1}P$ are

$$\begin{aligned}\eta_{i,i+k} &= \sum_{p=1}^{n} s_{i,p}r_{p,i+k}\\ &= \sum_{p=1}^{i-1} s_{i,p}r_{p,i+k} + s_{i,i}r_{i,i+k} + \sum_{p=i+1}^{n} s_{i,p}r_{p,i+k}\\ &= \sum_{p=1}^{i-1} 0\cdot r_{p,i+k} + s_{i,i}r_{i,i+k} + \sum_{p=i+1}^{n} s_{i,p}\cdot 0\\ &= s_{i,i}r_{i,i+k}.\end{aligned}$$

By commutativity of R we have

$$\zeta_{i,i+k} - \eta_{i,i+k} = r_{i,i+k}s_{i+k,i+k} - s_{i,i}r_{i,i+k} = 0$$

since $s_{i,i} = s_{i+k,i+k}$.

In general we have $\zeta_{i,i+k} - \eta_{i,i+k} = 0$ for every i in $\mathbb{Z}$ such that $1 \leq i \leq n-k$. That is, the kth diagonal above the principal diagonal of $[P, X_{k+1}]$ is zero. Therefore the principal diagonal and the jth diagonal above the main diagonal of

$$[P, X_{k+1}] = [[[\ldots[[X_1, X_2], X_3], \ldots], X_k], X_{k+1}]$$

are zero for $j = 1, 2, \ldots, k$. It follows from induction that the principal diagonal and the jth diagonal above the principal diagonal of

$$[[[\ldots[[X_1, X_2], X_3], \ldots], X_{n-1}], X_n]$$

are zero for $j = 1, 2, \ldots, n-1$ and hence

$$[[[\ldots[[X_1, X_2], X_3], \ldots], X_{n-1}], X_n] = 0.$$

□

Proposition 42. *Let R be a commutative ring. Then the ring $U_n^*(R)$ does not satisfy the identity*

$$[[\ldots[[X_1, X_2], X_3], \ldots], X_{n-1}] = 0.$$

Proof. Let R be a commutative ring and Y_i be the matrix $E_{i,i+1}$ in $U_n^*(R)$, $i = 1, \ldots, n-1$. Then $[Y_1, Y_2] = E_{1,2}E_{2,3} - 0 = E_{1,3}$ and $[[Y_1, Y_2], Y_3] = E_{1,3}E_{3,4} - 0 = E_{1,4}$. Assume that for some $k \leq n-2$ we have

$$[[...[[Y_1, Y_2], Y_3], ...], Y_k] = E_{1,k+1}.$$

Then

$$[[[...[[Y_1, Y_2], Y_3], ...], Y_k], Y_{k+1}] = E_{1,k+1}E_{k+1,k+2} - 0 = E_{1,k+2}.$$

It follows from induction that

$$[[...[[Y_1, Y_2], Y_3], ...], Y_{n-1}] = E_{1,n} \neq 0.$$

□

Remark 43. It follows from Propositions 41 and 42 that if R is commutative then n is the smallest k such that $U_n^*(R)$ satisfies the identity

$$[[\ldots[[X_1, X_2], X_3], \ldots], X_k] = 0,$$

and so we arrive at the following theorem.

Theorem 44. *If R is a commutative ring, then $U_n^*(R)$ is Lie nilpotent of index $n-1$.*

Proof. Let R be a commutative ring. Then by Proposition 42 the ring $U_n^*(R)$ does not satisfy the identity

$$[[...[[X_1, X_2], X_3], ...], X_{n-1}] = 0.$$

Thus, it follows from Proposition 41 that $U_n^*(R)$ is Lie nilpotent of index $n-1$. □

In the following example we show that if, in Theorem 44, the hypothesis changes to a ring R being noncommutative, then the conclusion does not necessarily hold.

Example 45. Consider the elements $E_{1,2}$ and $E_{2,2}$ in the noncommutative ring $M_2(R)$, and let I denote the identity in $U_n^*(M_2(R))$. Then $[E_{1,2}I, E_{2,2}I] = E_{1,2}I$, and so

$$[[[\ldots[[E_{1,2}I, E_{2,2}I], E_{2,2}I], \ldots], E_{2,2}I], E_{2,2}I] \quad = \quad E_{1,2}I \neq 0.$$

Therefore, $U_n^*(M_2(R))$ is not Lie nilpotent. Thus, if R is not commutative, then the ring $U_n^*(R)$ is not necessarily Lie nilpotent.

4.2 The ring $U_4^*(R)$

In this section we consider the ring $U_4^*(R)$ and look at results similar to Proposition 34 and Theorem 36. We also look at Lie solvability for $U_4^*(R)$. We note that, by Theorem 44, the ring $U_4^*(R)$ with R commutative is Lie nilpotent of index 3 and hence does not satisfy $[[x, y], z] = 0$.

Consider elements X and Y in $U_4^*(R)$ such that

$$X = \begin{bmatrix} 0 & x_1 & x_2 & x_3 \\ 0 & 0 & x_4 & x_5 \\ 0 & 0 & 0 & x_6 \\ 0 & 0 & 0 & 0 \end{bmatrix}, Y = \begin{bmatrix} 0 & y_1 & y_2 & y_3 \\ 0 & 0 & y_4 & y_5 \\ 0 & 0 & 0 & y_6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then

$$XY = \begin{bmatrix} 0 & x_1 & x_2 & x_3 \\ 0 & 0 & x_4 & x_5 \\ 0 & 0 & 0 & x_6 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & y_1 & y_2 & y_3 \\ 0 & 0 & y_4 & y_5 \\ 0 & 0 & 0 & y_6 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & x_1y_4 & x_1y_5 + x_2y_6 \\ 0 & 0 & 0 & x_4y_6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Proposition 46. *If R is a commutative ring, then $U_4^*(R)$ satisfies the polynomial identity*

$$[x, y][z, w] = 0,$$

i.e., $[A, B][C, D] = 0$ for all matrices A, B, C and D in $U_4^*(R)$.

Proof. Consider any matrices A, B, C, D in $U_4^*(R)$ with

$$A = \begin{bmatrix} a & x_1 & x_2 & x_3 \\ 0 & a & x_4 & x_5 \\ 0 & 0 & a & x_6 \\ 0 & 0 & 0 & a \end{bmatrix}, B = \begin{bmatrix} b & y_1 & y_2 & y_3 \\ 0 & b & y_4 & y_5 \\ 0 & 0 & b & y_6 \\ 0 & 0 & 0 & b \end{bmatrix},$$

$$C = \begin{bmatrix} c & u_1 & u_2 & u_3 \\ 0 & c & u_4 & u_5 \\ 0 & 0 & c & u_6 \\ 0 & 0 & 0 & c \end{bmatrix}, D = \begin{bmatrix} d & v_1 & v_2 & v_3 \\ 0 & d & v_4 & v_5 \\ 0 & 0 & d & v_6 \\ 0 & 0 & 0 & d \end{bmatrix}.$$

By (4.1), $A = aI+X$, $B = bI+Y$, $C = cI+U$, $D = dI+V$ with strictly upper triangular matrices X, Y, U and V. Thus,

$$\begin{aligned} [A, B] &= [aI + X, bI + Y] \\ &= [X, Y] \\ &= XY - YX \\ &= t_1E_{1,3} + t_2E_{1,4} + t_3E_{2,4}, \end{aligned}$$

for suitable $t_1, t_2, t_3 \in R$. Similar calculations yield

$$[C, D] = q_1 E_{1,3} + q_2 E_{1,4} + q_3 E_{2,4},$$

for suitable $q_1, q_2, q_3 \in R$.

Now, by (1.2),

$$\begin{aligned} [A, B][C, D] &= (t_1 E_{1,3} + t_2 E_{1,4} + t_3 E_{2,4})(q_1 E_{1,3} + q_2 E_{1,4} + q_3 E_{2,4}) \\ &= 0. \end{aligned}$$

□

We now gather some information about the number of zero diagonals above the principal diagonal in a product of strictly upper triangular matrices. This information is helpful in proving Theorem 52.

Lemma 47. *Let R be an arbitrary ring with unity 1 and $A = (a_{i,p})$, $B = (b_{p,q}) \in U_n^*(R)$. Suppose the principal diagonal and the jth diagonal above the principal diagonal of A and B are zero for $j = 1, ..., k-1$. Then the principal diagonal and the first $2k-1$ diagonals above the principal diagonal of $AB = (\alpha_{i,q}) = \sum_{p=1}^{n} a_{i,p} b_{p,q}$ are zero for $2 \leq k < 2k-1 \leq n$.*

Proof. We need only show that the entries of the arbitrary *mth* diagonal above the principal diagonal of AB are zero for $k-1 < m \leq 2k-1$. We have $a_{i,p} = 0$ whenever $i > p$ or $|i-p| < k$ and $b_{p,q} = 0$ whenever $p > q$ or $|p-q| < k$. The entries of the *mth* diagonal above the principal diagonal of AB are

$$\begin{aligned} \alpha_{i,i+m} &= \sum_{p=1}^{n} a_{i,p} b_{p,i+m} \\ &= \sum_{p=1}^{i} a_{i,p} b_{p,i+m} + \sum_{p=i+1}^{i+(k-1)} a_{i,p} b_{p,i+m} + \sum_{p=i+k}^{i+(m-1)} a_{i,p} b_{p,i+m} + \sum_{p=i+m}^{n} a_{i,p} b_{p,i+m} \\ &= \sum_{p=1}^{i} 0 \cdot b_{p,i+m} + \sum_{p=i+1}^{i+(k-1)} 0 \cdot b_{p,i+m} + \sum_{p=i+k}^{i+(m-1)} a_{i,p} \cdot 0 + \sum_{p=i+m}^{n} a_{i,p} \cdot 0 \\ &= 0. \end{aligned}$$

We have $b_{p,i+m} = 0$ for $p = i+k, ..., i+(m-1)$ since $|p-(i+m)| = |i+t-(i+m)| = |t-m|$ for $t = k, \ldots, m-1$, and

$$|t - m| = |k + x - m| = |k - (m - x)| < k$$

for $t = k + x$ with $x = 0, 1, \ldots, m-(k+1)$. Or we can let $t = m - x$ with $x = 1, \ldots, m-k$ so that

$$|t - m| = |m - x - m| = |-x| = x < k.$$

□

Remark 48. We note that if, in Lemma 47, $2k-1=n-1$, then AB is the zero matrix.

Let A_1, A_2, B_1 and B_2 be matrices in $U_n^*(R)$. By Proposition 40, the principal diagonal and the first diagonal above the principal diagonal of $[A_1, B_1]$ and $[A_2, B_2]$ are zero. Thus, it follows from Lemma 47 above (with $k=2$) that the principal diagonal and the first three diagonals above the principal diagonal of $[A_1, B_1][A_2, B_2]$ are zero. Thus, if $n=4$ we have $[A_1, B_1][A_2, B_2]=0$ (which agrees with Proposition 46).

Now, let $[A_1, B_1][A_2, B_2]=(a_{i,j})$ and $[A_3, B_3]=(b_{j,k})$. Then the entries of the $4th$ diagonal above the principal diagonal of $[A_1, B_1][A_2, B_2][A_3, B_3]=(\alpha_{i,k})$ are

$$\begin{aligned}
\alpha_{i,i+4} &= \sum_{p=1}^{n} a_{i,p}b_{p,i+4} \\
&= \sum_{p=1}^{i} a_{i,p}b_{p,i+4} + \sum_{p=i+1}^{i+3} a_{i,p}b_{p,i+4} + \sum_{p=i+4}^{n} a_{i,p}b_{p,i+4} \\
&= \sum_{p=1}^{i} 0 \cdot b_{p,i+4} + \sum_{p=i+1}^{i+3} 0 \cdot b_{p,i+4} + \sum_{p=i+4}^{n} a_{i,p} \cdot 0 \\
&= 0
\end{aligned}$$

and the entries of the $5th$ diagonal above the principal diagonal of $[A_1, B_1][A_2, B_2][A_3, B_3]$ are

$$\begin{aligned}
\alpha_{i,i+5} &= \sum_{p=1}^{n} a_{i,p}b_{p,i+5} \\
&= \sum_{p=1}^{i} a_{i,p}b_{p,i+5} + \sum_{p=i+1}^{i+3} a_{i,p}b_{p,i+5} + \sum_{p=i+4}^{n} a_{i,p}b_{p,i+5} \\
&= \sum_{p=1}^{i} 0 \cdot b_{p,i+5} + \sum_{p=i+1}^{i+3} 0 \cdot b_{p,i+5} + \sum_{p=i+4}^{n} a_{i,p} \cdot 0 \\
&= 0.
\end{aligned}$$

Thus if $n=5$ or 6 we have $[A_1, B_1][A_2, B_2][A_3, B_3]=0$. The entries of the $6th$ diagonal above the principal diagonal of $[A_1, B_1][A_2, B_2][A_3, B_3]$ are

$$\begin{aligned}
\alpha_{i,i+6} &= \sum_{p=1}^{n} a_{i,p}b_{p,i+6} \\
&= \sum_{p=1}^{i} a_{i,p}b_{p,i+6} + \sum_{p=i+1}^{i+3} a_{i,p}b_{p,i+6} + a_{i,i+4}b_{i+4,i+6} + \sum_{p=i+5}^{n} a_{i,p}b_{p,i+6} \\
&= \sum_{p=1}^{i} 0 \cdot b_{p,i+6} + \sum_{p=i+1}^{i+3} 0 \cdot b_{p,i+6} + a_{i,i+4}b_{i+4,i+6} + \sum_{p=i+5}^{n} a_{i,p} \cdot 0.
\end{aligned}$$

The first sum $\sum_{p=1}^{i} 0 \cdot b_{p,i+6}$ is zero because $a_{i,p} = 0$ when $i \geq p$, the second sum $\sum_{p=i+1}^{i+3} 0 \cdot b_{p,i+6}$ is zero because $a_{i,p} = 0$ when $|i-p| < 4$ and the last sum $\sum_{p=i+5}^{n} a_{i,p} \cdot 0$ is zero because $b_{p,i+6} = 0$ when $p > i+6$ or $|p-(i+6)| < 2$. Thus $\alpha_{i,i+6} = a_{i,i+4}b_{i+4,i+6}$, and since we are dealing with a ring R with unity 1, by choosing $a_{i,i+4} = 1 = b_{i+4,i+6}$ we see that $\alpha_{i,i+6} \neq 0$. This shows that for $n \geq 7$, $[A_1, B_1][A_2, B_2][A_3, B_3] \neq 0$.

Proposition 49. *Let R be an arbitrary ring with unity 1 and $A = (a_{i,p})$, $B = (b_{p,q}) \in U_n^*(R)$. Suppose the principal diagonal and the first t diagonals above the principal diagonal of A are zero and the principal diagonal and the first s diagonals above the principal diagonal of B are zero. Then the principal diagonal and the first $t+s+1$ diagonals above the principal diagonal of AB are zero.*

Proof. Without loss of generality, let $t < s$. Obviously, the principal diagonal and the first s diagonals above the principal diagonal of

$$AB = (\alpha_{i,q}) = \left(\sum_{p=1}^{n} a_{i,p}b_{p,q}\right)$$

are zero. We need only show that the entries of the arbitrary mth diagonal above the principal diagonal of AB are zero for $s < m \leq t+s+1$. We have $a_{i,p} = 0$ whenever $i > p$ or $|i-p| < t+1$ and $b_{p,q} = 0$ whenever $p > q$ or $|p-q| < s+1$. The entries of the mth diagonal above the principal diagonal of AB are

$$\begin{aligned}
\alpha_{i,i+m} &= \sum_{p=1}^{n} a_{i,p}b_{p,i+m} \\
&= \sum_{p=1}^{i} a_{i,p}b_{p,i+m} + \sum_{p=i+1}^{i+t} a_{i,p}b_{p,i+m} + \sum_{p=i+t+1}^{i+m-1} a_{i,p}b_{p,i+m} + \sum_{p=i+m}^{n} a_{i,p}b_{p,i+m} \\
&= \sum_{p=1}^{i} 0 \cdot b_{p,i+m} + \sum_{p=i+1}^{i+t} 0 \cdot b_{p,i+m} + \sum_{p=i+t+1}^{i+m-1} a_{i,p} \cdot 0 + \sum_{p=i+m}^{n} a_{i,p} \cdot 0 \\
&= 0.
\end{aligned}$$

(The reason why, $b_{p,i+m} = 0$ in the third sum for $p = i+t+1, ..., i+(m-1)$ is because

$$|p-(i+m)| = |i+u-i-m| = |u-m|$$

for $u = t+1, \ldots, m-1$, and

$$|u-m| = |m-x-m| = |-x| = x < s+1$$

for $u = m-x$ with $x = 1, \ldots, m-(t+1)$.) □

Remark 50.

(i) In Proposition 49, if $t+s+1 \geq n-1$, then AB is the zero matrix.

(ii) Lemma 47 is a special case of Proposition 49 with $t = s = k-1$.

The example below (with $n = 7$), illustrates the observation stated in Remark 50 (i).

Example 51.

$$\begin{bmatrix} 0&0&0&0&a&b&c\\ 0&0&0&0&0&d&e\\ 0&0&0&0&0&0&f\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&0&0&g&h&i&j\\ 0&0&0&0&k&l&m\\ 0&0&0&0&0&p&q\\ 0&0&0&0&0&0&r\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} = 0$$

with $t + s + 1 = 3 + 2 + 1 = 6$ and

$$\begin{bmatrix} 0&0&0&g&h&i&j\\ 0&0&0&0&k&l&m\\ 0&0&0&0&0&p&q\\ 0&0&0&0&0&0&r\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&0&0&0&a&b&c\\ 0&0&0&0&0&d&e\\ 0&0&0&0&0&0&f\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} = 0$$

with $t + s + 1 = 2 + 3 + 1 = 6$.

We observe that the number of zero diagonals above the principal diagonal in a product of strictly upper triangular matrices is always greater than the number of such diagonals in individual factors.

Theorem 52. *Let R be a commutative ring and $A_i, B_i \in U_n^*(R)$, $i = 1, 2, \ldots, k$.*

(i) The first $2k - 1$ diagonals above the principal diagonal of $[A_1, B_1][A_2, B_2] \ldots [A_k, B_k]$ are zero diagonals.
(ii) The ring $U_n^*(R)$ satisfies the identity $[A_1, B_1][A_2, B_2] \ldots [A_k, B_k] = 0$ for $k = \frac{1}{2}n$ when n is even, and for $k = \frac{1}{2}(n + 1)$ when n is odd.

Proof. (i) Since, by Proposition 40, the first diagonal above the principal diagonal of a matrix $[A_i, B_i]$ is zero for all matrices $A_i, B_i \in U_n^*(R)$, it follows from Proposition 49 that

(a) the first $1 + 1 + 1 = 3$ diagonals above the principal diagonal of $[A_1, B_1][A_2, B_2]$ are zero and
(b) the first $3+1+1 = 5$ diagonals above the principal diagonal of $[A_1, B_1][A_2, B_2][A_3, B_3]$ are zero.

Assume that the first $2t - 1$ diagonals above the principal diagonal of

$$[A_1, B_1][A_2, B_2] \ldots [A_t, B_t]$$

are zero. Now, by Proposition 49 above, the first $(2t-1)+1+1 = 2(t+1)-1$ diagonals above the principal diagonal of

$$[A_1, B_1][A_2, B_2] \ldots [A_t, B_t][A_{t+1}, B_{t+1}]$$

are zero. Thus it follows, by induction, that the first $2k-1$ diagonals above the principal diagonal of $[A_1, B_1][A_2, B_2] \ldots [A_k, B_k]$ are zero (and diagonals $2k, 2k+1, \ldots, n-1$ are not necessarily zero).

(ii) For an $n \times n$ matrix the number of diagonals above the principal diagonal is $n-1$. Hence, for matrices $A_i, B_i \in U_n^*(R)$,

$$[A_1, B_1][A_2, B_2] \ldots [A_k, B_k] = 0$$

implies $2k-1 \geq n-1$, i.e., $2k \geq n$. Thus the smallest k for which the inequality $2k \geq n$ holds is $k = \frac{1}{2}n$ when n is even and $k = \frac{1}{2}(n+1)$ when n is odd.

Therefore $U_n^*(R)$ satisfies the identity $[A_1, B_1][A_2, B_2] \ldots [A_k, B_k] = 0$ for $k = \frac{1}{2}n$ when n is even, and for $k = \frac{1}{2}(n+1)$ when n is odd. □

Theorem 53. *Let R be a commutative ring and $A_k, B_k \in U_n^*(R)$ for $k \in \mathbb{Z}^+$. Then the ring $U_n^*(R)$ does not necessarily satisfy the identity $[A_1, B_1][A_2, B_2] \ldots [A_j, B_j] = 0$ if $j < \frac{1}{2}n$.*

Proof. Let $A_k, B_k \in U_n^*(R)$ such that $A_k = E_{2k-1,2k}$ and $B_k = E_{2k,2k+1}$. Then

$$\begin{aligned} [A_k, B_k] &= A_k B_k - B_k A_k \\ &= E_{2k-1,2k} E_{2k,2k+1} - E_{2k,2k+1} E_{2k-1,2k} \\ &= E_{2k-1,2k+1}. \end{aligned}$$

Thus, if $j < \frac{1}{2}n$, i.e., $2j+1 \leq n$, we have

$$\begin{aligned} & [A_1, B_1][A_2, B_2][A_3, B_3] \ldots [A_{j-1}, B_{j-1}][A_j, B_j] \\ = \; & E_{1,3} \cdot E_{3,5} \cdot E_{5,7} \cdots E_{2j-3,2j-1} \cdot E_{2j-1,2j+1} \\ = \; & E_{1,2j+1} \\ \neq \; & 0. \end{aligned}$$

□

Remark 54. It follows from Theorems 52 (ii) and 53 that when R is commutative, then $k = \frac{1}{2}n$ (if n is even) or $k = \frac{1}{2}(n+1)$ (if n odd) is the smallest value such that $[A_1, B_1][A_2, B_2] \ldots [A_k, B_k] = 0$ for $A_i, B_i \in U_n^*(R)$.

The example below shows that if the ring R is not commutative, then
(i) Proposition 46 does not necessarily hold, i.e., there are matrices in $U_4^*(R)$ for which the identity in Proposition 46 does not hold.
(ii) The ring $U_4^*(R)$ does not necessarily satisfy the identity $[[[X_1, X_2], X_3], X_4] = 0$.

Example 55. Let R be a noncommutative ring and I denote the identity in $U_4^*(R)$. Then $[r,s] \neq 0$ for some $r, s \in R$, and so for $X = rI$, $Y = sE_{1,2}$, $U = Z = E_{2,3}$, $V = E_{3,4}$ in $U_4^*(R)$, we have

$$\begin{aligned}
[X,Y][U,V] &= [rI, sE_{1,2}][E_{2,3}, E_{3,4}] \\
&= ((rI)(sE_{1,2}) - (sE_{1,2})(rI))(E_{2,3}E_{3,4} - E_{3,4}E_{2,3}) \\
&= (rs(IE_{1,2}) - sr(E_{1,2}I))(E_{2,4} - 0) \\
&= (rsE_{1,2} - srE_{1,2})E_{2,4} \\
&= ((rs - sr)E_{1,2})E_{2,4} \\
&= (rs - sr)E_{1,2}E_{2,4} \\
&= [r,s]E_{1,4} \\
&\neq 0
\end{aligned}$$

and

$$\begin{aligned}
[[[X,Y],Z],V] &= [[[r,s]E_{1,2}, E_{2,3}], E_{3,4}] \\
&= [[r,s]E_{1,3}, E_{3,4}] \\
&= [r,s]E_{1,3}E_{3,4} - [r,s]E_{3,4}E_{1,3} \\
&= [r,s]E_{1,4} \\
&\neq 0.
\end{aligned}$$

Remark 56. Suppose $[r,s] \neq 0$ for some $r, s \in R$ and let $A_k, B_k \in U_n^*(R)$ such that $A_k = E_{2k-1,2k}$ and $B_k = E_{2k,2k+1}$ for $k = 1, \ldots, \frac{1}{2}n - 1$ and $A_k = B_k = I$ (I the identity in $U_n^*(R)$) for $k = \frac{1}{2}n$, (n is even). Then, for $k = 1, \ldots, \frac{1}{2}n - 1$, we have

$$[A_k, B_k] = E_{2k-1,2k+1}.$$

Thus,

$$\begin{aligned}
& [A_1,B_1][A_2,B_2][A_3,B_3]\ldots[A_{\frac{1}{2}n-2}, B_{\frac{1}{2}n-2}][A_{\frac{1}{2}n-1}, B_{\frac{1}{2}n-1}][rI, sI] \\
= {} & E_{1,3} \cdot E_{3,5} \cdot E_{5,7} \cdot \ldots \cdot E_{n-5,n-3} \cdot E_{n-3,n-1} \cdot [r,s]I \\
= {} & E_{1,n-1} \cdot [r,s]I \\
= {} & [r,s]E_{1,n-1} \\
\neq {} & 0,
\end{aligned}$$

showing that for a noncommutative ring R the ring $U_n^*(R)$ does not necessarily satisfy the identity $[A_1,B_1][A_2,B_2]\ldots[A_k,B_k] = 0$ for all $k \leq \frac{1}{2}n$.

We note that if $C \in U_4^*(R)$ is a strictly upper triangular matrix, then

(i) $E_{2,4}C = CE_{1,3} = 0$,
(ii) $CE_{2,4} = q_1E_{1,4}$, for some $q_1 \in R$,
(iii) $E_{1,3}C = q_2E_{1,4}$ for some $q_2 \in R$.

The items (i) - (iii) just mentioned above will be used in the proof of the following theorem which is similar to Theorem 36.

Theorem 57. *If S satisfies $[x,y][u,v]=0$ and $[[x,y],z]=0$ then $U_4^*(S)$ does not satisfy the polynomial identity $[[x,y],[u,v]]=0$.*

Proof. For

$$X=\begin{bmatrix} a & b & c & d\\ 0 & a & e & f\\ 0 & 0 & a & g\\ 0 & 0 & 0 & a\end{bmatrix} \text{ and } Y=\begin{bmatrix} h & i & j & k\\ 0 & h & l & m\\ 0 & 0 & h & n\\ 0 & 0 & 0 & h\end{bmatrix}$$

in $U_4^*(S)$ we have

$$\begin{aligned}
& [X,Y]\\
= \ & XY-YX\\
= \ & \begin{bmatrix} a & b & c & d\\ 0 & a & e & f\\ 0 & 0 & a & g\\ 0 & 0 & 0 & a\end{bmatrix}\begin{bmatrix} h & i & j & k\\ 0 & h & l & m\\ 0 & 0 & h & n\\ 0 & 0 & 0 & h\end{bmatrix}-\begin{bmatrix} h & i & j & k\\ 0 & h & l & m\\ 0 & 0 & h & n\\ 0 & 0 & 0 & h\end{bmatrix}\begin{bmatrix} a & b & c & d\\ 0 & a & e & f\\ 0 & 0 & a & g\\ 0 & 0 & 0 & a\end{bmatrix}\\
= \ & \begin{bmatrix} ah & ai+bh & aj+bl+ch & ak+bm+cn+dh\\ 0 & ah & al+eh & am+en+fh\\ 0 & 0 & ah & an+gh\\ 0 & 0 & 0 & ah\end{bmatrix}\\
& -\begin{bmatrix} ha & hb+ia & hc+ie+ja & hd+if+jg+ka\\ 0 & ha & he+la & hf+lg+ma\\ 0 & 0 & ha & hg+na\\ 0 & 0 & 0 & ha\end{bmatrix}\\
= \ & \begin{bmatrix} [a,h] & [a,i]+[b,h] & [a,j]+[c,h]+bl-ie & [a,k]+[d,h]+bm+cn-if-jg\\ 0 & [a,h] & [a,l]+[e,h] & [a,m]+[f,h]+en-lg\\ 0 & 0 & [a,h] & [a,n]+[g,h]\\ 0 & 0 & 0 & [a,h]\end{bmatrix}\\
= \ & [a,h]\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}+\begin{bmatrix} 0 & [a,i]+[b,h] & [a,j]+[c,h] & [a,k]+[d,h]\\ 0 & 0 & [a,l]+[e,h] & [a,m]+[f,h]\\ 0 & 0 & 0 & [a,n]+[g,h]\\ 0 & 0 & 0 & 0\end{bmatrix}\\
& +\begin{bmatrix} 0 & 0 & bl-ie & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0 & 0 & bm+cn-if-jg\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}+\begin{bmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & en-lg\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}.
\end{aligned}$$

Therefore

$$[X,Y] \quad = \quad [a,h]I+C+\alpha E_{1,3}+\beta E_{1,4}+\omega E_{2,4}, \tag{4.5}$$

for suitable $\alpha,\beta,\omega\in S$, where C is a strictly upper triangular matrix whose entries are in the subgroup $[S,S]$ of S generated by all commutators.

Similarly,

$$[U,V] = [a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4}$$

for some $U, V \in U_4^*(S)$ and suitable $\alpha', \beta', \omega' \in S$. Again, C' is a strictly upper triangular matrix whose entries are in the subgroup $[S,S]$ of S generated by all commutators. By hypothesis $[[a,h],s] = 0$ which implies $[a,h]s = s[a,h]$ for every $s \in S$. Thus, $[a,h]I$ is central in $U_4^*(S)$ and moreover also in $M_4(S)$. Now,

$$\begin{aligned}
& [[X,Y],[U,V]] \\
= & [[a,h]I + C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4}, [a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4}] \\
= & [a,h]I([a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4}) \\
& -([a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4})[a,h]I \\
& +(C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4})([a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4}) \\
& -([a',h']I + C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4})(C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4}) \\
= & (C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4})[a',h']I - [a',h']I(C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4}) \\
& +(C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4})(C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4}) \\
& -(C' + \alpha' E_{1,3} + \beta' E_{1,4} + \omega' E_{2,4})(C + \alpha E_{1,3} + \beta E_{1,4} + \omega E_{2,4}) \\
= & CC' + CE_{1,3}\alpha' + CE_{1,4}\beta' + CE_{2,4}\omega' + \alpha E_{1,3}C' + \beta E_{1,4}C' + \omega E_{2,4}C' - C'C \\
& -C'E_{1,3}\alpha - C'E_{1,4}\beta - C'E_{2,4}\omega - \alpha' E_{1,3}C - \beta' E_{1,4}C - \omega' E_{2,4}C \\
= & CE_{2,4}\omega' + \alpha E_{1,3}C' - C'E_{2,4}\omega - \alpha' E_{1,3}C \\
= & E_{1,3}(\alpha C' - \alpha' C) + (C\omega' - C'\omega)E_{2,4} \\
= & tE_{1,4} \\
\neq & 0 \qquad \text{provided } t \neq 0.
\end{aligned}$$

That is, $U_4^*(S)$ does not satisfy $[[X,Y],[U,V]] = 0$ whenever S satisfies $[x,y][u,v] = 0$ and $[[x,y],z] = 0$.

□

In the following example we show an instance where the value of t in the proof of Theorem 57 is not zero.

Example 58. Let R be a commutative ring. By Propositions 33 and 34, the ring $U_3^*(R)$ satisfies the identities $[[x,y],z] = 0$ and $[x,y][u,v] = 0$, respectively. Since $U_3^*(R)$ is never commutative, there exist matrices, say r and s in $U_3^*(R)$ such that $[r,s] \neq 0$. Now, for matrices $X = rI$, $Y = sE_{1,2}$, $U = E_{2,3}$, $V = E_{3,4}$ in $U_4^*(U_3^*(R))$ we have

$$\begin{aligned}
[X,Y][U,V] &= [rI, sE_{1,2}][E_{2,3}, E_{3,4}] \\
&= ((rs - sr)E_{1,2})E_{2,4} \\
&= [r,s]E_{1,4}
\end{aligned}$$

and

$$\begin{aligned} [U,V][X,Y] &= [E_{2,3},E_{3,4}][rI,sE_{1,2}] \\ &= E_{2,4}((rs-sr)E_{1,2}) \\ &= [r,s]E_{2,4}E_{1,2} \\ &= 0. \end{aligned}$$

Thus,

$$\begin{aligned} [[X,Y],[U,V]] &= [X,Y][U,V]-[U,V][X,Y] \\ &= [r,s]E_{1,4}, \end{aligned}$$

that is,

$$[[X,Y],[U,V]] = \begin{bmatrix} 0 & 0 & 0 & [r,s] \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Returning back to the proof of Theorem 57, we find that the value of t in the expression $tE_{1,4}$ is

$$t = \alpha[a',n'] + \alpha[g',h'] - \alpha'[a,n] - \alpha'[g,h] + [a,i]\omega' + [b,h]\omega' - [a',i']\omega - [b',h']\omega.$$

Thus,

$$tE_{1,4} = \begin{bmatrix} 0 & 0 & 0 & \alpha([a',n']+[g',h']) - \alpha'([a,n]+[g,h]) \\ & & & +([a,i]+[b,h])\omega' - ([a',i']+[b',h'])\omega \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For any $W = wI + P \in U_4^*(S)$ with P a strictly upper triangular matrix in $U_4^*(S)$, we find that

$$\begin{aligned} & [[[X,Y],[U,V]],W] \\ =\ & [tE_{1,4}, wI+P] \\ =\ & tE_{1,4}(wI+P) - (wI+P)tE_{1,4} \\ =\ & twE_{1,4} + tE_{1,4}P - wtE_{1,4} - PE_{1,4}t \\ =\ & twE_{1,4} - wtE_{1,4} \\ =\ & [t,w]E_{1,4} \end{aligned}$$

($E_{1,4}P = PE_{1,4} = 0$ since P is strictly upper triangular) and

$$\begin{aligned}
[t,w] &= [\alpha[a',n'] + \alpha[g',h'] - \alpha'[a,n] - \alpha'[g,h] + [a,i]\omega' + [b,h]\omega' - [a',i']\omega - [b',h']\omega, w] \\
&= (\alpha[a',n'] + \alpha[g',h'] - \alpha'[a,n] - \alpha'[g,h] + [a,i]\omega' + [b,h]\omega' - [a',i']\omega - [b',h']\omega)w \\
&\quad -w(\alpha[a',n'] + \alpha[g',h'] - \alpha'[a,n] - \alpha'[g,h] + [a,i]\omega' + [b,h]\omega' - [a',i']\omega - [b',h']\omega) \\
&= \alpha[a',n']w - w\alpha[a',n'] + \alpha[g',h']w - w\alpha[g',h'] - \alpha'[a,n]w + w\alpha'[a,n] - \alpha'[g,h]w \\
&\quad +w\alpha'[g,h] + [a,i]\omega'w - w[a,i]\omega' + [b,h]\omega'w - w[b,h]\omega' - [a',i']\omega w + w[a',i']\omega \\
&\quad -[b',h']\omega w + w[b',h']\omega \\
&= [\alpha[a',n'],w] + [\alpha[g',h'],w] - [\alpha'[a,n],w] - [\alpha'[g,h],w] + [[a,i]\omega',w] + [[b,h]\omega',w] \\
&\quad -[[a',i']\omega,w] - [[b',h']\omega,w].
\end{aligned}$$

Thus,

$$\begin{aligned}
[t,w] &= [\alpha[a',n'],w] + [\alpha[g',h'],w] - [\alpha'[a,n],w] - [\alpha'[g,h],w] + [[a,i]\omega',w] \\
&\quad +[[b,h]\omega',w] - [[a',i']\omega,w] - [[b',h']\omega,w]. \qquad (4.6)
\end{aligned}$$

Using commutator identity $[AB,C] = A[B,C] + [A,C]B$ we find that each of the first four terms on the right of (4.6) is of the form

$$[r[x,y],s] = r[[x,y],s] + [r,s][x,y]$$

and each the last four terms is of the form

$$[[x,y]r,s] = [x,y][r,s] + [[x,y],s]r.$$

Since ring S in Theorem 57 satisfies the identities $[x,y][u,v] = 0$ and $[[x,y],z] = 0$ it follows that the eight terms on the right of (4.6) are all zero and hence $[t,w] = 0$. Therefore $[[[X,Y],[U,V]],W] = 0$.

Definition 59. A ring R is said to be Lie solvable of index 3 if it satisfies the identity

$$[[[x,y],[u,v]],[[w,z],[p,q]]] = 0.$$

We observe in this chapter that when a ring R is commutative
(i) the algebra $U_n^*(R)$ is Lie nilpotent of index $n-1$,
(ii) the algebras $U_3^*(R)$ and $U_4^*(R)$ satisfy the identity $[x,y][w,z] = 0$ (and hence $[[x,y],[w,z]] = 0$) and when R is not commutative Examples 35, 45 and 55 show that (i) and (ii) do not necessarily hold. We also observe that whenever a ring S satisfies the identities $[x,y][u,v] = 0$ and $[[x,y],z] = 0$, then $U_3^*(S)$ is Lie solvable of index 2 and $U_4^*(S)$ satisfies the identity $[[[x,y],[u,v]],w] = 0$ and it follows that $U_4^*(S)$ is Lie solvable of index 3.

Chapter 5

Cayley-Hamilton Trace Identity for 2×2 Matrices

In this chapter we further dissect the paper [8] by Meyer, et al. We discuss the Cayley-Hamilton trace identity for 2×2 matrices over a ring R with $\frac{1}{2} \in R$ that satisfy the identity $[[x, y], [x, z]] = 0$. We also look at the result showing that for a ring containing $\frac{1}{2}$, the identity $[[x, y], [x, z]] = 0$ implies $[[x, y], [u, v]] = 0$.

We will use the following commutator identities throughout the text.
(i) $[a, -a] = a(-a) - (-a)a = 0$,
(ii) $[a, b] = ab - ba = -(ba - ab) = -[b, a]$,
(iii) $[a, -b] = a(-b) - (-b)a = -(ab - ba) = -[a, b]$, and similarly $[-a, b] = -[a, b]$.
From identity (iii) it follows that
(iv) $[-[a, b], [c, d]] = -[[a, b], [c, d]]$,
(v) $-[-[a, b], \frac{1}{2}[c, d]] = \frac{1}{2}[[a, b], [c, d]]$.

5.1 Trace identities for 2×2 matrices

It is stated in Chapter 3 that the right hand side of (3.19) is not necessarily zero when the underlying ring R is noncommutative. The following proposition which can be considered as a "real" 2×2 Cayley-Hamilton trace identity gives the right hand side of (3.19) as a matrix with commutator entries. Note that if R is commutative, then the matrix on the right of Proposition 60 (i) becomes zero.

Proposition 60. *If* $\frac{1}{2} \in R$ *and* $A = (a_{ij}) \in M_2(R)$, *then*

(i) $A^2 - \mathrm{tr}(A)A + \frac{1}{2}(\mathrm{tr}^2(A) - \mathrm{tr}(A^2))I = \begin{bmatrix} \frac{1}{2}[a_{11}, a_{22}] + \frac{1}{2}[a_{12}, a_{21}] & [a_{12}, a_{22}] \\ [a_{21}, a_{11}] & -\frac{1}{2}[a_{11}, a_{22}] - \frac{1}{2}[a_{12}, a_{21}] \end{bmatrix}.$
(ii) $\mathrm{tr}(A^2 - \mathrm{tr}(A)A + \frac{1}{2}(\mathrm{tr}^2(A) - \mathrm{tr}(A^2))I) = 0.$

Proof. Consider a matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in M_2(R).$$

(i) Then $A^2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}a_{11} + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{21}a_{11} + a_{22}a_{21} & a_{21}a_{12} + a_{22}a_{22} \end{bmatrix}$,

$\text{tr}(A) = a_{11} + a_{22}$ and $\text{tr}(A^2) = a_{11}a_{11} + a_{12}a_{21} + a_{21}a_{12} + a_{22}a_{22}$.

Now

$$\begin{aligned}
& A^2 - \text{tr}(A)A + \frac{1}{2}(\text{tr}^2(A) - \text{tr}(A^2))I \\
= & \begin{bmatrix} a_{11}a_{11} + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{21}a_{11} + a_{22}a_{21} & a_{21}a_{12} + a_{22}a_{22} \end{bmatrix} - (a_{11} + a_{22}) \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\
& + \frac{1}{2}((a_{11} + a_{22})^2 - (a_{11}a_{11} + a_{12}a_{21} + a_{21}a_{12} + a_{22}a_{22})) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
= & \begin{bmatrix} a_{11}a_{11} + a_{12}a_{21} & a_{11}a_{12} + a_{12}a_{22} \\ a_{21}a_{11} + a_{22}a_{21} & a_{21}a_{12} + a_{22}a_{22} \end{bmatrix} - \begin{bmatrix} a_{11}a_{11} + a_{22}a_{11} & a_{11}a_{12} + a_{22}a_{12} \\ a_{11}a_{21} + a_{22}a_{21} & a_{11}a_{22} + a_{22}a_{22} \end{bmatrix} \\
& + \frac{1}{2}(a_{11}a_{11} + a_{11}a_{22} + a_{22}a_{11} + a_{22}a_{22} - a_{11}a_{11} - a_{12}a_{21} - a_{21}a_{12} - a_{22}a_{22}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
= & \begin{bmatrix} a_{11}a_{11} + a_{12}a_{21} - (a_{11}a_{11} + a_{22}a_{11}) & a_{11}a_{12} + a_{12}a_{22} - (a_{11}a_{12} + a_{22}a_{12}) \\ a_{21}a_{11} + a_{22}a_{21} - (a_{11}a_{21} + a_{22}a_{21}) & a_{21}a_{12} + a_{22}a_{22} - (a_{11}a_{22} + a_{22}a_{22}) \end{bmatrix} \\
& + \frac{1}{2}(a_{11}a_{22} + a_{22}a_{11} - a_{12}a_{21} - a_{21}a_{12}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
= & \begin{bmatrix} a_{11}a_{11} + a_{12}a_{21} - a_{11}a_{11} - a_{22}a_{11} & a_{11}a_{12} + a_{12}a_{22} - a_{11}a_{12} - a_{22}a_{12} \\ a_{21}a_{11} + a_{22}a_{21} - a_{11}a_{21} - a_{22}a_{21} & a_{21}a_{12} + a_{22}a_{22} - a_{11}a_{22} - a_{22}a_{22} \end{bmatrix} \\
& + \begin{bmatrix} \frac{1}{2}(a_{11}a_{22} + a_{22}a_{11} - a_{12}a_{21} - a_{21}a_{12}) & 0 \\ 0 & \frac{1}{2}(a_{11}a_{22} + a_{22}a_{11} - a_{12}a_{21} - a_{21}a_{12}) \end{bmatrix} \\
= & \begin{bmatrix} a_{12}a_{21} - a_{22}a_{11} & a_{12}a_{22} - a_{22}a_{12} \\ a_{21}a_{11} - a_{11}a_{21} & a_{21}a_{12} - a_{11}a_{22} \end{bmatrix} \\
& + \begin{bmatrix} \frac{1}{2}(a_{11}a_{22} + a_{22}a_{11} - a_{12}a_{21} - a_{21}a_{12}) & 0 \\ 0 & \frac{1}{2}(a_{11}a_{22} + a_{22}a_{11} - a_{12}a_{21} - a_{21}a_{12}) \end{bmatrix} \\
= & \begin{bmatrix} \frac{1}{2}a_{11}a_{22} - \frac{1}{2}a_{22}a_{11} + \frac{1}{2}a_{12}a_{21} - \frac{1}{2}a_{21}a_{12} & a_{12}a_{22} - a_{22}a_{12} \\ a_{21}a_{11} - a_{11}a_{21} & -\frac{1}{2}a_{11}a_{22} + \frac{1}{2}a_{22}a_{11} - \frac{1}{2}a_{12}a_{21} + \frac{1}{2}a_{21}a_{12} \end{bmatrix} \\
= & \begin{bmatrix} \frac{1}{2}(a_{11}a_{22} - a_{22}a_{11}) + \frac{1}{2}(a_{12}a_{21} - a_{21}a_{12}) & a_{12}a_{22} - a_{22}a_{12} \\ a_{21}a_{11} - a_{11}a_{21} & -\frac{1}{2}(a_{11}a_{22} - a_{22}a_{11}) - \frac{1}{2}(a_{12}a_{21} - a_{21}a_{12}) \end{bmatrix} \\
= & \begin{bmatrix} \frac{1}{2}[a_{11}, a_{22}] + \frac{1}{2}[a_{12}, a_{21}] & [a_{12}, a_{22}] \\ [a_{21}, a_{11}] & -\frac{1}{2}[a_{11}, a_{22}] - \frac{1}{2}[a_{12}, a_{21}] \end{bmatrix}.
\end{aligned}$$

(ii) Clearly $\text{tr}(A^2 - \text{tr}(A)A + \frac{1}{2}(\text{tr}^2(A) - \text{tr}(A^2))I) = 0$. □

When the trace of a matrix is zero then the Cayley-Hamilton trace identity in Proposition 60 (i) reduces to a "simpler" form as in Corollary 61 below.

Corollary 61. *If* $\frac{1}{2} \in R$ *and* $B = (b_{ij}) \in M_2(R)$ *with* $\operatorname{tr}(B) = 0$, *then*

$$B^2 - \tfrac{1}{2}\operatorname{tr}(B^2)I = \begin{bmatrix} \frac{1}{2}[b_{12}, b_{21}] & -[b_{12}, b_{11}] \\ [b_{21}, b_{11}] & -\frac{1}{2}[b_{12}, b_{21}] \end{bmatrix}.$$

Proof. Consider a matrix

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

By Proposition 60 (i),

$$B^2 - \operatorname{tr}(B)B + \frac{1}{2}(\operatorname{tr}^2(B) - \operatorname{tr}(B^2))I \tag{5.1}$$

$$= \begin{bmatrix} \frac{1}{2}[b_{11}, b_{22}] + \frac{1}{2}[b_{12}, b_{21}] & [b_{12}, b_{22}] \\ [b_{21}, b_{11}] & -\frac{1}{2}[b_{11}, b_{22}] - \frac{1}{2}[b_{12}, b_{21}] \end{bmatrix}. \tag{5.2}$$

Now, suppose $\operatorname{tr}(B) = 0$. Then $\operatorname{tr}(B) = b_{11} + b_{22} = 0$ which implies that $b_{11} = -b_{22}$. By commutator identities (i) and (iii), it follows that

$$[b_{11}, b_{22}] = 0 \tag{5.3}$$

and

$$[b_{12}, b_{22}] = -[b_{12}, b_{11}]. \tag{5.4}$$

Substituting (5.3) and (5.4) into (5.2) gives

$$B^2 - \frac{1}{2}\operatorname{tr}(B^2)I = \begin{bmatrix} \frac{1}{2}[b_{12}, b_{21}] & -[b_{12}, b_{11}] \\ [b_{21}, b_{11}] & -\frac{1}{2}[b_{12}, b_{21}] \end{bmatrix}.$$

□

Lemma 62. *If* $\frac{1}{2} \in R$, *then* $\operatorname{tr}(X - \frac{1}{2}\operatorname{tr}(X)I) = 0$ *for every matrix* $X \in M_2(R)$.

Proof. Let $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $\operatorname{tr}(X) = a + d$ and $\frac{1}{2}\operatorname{tr}(X)I = \frac{1}{2}(a+d)I$. Thus

$$\begin{aligned} \operatorname{tr}(X - \frac{1}{2}\operatorname{tr}(X)I) &= \operatorname{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \frac{1}{2}(a+d) & 0 \\ 0 & \frac{1}{2}(a+d) \end{bmatrix}\right) \\ &= \operatorname{tr}\left(\begin{bmatrix} \frac{1}{2}(a-d) & b \\ c & -\frac{1}{2}(a-d) \end{bmatrix}\right) \\ &= \frac{1}{2}(a-d) - \frac{1}{2}(a-d) \\ &= 0. \end{aligned}$$

□

The following theorem is important for the proof of Corollary 65.

Theorem 63. *If* $\frac{1}{2} \in R$ *and* R *satisfies* $[[x,y],[x,z]] = 0$, *then*

$$\left(C^2 - \frac{1}{2}\,\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\,\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\,\mathrm{tr}(C^2)I\right)^2\right)I = 0$$

for all $C \in M_2(R)$ *with* $\mathrm{tr}(C) = 0$.

Proof. Take $C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$. In view of Corollary 61, we have

$$C^2 - \frac{1}{2}\mathrm{tr}(C^2)I = \begin{bmatrix} \frac{1}{2}[c_{12},c_{21}] & -[c_{12},c_{11}] \\ [c_{21},c_{11}] & -\frac{1}{2}[c_{12},c_{21}] \end{bmatrix}$$

where

$$C^2 = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} c_{11}c_{11} + c_{12}c_{21} & c_{11}c_{12} + c_{12}c_{22} \\ c_{21}c_{11} + c_{22}c_{21} & c_{21}c_{12} + c_{22}c_{22} \end{bmatrix},$$

$\mathrm{tr}(C^2) = c_{11}c_{11} + c_{12}c_{21} + c_{21}c_{12} + c_{22}c_{22}$ and

$$\begin{aligned} & \frac{1}{2}\mathrm{tr}(C^2)I \\ = \; & \frac{1}{2}(c_{11}c_{11} + c_{12}c_{21} + c_{21}c_{12} + c_{22}c_{22})\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ = \; & \begin{bmatrix} \frac{1}{2}(c_{11}c_{11} + c_{12}c_{21} + c_{21}c_{12} + c_{22}c_{22}) & 0 \\ 0 & \frac{1}{2}(c_{11}c_{11} + c_{12}c_{21} + c_{21}c_{12} + c_{22}c_{22}) \end{bmatrix}. \end{aligned}$$

By Lemma 62 above $\mathrm{tr}(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I) = 0$. Applying Corollary 61 to the matrix $B = C^2 - \frac{1}{2}\mathrm{tr}(C^2)I$ with $b_{11} = \frac{1}{2}[c_{12},c_{21}]$, $b_{12} = -[c_{12},c_{11}]$, $b_{21} = [c_{21},c_{11}]$ and $b_{22} = -\frac{1}{2}[c_{12},c_{21}]$ and using commutator identities (iv) and (v) gives

$$\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I \tag{5.5}$$

$$= \begin{bmatrix} \frac{1}{2}[-[c_{12},c_{11}],[c_{21},c_{11}]] & -[-[c_{12},c_{11}],\frac{1}{2}[c_{12},c_{21}]] \\ [[c_{21},c_{11}],\frac{1}{2}[c_{12},c_{21}]] & -\frac{1}{2}[-[c_{12},c_{11}],[c_{21},c_{11}]] \end{bmatrix} \tag{5.6}$$

$$= \frac{1}{2}\begin{bmatrix} -[[c_{12},c_{11}],[c_{21},c_{11}]] & [[c_{12},c_{11}],[c_{12},c_{21}]] \\ [[c_{21},c_{11}],[c_{12},c_{21}]] & [[c_{12},c_{11}],[c_{21},c_{11}]] \end{bmatrix}. \tag{5.7}$$

By commutator identity (ii) we have $[c_{12},c_{11}] = -[c_{11},c_{12}]$ and $[c_{21},c_{11}] = -[c_{11},c_{21}]$.

Thus substituting

$$\begin{aligned}[[c_{12}, c_{11}], [c_{21}, c_{11}]] &= [-[c_{11}, c_{12}], -[c_{11}, c_{21}]] \\ &= [[c_{11}, c_{12}], [c_{11}, c_{21}]]\end{aligned}$$

and

$$\begin{aligned}[[c_{21}, c_{11}], [c_{12}, c_{21}]] &= [[c_{21}, c_{11}], -[c_{21}, c_{12}]] \\ &= -[[c_{21}, c_{11}], [c_{21}, c_{12}]]\end{aligned}$$

into (5.7) yields

$$\begin{aligned}&\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I \\ = \ &\frac{1}{2}\begin{bmatrix} -[[c_{11}, c_{12}], [c_{11}, c_{21}]] & [[c_{12}, c_{11}], [c_{12}, c_{21}]] \\ -[[c_{21}, c_{11}], [c_{21}, c_{12}]] & [[c_{11}, c_{12}], [c_{11}, c_{21}]] \end{bmatrix}.\end{aligned}$$

Since R satisfies $[[x, y], [x, z]] = 0$, and each entry of the above 2×2 matrix is of the form $\pm[[x, y], [x, z]] = 0$, it follows that

$$\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I = 0.$$

□

Corollary 64. *Suppose $\frac{1}{2} \in R$ and R satisfies the identity $[[x, y], [x, z]] = 0$. If $C \in M_2(R)$ with $\mathrm{tr}(C) = \mathrm{tr}(C^2) = \mathrm{tr}(C^4) = 0$, then $C^4 = 0$.*

Proof. Since the hypotheses of Theorem 63 above are satisfied, it follows that

$$\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I = 0.$$

Thus,

$$\begin{aligned}0 &= C^2\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right) - \frac{1}{2}\mathrm{tr}(C^2)I\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right) - \frac{1}{2}\mathrm{tr}\Big[C^2\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right) \\ &\quad -\frac{1}{2}\mathrm{tr}(C^2)I\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)\Big]I \\ &= C^4 - \frac{1}{2}C^2\mathrm{tr}(C^2)I - \frac{1}{2}\mathrm{tr}(C^2)IC^2 + \frac{1}{4}\mathrm{tr}^2(C^2)I^2 - \frac{1}{2}\mathrm{tr}\Big[C^4 - \frac{1}{2}C^2\mathrm{tr}(C^2)I - \frac{1}{2}\mathrm{tr}(C^2)IC^2 \\ &\quad +\frac{1}{4}\mathrm{tr}^2(C^2)I^2\Big]I \\ &= C^4,\end{aligned}$$

where the last equality follows from $\mathrm{tr}(C^2) = \mathrm{tr}(C^4) = 0$. □

Corollary 65. *If $\frac{1}{2} \in R$ and R is a ring satisfying $[[x,y],[x,z]] = 0$, then for all $A \in M_2(R)$ we have*

$$\begin{aligned}
& A^4 - \frac{1}{2}A^2\operatorname{tr}(A)A - \frac{1}{2}A\operatorname{tr}(A)A^2 - \frac{1}{2}A^3\operatorname{tr}(A) - \frac{1}{2}\operatorname{tr}(A)A^3 + \frac{1}{2}A^2\operatorname{tr}^2(A) + \frac{1}{2}\operatorname{tr}^2(A)A^2 \\
& -\frac{1}{2}A^2\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}(A^2)A^2 + \frac{1}{4}A\operatorname{tr}(A)A\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}(A)A + \frac{1}{4}\operatorname{tr}(A)A^2\operatorname{tr}(A) \\
& +\frac{1}{4}A\operatorname{tr}^2(A)A - \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}^2(A) - \frac{1}{4}\operatorname{tr}^2(A)A\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}(A^2) + \frac{1}{4}\operatorname{tr}(A^2)A\operatorname{tr}(A) \\
& -\frac{1}{4}A\operatorname{tr}^3(A) - \frac{1}{4}\operatorname{tr}^3(A)A + \frac{1}{4}A\operatorname{tr}(A)\operatorname{tr}(A^2) + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}(A)A - \frac{1}{2}\operatorname{tr}^2(A)\operatorname{tr}(A^2)I \\
& -\frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}^2(A)I + \frac{1}{2}\operatorname{tr}^2(A^2)I + \frac{1}{4}\operatorname{tr}(A^2\operatorname{tr}(A)A)I + \frac{1}{4}\operatorname{tr}((A)tr(A)A^2)I + \frac{1}{4}\operatorname{tr}(A^3)\operatorname{tr}(A)I \\
& +\frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A^3)I - \frac{1}{8}\operatorname{tr}(A)\operatorname{tr}(A\operatorname{tr}(A)A)I - \frac{1}{8}\operatorname{tr}(A\operatorname{tr}(A)A)\operatorname{tr}(A)I - \frac{1}{8}\operatorname{tr}(A\operatorname{tr}^2(A)A)I \\
& -\frac{1}{8}\operatorname{tr}(A)\operatorname{tr}(A^2)\operatorname{tr}(A)I + \frac{1}{2}\operatorname{tr}^4(A)I - \frac{1}{2}\operatorname{tr}(A^4)I \\
= {} & 0.
\end{aligned}$$

Proof. By Lemma 62, $\operatorname{tr}(A - \frac{1}{2}\operatorname{tr}(A)I) = 0$ and hence the hypotheses of Theorem 63 hold for $C = A - \frac{1}{2}\operatorname{tr}(A)I$. Then

$$\begin{aligned}
& C^2 - \frac{1}{2}\operatorname{tr}(C^2)I \\
= {} & \left(A - \frac{1}{2}\operatorname{tr}(A)I\right)^2 - \frac{1}{2}\operatorname{tr}\left(\left(A - \frac{1}{2}\operatorname{tr}(A)I\right)^2\right)I \\
= {} & \left(A - \frac{1}{2}\operatorname{tr}(A)I\right)\left(A - \frac{1}{2}\operatorname{tr}(A)I\right) - \frac{1}{2}\operatorname{tr}\left[\left(A - \frac{1}{2}\operatorname{tr}(A)I\right)\left(A - \frac{1}{2}\operatorname{tr}(A)I\right)\right]I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)IA - \frac{1}{2}A\operatorname{tr}(A)I + \frac{1}{4}\operatorname{tr}(A)I\operatorname{tr}(A)I - \frac{1}{2}\operatorname{tr}\Big[A^2 - \frac{1}{2}\operatorname{tr}(A)IA - \frac{1}{2}A\operatorname{tr}(A)I \\
& +\frac{1}{4}\operatorname{tr}(A)I\operatorname{tr}(A)I\Big]I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)IA - \frac{1}{2}A\operatorname{tr}(A)I + \frac{1}{4}\operatorname{tr}(A)I\operatorname{tr}(A)I - \frac{1}{2}\operatorname{tr}(A^2)I + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A)IA)I + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)I)I \\
& -\frac{1}{8}\operatorname{tr}(\operatorname{tr}(A)I\operatorname{tr}(A)I)I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)A - \frac{1}{2}A\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}^2(A)I - \frac{1}{2}\operatorname{tr}(A^2)I + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A)I + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A)I \\
& -\frac{1}{8}\operatorname{tr}(\operatorname{tr}^2(A)I)I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)A - \frac{1}{2}A\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}^2(A)I - \frac{1}{2}\operatorname{tr}(A^2)I + \frac{1}{4}\operatorname{tr}^2(A)I + \frac{1}{4}\operatorname{tr}^2(A)I - \frac{1}{8}(\operatorname{tr}^2(A) + \operatorname{tr}^2(A))I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)A - \frac{1}{2}A\operatorname{tr}(A) + \frac{1}{2}\operatorname{tr}^2(A)I - \frac{1}{2}\operatorname{tr}(A^2)I + \frac{1}{4}\operatorname{tr}^2(A)I - \frac{1}{4}\operatorname{tr}^2(A)I \\
= {} & A^2 - \frac{1}{2}\operatorname{tr}(A)A - \frac{1}{2}A\operatorname{tr}(A) + \frac{1}{2}\operatorname{tr}^2(A)I - \frac{1}{2}\operatorname{tr}(A^2)I,
\end{aligned}$$

and so

$$
\begin{aligned}
&\left(C^2 - \frac{1}{2}\text{tr}(C^2)I\right)^2 \\
= \;& \left(A^2 - \frac{1}{2}A\text{tr}(A) - \frac{1}{2}\text{tr}(A)A + \frac{1}{2}\text{tr}^2(A)I - \frac{1}{2}\text{tr}(A^2)I\right)^2 \\
= \;& \left(A^2 - \frac{1}{2}A\text{tr}(A) - \frac{1}{2}\text{tr}(A)A + \frac{1}{2}\text{tr}^2(A)I - \frac{1}{2}\text{tr}(A^2)I\right)\left(A^2 - \frac{1}{2}A\text{tr}(A) - \frac{1}{2}\text{tr}(A)A + \frac{1}{2}\text{tr}^2(A)I\right. \\
& \left.-\frac{1}{2}\text{tr}(A^2)I\right) \\
= \;& A^4 - \frac{1}{2}A\text{tr}(A)A^2 - \frac{1}{2}\text{tr}(A)A^3 + \frac{1}{2}\text{tr}^2(A)IA^2 - \frac{1}{2}\text{tr}(A^2)IA^2 - \frac{1}{2}A^3\text{tr}(A) + \frac{1}{4}A\text{tr}(A)A\text{tr}(A) \\
& + \frac{1}{4}\text{tr}(A)A^2\text{tr}(A) - \frac{1}{4}\text{tr}^2(A)IA\text{tr}(A) + \frac{1}{4}\text{tr}(A^2)IA\text{tr}(A) - \frac{1}{2}A^2\text{tr}(A)A + \frac{1}{4}A\text{tr}^2(A)A \\
& + \frac{1}{4}\text{tr}(A)A\text{tr}(A)A - \frac{1}{4}\text{tr}^2(A)I\text{tr}(A)A + \frac{1}{4}\text{tr}(A^2)I\text{tr}(A)A + \frac{1}{2}A^2\text{tr}^2(A)I - \frac{1}{4}A\text{tr}^3(A)I \\
& - \frac{1}{4}\text{tr}(A)A\text{tr}^2(A)I + \frac{1}{4}\text{tr}^2(A)I\text{tr}^2(A)I - \frac{1}{4}\text{tr}(A^2)I\text{tr}^2(A)I - \frac{1}{2}A^2\text{tr}(A^2)I + \frac{1}{4}A\text{tr}(A)\text{tr}(A^2)I \\
& + \frac{1}{4}\text{tr}(A)A\text{tr}(A^2)I - \frac{1}{4}\text{tr}^2(A)I\text{tr}(A^2)I + \frac{1}{4}\text{tr}(A^2)I\text{tr}(A^2)I \\
= \;& A^4 - \frac{1}{2}A\text{tr}(A)A^2 - \frac{1}{2}\text{tr}(A)A^3 + \frac{1}{2}\text{tr}^2(A)A^2 - \frac{1}{2}\text{tr}(A^2)A^2 - \frac{1}{2}A^3\text{tr}(A) + \frac{1}{4}A\text{tr}(A)A\text{tr}(A) \\
& + \frac{1}{4}\text{tr}(A)A^2\text{tr}(A) - \frac{1}{4}\text{tr}^2(A)A\text{tr}(A) + \frac{1}{4}\text{tr}(A^2)A\text{tr}(A) - \frac{1}{2}A^2\text{tr}(A)A + \frac{1}{4}A\text{tr}^2(A)A \\
& + \frac{1}{4}\text{tr}(A)A\text{tr}(A)A - \frac{1}{4}\text{tr}^3(A)A + \frac{1}{4}\text{tr}(A^2)\text{tr}(A)A + \frac{1}{2}A^2\text{tr}^2(A) - \frac{1}{4}A\text{tr}^3(A) \\
& - \frac{1}{4}\text{tr}(A)A\text{tr}^2(A) + \frac{1}{4}\text{tr}^4(A)I - \frac{1}{4}\text{tr}(A^2)\text{tr}^2(A)I - \frac{1}{2}A^2\text{tr}(A^2) + \frac{1}{4}A\text{tr}(A)\text{tr}(A^2) \\
& + \frac{1}{4}\text{tr}(A)A\text{tr}(A^2) - \frac{1}{4}\text{tr}^2(A)\text{tr}(A^2)I + \frac{1}{4}\text{tr}^2(A^2)I.
\end{aligned}
$$

Thus

$$
\begin{aligned}
& \operatorname{tr}\left(C^2 - \frac{1}{2}\operatorname{tr}(C^2)I\right)^2 \\
= & \operatorname{tr}\Big(A^4 - \frac{1}{2}A\operatorname{tr}(A)A^2 - \frac{1}{2}\operatorname{tr}(A)A^3 + \frac{1}{2}\operatorname{tr}^2(A)A^2 - \frac{1}{2}\operatorname{tr}(A^2)A^2 - \frac{1}{2}A^3\operatorname{tr}(A) + \frac{1}{4}A\operatorname{tr}(A)A\operatorname{tr}(A) \\
& + \frac{1}{4}\operatorname{tr}(A)A^2\operatorname{tr}(A) - \frac{1}{4}\operatorname{tr}^2(A)A\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A^2)A\operatorname{tr}(A) - \frac{1}{2}A^2\operatorname{tr}(A)A + \frac{1}{4}A\operatorname{tr}^2(A)A \\
& + \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}(A)A - \frac{1}{4}\operatorname{tr}^3(A)A + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}(A)A + \frac{1}{2}A^2\operatorname{tr}^2(A) - \frac{1}{4}A\operatorname{tr}^3(A) \\
& - \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}^2(A) + \frac{1}{4}\operatorname{tr}^4(A)I - \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}^2(A)I - \frac{1}{2}A^2\operatorname{tr}(A^2) + \frac{1}{4}A\operatorname{tr}(A)\operatorname{tr}(A^2) \\
& + \frac{1}{4}\operatorname{tr}(A)A\operatorname{tr}(A^2) - \frac{1}{4}\operatorname{tr}^2(A)\operatorname{tr}(A^2)I + \frac{1}{4}\operatorname{tr}^2(A^2)I\Big) \\
= & \operatorname{tr}(A^4) - \frac{1}{2}\operatorname{tr}(A\operatorname{tr}(A)A^2) - \frac{1}{2}\operatorname{tr}(\operatorname{tr}(A)A^3) + \frac{1}{2}\operatorname{tr}(\operatorname{tr}^2(A)A^2) - \frac{1}{2}\operatorname{tr}(\operatorname{tr}(A^2)A^2) - \frac{1}{2}\operatorname{tr}(A^3\operatorname{tr}(A)) \\
& + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)A\operatorname{tr}(A)) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A)A^2\operatorname{tr}(A)) - \frac{1}{4}\operatorname{tr}(\operatorname{tr}^2(A)A\operatorname{tr}(A)) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A^2)A\operatorname{tr}(A)) \\
& - \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}(A)A) + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}^2(A)A) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A)A\operatorname{tr}(A)A) - \frac{1}{4}\operatorname{tr}(\operatorname{tr}^3(A)A) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A^2)\operatorname{tr}(A)A) \\
& + \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}^2(A)) - \frac{1}{4}\operatorname{tr}(A\operatorname{tr}^3(A)) - \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A)A\operatorname{tr}^2(A)) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}^4(A)I) - \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A^2)\operatorname{tr}^2(A)I) \\
& - \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}(A^2)) + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)\operatorname{tr}(A^2)) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}(A)A\operatorname{tr}(A^2)) - \frac{1}{4}\operatorname{tr}(\operatorname{tr}^2(A)\operatorname{tr}(A^2)I) + \frac{1}{4}\operatorname{tr}(\operatorname{tr}^2(A^2)I) \\
= & \operatorname{tr}(A^4) - \frac{1}{2}\operatorname{tr}(A\operatorname{tr}(A)A^2) - \frac{1}{2}\operatorname{tr}(A)\operatorname{tr}(A^3) + \frac{1}{2}\operatorname{tr}^2(A)\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}(A^3)\operatorname{tr}(A) \\
& + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)A)\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A^2)\operatorname{tr}(A) - \frac{1}{4}\operatorname{tr}^2(A)\operatorname{tr}(A)\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}(A)\operatorname{tr}(A) \\
& - \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}(A)A) + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}^2(A)A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A\operatorname{tr}(A)A) - \frac{1}{4}\operatorname{tr}^3(A)\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}(A)\operatorname{tr}(A) \\
& + \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}^2(A) - \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}^3(A) - \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A)\operatorname{tr}^2(A) + \frac{1}{4}(\operatorname{tr}^4(A) + \operatorname{tr}^4(A)) \\
& - \frac{1}{4}(\operatorname{tr}(A^2)\operatorname{tr}^2(A) + \operatorname{tr}(A^2)\operatorname{tr}^2(A)) - \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}(A^2) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A)\operatorname{tr}(A^2) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A)\operatorname{tr}(A^2) \\
& - \frac{1}{4}(\operatorname{tr}^2(A)\operatorname{tr}(A^2) + \operatorname{tr}^2(A)\operatorname{tr}(A^2)) + \frac{1}{4}(\operatorname{tr}^2(A^2) + \operatorname{tr}^2(A^2)) \\
= & \operatorname{tr}(A^4) - \frac{1}{2}\operatorname{tr}(A\operatorname{tr}(A)A^2) - \frac{1}{2}\operatorname{tr}(A)\operatorname{tr}(A^3) + \frac{1}{2}\operatorname{tr}^2(A)\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}^2(A^2) - \frac{1}{2}\operatorname{tr}(A^3)\operatorname{tr}(A) \\
& + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)A)\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A^2)\operatorname{tr}(A) - \frac{1}{4}\operatorname{tr}^4(A) + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}^2(A) \\
& - \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}(A)A) + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}^2(A)A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A\operatorname{tr}(A)A) - \frac{1}{4}\operatorname{tr}^4(A) + \frac{1}{4}\operatorname{tr}(A^2)\operatorname{tr}^2(A) \\
& + \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}^2(A) - \frac{1}{4}\operatorname{tr}^4(A) - \frac{1}{4}\operatorname{tr}^4(A) + \frac{1}{2}\operatorname{tr}^4(A) - \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}^2(A) - \frac{1}{2}\operatorname{tr}^2(A^2) + \frac{1}{4}\operatorname{tr}^2(A)\operatorname{tr}(A^2) \\
& + \frac{1}{4}\operatorname{tr}^2(A)\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}^2(A)\operatorname{tr}(A^2) + \frac{1}{2}\operatorname{tr}^2(A^2) \\
= & \operatorname{tr}(A^4) - \frac{1}{2}\operatorname{tr}(A\operatorname{tr}(A)A^2) - \frac{1}{2}\operatorname{tr}(A)\operatorname{tr}(A^3) + \frac{1}{2}\operatorname{tr}^2(A)\operatorname{tr}(A^2) - \frac{1}{2}\operatorname{tr}^2(A^2) - \frac{1}{2}\operatorname{tr}(A^3)\operatorname{tr}(A) \\
& + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}(A)A)\operatorname{tr}(A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A^2)\operatorname{tr}(A) - \frac{1}{2}\operatorname{tr}^4(A) + \frac{1}{2}\operatorname{tr}(A^2)\operatorname{tr}^2(A) \\
& - \frac{1}{2}\operatorname{tr}(A^2\operatorname{tr}(A)A) + \frac{1}{4}\operatorname{tr}(A\operatorname{tr}^2(A)A) + \frac{1}{4}\operatorname{tr}(A)\operatorname{tr}(A\operatorname{tr}(A)A),
\end{aligned}
$$

which implies that

$$
\begin{aligned}
&\frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I \\
= \;& \frac{1}{2}\Bigg(\mathrm{tr}(A^4) - \frac{1}{2}\mathrm{tr}(A\mathrm{tr}(A)A^2) - \frac{1}{2}\mathrm{tr}(A)\mathrm{tr}(A^3) + \frac{1}{2}\mathrm{tr}^2(A)\mathrm{tr}(A^2) - \frac{1}{2}\mathrm{tr}^2(A^2) \\
& -\frac{1}{2}\mathrm{tr}(A^3)\mathrm{tr}(A) + \frac{1}{4}\mathrm{tr}(A\mathrm{tr}(A)A)\mathrm{tr}(A) + \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A^2)\mathrm{tr}(A) - \frac{1}{2}\mathrm{tr}^4(A) \\
& +\frac{1}{2}\mathrm{tr}(A^2)\mathrm{tr}^2(A) - \frac{1}{2}\mathrm{tr}(A^2\mathrm{tr}(A)A) + \frac{1}{4}\mathrm{tr}(A\mathrm{tr}^2(A)A) + \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A\mathrm{tr}(A)A)\Bigg)I \\
= \;& \frac{1}{2}\mathrm{tr}(A^4)I - \frac{1}{4}\mathrm{tr}(A\mathrm{tr}(A)A^2)I - \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A^3)I + \frac{1}{4}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I - \frac{1}{4}\mathrm{tr}^2(A^2)I \\
& -\frac{1}{4}\mathrm{tr}(A^3)\mathrm{tr}(A)I + \frac{1}{8}\mathrm{tr}(A\mathrm{tr}(A)A)\mathrm{tr}(A)I + \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A^2)\mathrm{tr}(A)I - \frac{1}{4}\mathrm{tr}^4(A)I \\
& +\frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I - \frac{1}{4}\mathrm{tr}(A^2\mathrm{tr}(A)A)I + \frac{1}{8}\mathrm{tr}(A\mathrm{tr}^2(A)A)I + \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A\mathrm{tr}(A)A)I.
\end{aligned}
$$

We conclude from Theorem 63 that

$$
\begin{aligned}
0 &= \left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2 - \frac{1}{2}\mathrm{tr}\left(\left(C^2 - \frac{1}{2}\mathrm{tr}(C^2)I\right)^2\right)I \\
&= A^4 - \frac{1}{2}A\mathrm{tr}(A)A^2 - \frac{1}{2}\mathrm{tr}(A)A^3 + \frac{1}{2}\mathrm{tr}^2(A)A^2 - \frac{1}{2}\mathrm{tr}(A^2)A^2 - \frac{1}{2}A^3\mathrm{tr}(A) + \frac{1}{4}A\mathrm{tr}(A)A\mathrm{tr}(A) \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A^2\mathrm{tr}(A) - \frac{1}{4}\mathrm{tr}^2(A)A\mathrm{tr}(A) + \frac{1}{4}\mathrm{tr}(A^2)A\mathrm{tr}(A) - \frac{1}{2}A^2\mathrm{tr}(A)A + \frac{1}{4}A\mathrm{tr}^2(A)A \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A)A - \frac{1}{4}\mathrm{tr}^3(A)A + \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}(A)A + \frac{1}{2}A^2\mathrm{tr}^2(A) - \frac{1}{4}A\mathrm{tr}^3(A) \\
&\quad - \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}^2(A) + \frac{1}{4}\mathrm{tr}^4(A)I - \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I - \frac{1}{2}A^2\mathrm{tr}(A^2) + \frac{1}{4}A\mathrm{tr}(A)\mathrm{tr}(A^2) \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A^2) - \frac{1}{4}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I + \frac{1}{4}\mathrm{tr}^2(A^2)I - \Bigg(\frac{1}{2}\mathrm{tr}(A^4)I - \frac{1}{4}\mathrm{tr}(A\mathrm{tr}(A)A^2)I \\
&\quad - \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A^3)I + \frac{1}{4}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I - \frac{1}{4}\mathrm{tr}^2(A^2)I - \frac{1}{4}\mathrm{tr}(A^3)\mathrm{tr}(A)I \\
&\quad + \frac{1}{8}\mathrm{tr}(A\mathrm{tr}(A)A)\mathrm{tr}(A)I + \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A^2)\mathrm{tr}(A)I - \frac{1}{4}\mathrm{tr}^4(A)I + \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I \\
&\quad - \frac{1}{4}\mathrm{tr}(A^2\mathrm{tr}(A)A)I + \frac{1}{8}\mathrm{tr}(A\mathrm{tr}^2(A)A)I + \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A\mathrm{tr}(A)A)I\Bigg) \\
&= A^4 - \frac{1}{2}A\mathrm{tr}(A)A^2 - \frac{1}{2}\mathrm{tr}(A)A^3 + \frac{1}{2}\mathrm{tr}^2(A)A^2 - \frac{1}{2}\mathrm{tr}(A^2)A^2 - \frac{1}{2}A^3\mathrm{tr}(A) + \frac{1}{4}A\mathrm{tr}(A)A\mathrm{tr}(A) \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A^2\mathrm{tr}(A) - \frac{1}{4}\mathrm{tr}^2(A)A\mathrm{tr}(A) + \frac{1}{4}\mathrm{tr}(A^2)A\mathrm{tr}(A) - \frac{1}{2}A^2\mathrm{tr}(A)A + \frac{1}{4}A\mathrm{tr}^2(A)A \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A)A - \frac{1}{4}\mathrm{tr}^3(A)A + \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}(A)A + \frac{1}{2}A^2\mathrm{tr}^2(A) - \frac{1}{4}A\mathrm{tr}^3(A) \\
&\quad - \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}^2(A) + \frac{1}{4}\mathrm{tr}^4(A)I - \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I - \frac{1}{2}A^2\mathrm{tr}(A^2) + \frac{1}{4}A\mathrm{tr}(A)\mathrm{tr}(A^2) \\
&\quad + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A^2) - \frac{1}{4}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I + \frac{1}{4}\mathrm{tr}^2(A^2)I - \frac{1}{2}\mathrm{tr}(A^4)I + \frac{1}{4}\mathrm{tr}(A\mathrm{tr}(A)A^2)I \\
&\quad + \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A^3)I - \frac{1}{4}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I + \frac{1}{4}\mathrm{tr}^2(A^2)I + \frac{1}{4}\mathrm{tr}(A^3)\mathrm{tr}(A)I \\
&\quad - \frac{1}{8}\mathrm{tr}(A\mathrm{tr}(A)A)\mathrm{tr}(A)I - \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A^2)\mathrm{tr}(A)I + \frac{1}{4}\mathrm{tr}^4(A)I - \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I \\
&\quad + \frac{1}{4}\mathrm{tr}(A^2\mathrm{tr}(A)A)I - \frac{1}{8}\mathrm{tr}(A\mathrm{tr}^2(A)A)I - \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A\mathrm{tr}(A)A)I
\end{aligned}
$$

and so

$$\begin{aligned}0 \;=\; & A^4 - \frac{1}{2}A\mathrm{tr}(A)A^2 - \frac{1}{2}\mathrm{tr}(A)A^3 + \frac{1}{2}\mathrm{tr}^2(A)A^2 - \frac{1}{2}\mathrm{tr}(A^2)A^2 - \frac{1}{2}A^3\mathrm{tr}(A) + \frac{1}{4}A\mathrm{tr}(A)A\mathrm{tr}(A) \\ & + \frac{1}{4}\mathrm{tr}(A)A^2\mathrm{tr}(A) - \frac{1}{4}\mathrm{tr}^2(A)A\mathrm{tr}(A) + \frac{1}{4}\mathrm{tr}(A^2)A\mathrm{tr}(A) - \frac{1}{2}A^2\mathrm{tr}(A)A + \frac{1}{4}A\mathrm{tr}^2(A)A \\ & + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A)A - \frac{1}{4}\mathrm{tr}^3(A)A + \frac{1}{4}\mathrm{tr}(A^2)\mathrm{tr}(A)A + \frac{1}{2}A^2\mathrm{tr}^2(A) - \frac{1}{4}A\mathrm{tr}^3(A) - \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}^2(A) \\ & - \frac{1}{2}\mathrm{tr}(A^2)\mathrm{tr}^2(A)I - \frac{1}{2}A^2\mathrm{tr}(A^2) + \frac{1}{4}A\mathrm{tr}(A)\mathrm{tr}(A^2) + \frac{1}{4}\mathrm{tr}(A)A\mathrm{tr}(A^2) - \frac{1}{2}\mathrm{tr}^2(A)\mathrm{tr}(A^2)I \\ & + \frac{1}{2}\mathrm{tr}^2(A^2)I - \frac{1}{2}\mathrm{tr}(A^4)I + \frac{1}{4}\mathrm{tr}(A\mathrm{tr}(A)A^2)I + \frac{1}{4}\mathrm{tr}(A)\mathrm{tr}(A^3)I + \frac{1}{4}\mathrm{tr}(A^3)\mathrm{tr}(A)I \\ & - \frac{1}{8}\mathrm{tr}(A\mathrm{tr}(A)A)\mathrm{tr}(A)I - \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A^2)\mathrm{tr}(A)I + \frac{1}{2}\mathrm{tr}^4(A)I + \frac{1}{4}\mathrm{tr}(A^2\mathrm{tr}(A)A)I \\ & - \frac{1}{8}\mathrm{tr}(A\mathrm{tr}^2(A)A)I - \frac{1}{8}\mathrm{tr}(A)\mathrm{tr}(A\mathrm{tr}(A)A)I.\end{aligned}$$

□

We observe that if $\mathrm{tr}(A) = 0$ for $A \in M_2(R)$ where R is a ring satisfying $[[x, y], [x, z]] = 0$ and $\frac{1}{2} \in R$, then the identity in Corollary 65 reduces to

$$A^4 - \frac{1}{2}\,\mathrm{tr}(A^2)A^2 - \frac{1}{2}A^2\,\mathrm{tr}(A^2) + \frac{1}{2}\,\mathrm{tr}^2(A^2)I - \frac{1}{2}\,\mathrm{tr}(A^4)I = 0,$$

which is also an identity of degree 4 in A.

In summary, the Cayley-Hamilton trace identity for any matrix $A \in M_2(R)$ assumes the expression in Proposition 60 (i) for an arbitrary ring R with $\frac{1}{2} \in R$ and reduces to the expression in Corollary 61 when the trace of the given matrix is zero. When R is commutative the right hand side of the expression in Proposition 60 becomes zero and gives the identity (3.19) in Chapter 3.

5.2 Relationship between identities

In Chapter 4 and also in this chapter we dealt with particular identities involving commutators in respect of matrices and entries of matrices. So we find it fit to end this thesis with a brief discussion on a relationship among some of such identities. We assume R to be noncommutative.

Suppose $[[x, y], z] = 0$ for all $x, y, z \in R$, then $[x, y]z = z[x, y]$. Now,

$$[x, y][x, z] = [x, y]xz - zx[x, y] = (x[x, y])z - z(x[x, y]) = [x[x, y], z].$$

But $x[x, y] = xxy - xyx = [x, xy]$, so $[x, y][x, z] = [[x, xy], z] = 0$ and consequently $[[x, y], [x, z]] = 0$.

We recall that a ring satisfying the identity $[[x,y],[u,v]] = 0$ is called Lie solvable of index 2. The next result shows that if a ring R, with $\frac{1}{2} \in R$, satisfies the identity $[[x,y],[x,z]] = 0$, then R satisfies the "seemingly stronger" identity $[[x,y],[u,v]] = 0$ (see [8]).

Proposition 66. *Let R be a ring with $\frac{1}{2} \in R$. If for all $x,y,z,w \in R$, $[[x,y],[x,z]] = 0$, then $[[x,y],[w,z]] = 0$.*

Proof. Set $F(x,y,z,w) := [[x,z],[y,w]] - [[x,y],[z,w]]$. Then

$$\begin{aligned}
& [[x+w,y],[x+w,z]] - [[x,y],[x,z]] - [[w,y],[w,z]] \\
= & [([x,y]+[w,y]),([x,z]+[w,z])] - [[x,y],[x,z]] - [[w,y],[w,z]] \\
= & ([x,y]+[w,y])([x,z]+[w,z]) - ([x,z]+[w,z])([x,y]+[w,y]) - [[x,y],[x,z]] - [[w,y],[w,z]] \\
= & [x,y][x,z] + [w,y][x,z] + [x,y][w,z] + [w,y][w,z] - [x,z][x,y] - [w,z][x,y] \\
& -[w,z][w,y] - [x,z][w,y] - [[x,y],[x,z]] - [[w,y],[w,z]] \\
= & [[x,y],[x,z]] + [[w,y],[x,z]] + [[x,y],[w,z]] + [[w,y],[w,z]] - [[x,y],[x,z]] - [[w,y],[w,z]] \\
= & [[w,y],[x,z]] + [[x,y],[w,z]] \\
= & -[[x,z],-[y,w]] + [[x,y],-[z,w]] \\
= & -(-[[x,z],[y,w]]) + (-[[x,y],[z,w]]) \\
= & [[x,z],[y,w]] - [[x,y],[z,w]] \\
= & F(x,y,z,w).
\end{aligned}$$

That is,

$$F(x,y,z,w) = [[x+w,y],[x+w,z]] - [[x,y],[x,z]] - [[w,y],[w,z]]. \tag{5.8}$$

Now, using the definition of F we find that

$$\begin{aligned}
& F(y,x,z,w) + F(x,y,w,z) \\
= & [[y,z],[x,w]] - [[y,x],[z,w]] + [[x,w],[y,z]] - [[x,y],[w,z]] \\
= & [[y,z],[x,w]] - [[y,x],[z,w]] - [[y,z],[x,w]] - [[x,y],[w,z]] \\
= & -[[y,x],[z,w]] - [[x,y],[w,z]] \\
= & -[-[x,y],[z,w]] - [[x,y],-[z,w]] \\
= & -(-[[x,y],[z,w]]) - (-[[x,y],[z,w]]) \\
= & 2[[x,y],[z,w]].
\end{aligned}$$

Thus

$$F(y,x,z,w) + F(x,y,w,z) = 2[[x,y],[z,w]]. \tag{5.9}$$

From (5.8) we find that

$$\begin{aligned}
& F(y,x,z,w) + F(x,y,w,z) \\
= & [[y+w,x],[y+w,z]] - [[y,x],[y,z]] - [[w,x],[w,z]] \\
& +[[x+z,y],[x+z,w]] - [[x,y],[x,w]] - [[z,y],[z,w]].
\end{aligned} \tag{5.10}$$

Equating (5.9) and (5.10) gives

$$\begin{aligned} & 2[[x,y],[z,w]] \\ = \; & [[y+w,x],[y+w,z]] - [[y,x],[y,z]] - [[w,x],[w,z]] \\ & +[[x+z,y],[x+z,w]] - [[x,y],[x,w]] - [[z,y],[z,w]]. \end{aligned} \tag{5.11}$$

Assuming that $[[x,y],[x,z]] = 0$ for every x, y, z in the ring R yields that each of the six terms on the right hand side of (5.11) is zero, and so (5.11) becomes

$$2[[x,y],[z,w]] = 0$$

and thus

$$[[x,y],[z,w]] = 0$$

if $\frac{1}{2} \in R$. Therefore $[[x,y],[x,z]] = 0$ implies $[[x,y],[z,w]] = 0$. □

For a ring R containing $\frac{1}{2}$, the observation following Corollary 65 together with Proposition 66 gives the following order of implication in respect of identities:

$$[[x,y],z] = 0 \Rightarrow [x,y][x,z] = 0 \Leftrightarrow [[x,y],[w,z]] = 0.$$

Thus, for a ring R with $\frac{1}{2} \in R$, Lie nilpotency of index 2 implies Lie solvability of index 2.

List of References

[1] Anton H., Elementary Linear Algebra 6th Ed., John Wiley, 1991.

[2] Gantmacher F.R., Matrix Theory vol. I, AMS Chelsea, 1959.

[3] Hungerford T.W., Algebra, Springer-Verlag, New York, 1974.

[4] Hoffman K. and Kunze R., Linear Algebra, Prentice-Hall, Inc. 1971.

[5] Jain S.K. and Gunawardena A.D., Linear Algebra: An Interactive Approach, Thompson Brooks/Cole, 2004.

[6] Lang S., Linear Algebra 3rd Ed., Springer-Verlag, New York, 1987.

[7] McCoy N.H., The Theory of Rings, Chelsea Publishing Company, New York, 1973.

[8] Meyer J., Szigeti J. and Van Wyk L., A Cayley-Hamilton trace identity for 2×2 matrices over Lie-solvable rings, Linear Algebra and its Applications 436 (2012), 2578-2582.

[9] Mirzakhani M., A Simple Proof of a Theorem of Schur, The American Mathematical Monthly, vol. 105, No. 3 (Mar., 1998), 260-262.

[10] Nering E.D., Linear Algebra and Matrix Theory 2nd Ed., John Wiley, 1970.

[11] Paria G., Classical Algebra, Arunabha Sen, Calcutta, 1998.

[12] Straubing H., A combinatorial proof of the Cayley-Hamilton Theorem, Discrete Math. 43 (1983), 273-279.

[13] Silva R.R., The Trace Formulas yield the Inverse Metric Formula, J. Math. Phys. 39 (1998), 6206-6213.

[14] Zhang F., Matrix Theory, Springer-Verlag, New York, 1961.

[15] Zhang H., Yan W. and Li X., Trace Formulae of Characteristic Polynomial and Cayley-Hamilton's Theorem, and Applications to Chiral Perturbation Theory and General Relativity, Commun. Theor. Phys. 49 (2008), 801-808.