Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Linear Algebra and its Applications 430 (2009) 1790–1805
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Commutative matrix subalgebras and length function ☆

A.E. Guterman *, O.V. Markova

Department of Algebra, Faculty of Mathematics and Mechanics, Moscow State University, Moscow 119991, GSP-1, Russia

Received 24 February 2008; accepted 4 July 2008
Available online 2 September 2008
Submitted by V. Sergeichuk

Dedicated to Thomas Laffey on the occasion of his 65th birthday.

Abstract

It is proved that if the length of a commutative matrix subalgebra is maximal then this subalgebra is maximal under inclusion. The examples are given showing that the converse does not hold. To establish this result, we prove several fundamental properties of the length function.

AMS classification: Primary: 13E10, 15A27; Secondary: 15A18

Keywords: Finite-dimensional algebras; Matrix length; Commutative matrix algebras

1. Introduction

A ring $\mathscr{A}$, which is a vector space over a field $\mathbb{F}$, is called an *algebra* over $\mathbb{F}$ or $\mathbb{F}$-*algebra*, if for any $\lambda \in \mathbb{F}$ and arbitrary $a, b \in \mathscr{A}$ it is true that $\lambda(ab) = (\lambda a)b = a(\lambda b)$. The algebra is called *finite-dimensional* if the corresponding vector space has a finite dimension over $\mathbb{F}$. The algebra is called *finitely generated* if all its elements can be represented as finite linear combinations with coefficients from the field $\mathbb{F}$ of finite products of some finite set of its elements. This system is called *a system of generators*. It is easy to see, that any finite-dimensional algebra is generated by its basis, i.e., it is finitely generated. Let us denote by $\mathscr{S} = \{a_1, \ldots, a_k\}$ a finite system of generators in algebra $\mathscr{A}$.

☆ This work is partially supported by the grants: RFBR 08-01-00693a, MK-2718.2007.1 and NSh-1983.2008.1.
* Corresponding author.
E-mail address: guterman@list.ru (A.E. Guterman).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2008.07.010

Notation 1.1. Let $\langle S\rangle$ denote the linear span of the set S, i.e., the set of all finite linear combinations of elements from S with coefficients from $\mathbb{F}$.

Definition 1.2. *A length* of the word $a_{i_1}\dots a_{i_t}$, where $a_{i_j}\in\mathscr{S}$, $a_{i_j}\neq 1$, is equal to the number t. If $\mathscr{A}$ is an algebra with 1, then we assume that 1 is the word in elements of the set $\mathscr{S}$ *of the length* 0.

Notation 1.3. Let $\mathscr{S}^i$ denote the set of all words in the alphabet $\{a_1,\dots,a_k\}$ of the length not greater than i, $i\geqslant 0$.

Notation 1.4. Let us denote by $L_i(\mathscr{S})$ a linear span of the words from $\mathscr{S}^i$, $L(\mathscr{S})=\bigcup_{i=0}^{\infty}L_i(\mathscr{S})$ is the linear span of all words in the alphabet $\{a_1,\dots,a_k\}$. Note that $L_0(\mathscr{S})=\mathbb{F}$, if $\mathscr{A}$ is unitary, and $L_0(\mathscr{S})=0$, otherwise.

Remark 1.5. Since $\mathscr{S}$ is the system of generators for the algebra $\mathscr{A}$, then any element of $\mathscr{A}$ can be represented as a finite linear combination of the words in the alphabet $\{a_1,\dots,a_k\}$, i.e. $\mathscr{A}=L(\mathscr{S})$. From the definition of $\mathscr{S}^i$ we have that

$$L_{i+j}(\mathscr{S})=\langle L_i(\mathscr{S})L_j(\mathscr{S})\rangle,$$

and

$$L_0(\mathscr{S})\subseteq L_1(\mathscr{S})\subseteq\cdots\subseteq L_h(\mathscr{S})\subseteq\cdots\subseteq L(\mathscr{S})=\mathscr{A}.$$

Since the algebra $\mathscr{A}$ is finite-dimensional, we have that there exist a number h, such that $L_h(\mathscr{A})=L_{h+1}(\mathscr{A})$.

Definition 1.6. *The length of the generating system* $\mathscr{S}$ is a minimal nonnegative integer k, such that $L_k(S)=L_{k+1}(S)$. Let us denote the length of the generating system $\mathscr{S}$ by $l(\mathscr{S})$.

Remark 1.7. If for some $h\geqslant 0$ we have $L_h(\mathscr{S})=L_{h+1}(\mathscr{S})$ then

$$L_{h+2}(\mathscr{S})=\langle L_1(\mathscr{S})L_{h+1}(\mathscr{S})\rangle=\langle L_1(\mathscr{S})L_h(\mathscr{S})\rangle=L_{h+1}(\mathscr{S})$$

and also $L_i(\mathscr{S})=L_h(\mathscr{S})$ for all $i\geqslant h$. Therefore $l(\mathscr{S})$ is correctly defined. Since $\mathscr{S}$ is a generating system for the algebra $\mathscr{A}$, we have $L_h(\mathscr{S})=L(\mathscr{S})=\mathscr{A}$.

Definition 1.8. *The length of the algebra* $\mathscr{A}$ is defined to be $l(\mathscr{A})=\max_{\mathscr{S}}l(\mathscr{S})$, where maximum is taken over all generating systems for this algebra.

Definition 1.9. The word $v\in L_j(\mathscr{S})$ is called *reducible over* $\mathscr{S}$, if there exist a number $i<j$, such that $v\in L_i(\mathscr{S})$ and $L_i(\mathscr{S})\neq L_j(\mathscr{S})$.

Notation 1.10. Let $M_{m,n}(\mathbb{F})$ denote the vector space of $m\times n$-matrices over a field $\mathbb{F}$. By $M_n(\mathbb{F})$ we denote the algebra of matrices of the order n over $\mathbb{F}$, $T_n(\mathbb{F})$ is the algebra of upper-triangular matrices of the size n over $\mathbb{F}$, $D_n(\mathbb{F})$ denotes the algebra of diagonal matrices of the order n over $\mathbb{F}$, and $N_n(\mathbb{F})$ denotes the subalgebra of nilpotent matrices in $T_n(\mathbb{F})$.

By I we denote an identity matrix and by $E_{i,j}$ we denote the (i,j)th matrix unit, i.e., the matrix with 1 on the (i,j)th position and 0 elsewhere.

The length of the algebra of 3×3-matrices was firstly studied by Spencer and Rivlin [15,16] in connections with the possible applications in mechanics. The following bounds for the length of the matrix algebra are due to Paz and Pappacena.

Theorem 1.11 [12, Theorem 1, Remark 1]. *Let $\mathbb{F}$ be an arbitrary field. Then $l(M_n(\mathbb{F})) \leqslant \lceil (n^2+2)/3 \rceil$.*

Theorem 1.12 [11, Corollary 3.2]. *Let $\mathbb{F}$ be an arbitrary field. Then $l(M_n(\mathbb{F})) < n\sqrt{2n^2/(n-1)+1/4}+n/2-2$.*

It was proved by Paz, see [12], that the upper bound for the length of a commutative matrix subalgebra in $M_n(\mathbb{C})$ is equal to $n-1$, here $\mathbb{C}$ denotes the field of complex numbers.

In Sections 2–5 we give some basic algebraic properties of the length function, which are necessary for the further investigations of the length of commutative matrix subalgebras. In particular, in Section 2 we show, that the length of a generating system is invariant with respect to invertible linear transformations of this system. In Section 3 we prove that, the length of an algebra is not changed under the adjunction of identity element. In Section 4 we obtain upper and lower bounds for the length of a direct sum of several matrix algebras and block-triangular matrix algebras. As a corollary the upper bounds for the lengths of subalgebras of the triangular matrix algebra are found. In Section 5 the length behavior under the ground field extensions is investigated, in particular, we study the case of the algebraic closure of a given field. In Section 6 we generalize the result by Paz on the length of commutative matrix algebras to the case of an arbitrary ground field. In Section 7 we characterize the class of algebras over an algebraically closed field, for which this bound is achieved, namely, it is shown that such algebras are maximal under the inclusion and are generated by a nonderogatory matrix. In Section 8 it is demonstrated that there are maximal commutative subalgebras of nonmaximal length. Also in Section 7 the length of an arbitrary finite-dimensional local algebra is estimated by a linear function in the nilpotency index of its Jacobson radical.

In this work we assume that all algebras under consideration are associative and finite dimensional. Also, since in Section 3 we prove that the length does not change under the adjunction of the identity element, starting from Section 4 we assume that the algebra $\mathscr{A}$ contains the unity element $1_{\mathscr{A}} \in \mathscr{A}$, $1_{\mathscr{A}} \neq 0$.

2. Transforming systems of generators

Proposition 2.1. *Let $\mathbb{F}$ be an arbitrary field and $\mathscr{A}$ be a finite-dimensional algebra over $\mathbb{F}$. If $\mathscr{S} = \{a_1, \ldots, a_k\}$ is a generating system of this algebra $\mathscr{A}$ and $C = \{c_{ij}\} \in M_k(\mathbb{F})$ is a nonsingular matrix, then the set of the coordinates of the vector*

$$C\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} c_{11}a_1 + c_{12}a_2 + \cdots + c_{1k}a_k \\ \vdots \\ c_{k1}a_1 + c_{k2}a_2 + \cdots + c_{kk}a_k \end{pmatrix}, \tag{1}$$

i.e. the set

$$\mathscr{S}_c = \{c_{11}a_1 + c_{12}a_2 + \cdots + c_{1k}a_k, \ldots, c_{k1}a_1 + c_{k2}a_2 + \cdots + c_{kk}a_k\},$$

is a system of generators for the algebra $\mathscr{A}$ and $l(\mathscr{S}_c) = l(\mathscr{S})$.

Proof. Let us prove the equality $L_n(\mathscr{S}) = L_n(\mathscr{S}_c)$ by the induction on n.

The base: $n = 1$. By the definition we have that for any $\gamma_1, \ldots, \gamma_k \in \mathbb{F}$ the inclusion $\gamma_1 a_1 + \cdots + \gamma_k a_k \in L_1(\mathscr{S})$ holds. Therefore, $L_1(\mathscr{S}_c) \subseteq L_1(\mathscr{S})$.

Since the matrix C is nonsingular, there exists the matrix $D := C^{-1}$. We denote the elements of this matrix by d_{ij}, i.e., $D = (d_{ij})$. Then by (1)

$$a_i = (d_{i1}, d_{i2}, \ldots, d_{ik}) \begin{pmatrix} c_{11}a_1 + c_{12}a_2 + \cdots + c_{1k}a_k \\ \vdots \\ c_{k1}a_1 + c_{k2}a_2 + \cdots + c_{kk}a_k \end{pmatrix} \in L_1(\mathscr{S}_c), \text{ for all } i = 1, \ldots, k,$$

according to the definition of linear span. Therefore, $L_1(\mathscr{S}) \subseteq L_1(\mathscr{S}_c)$. Hence $L_1(\mathscr{S}_c) = L_1(\mathscr{S})$.

The step. Let us assume that $n > 1$ and the statement of the proposition holds for all $m < n$. Therefore,

$$L_n(\mathscr{S}) = \langle L_1(\mathscr{S}) L_{n-1}(\mathscr{S}) \rangle = \langle L_1(\mathscr{S}_c) L_{n-1}(\mathscr{S}_c) \rangle = L_n(\mathscr{S}_c). \qquad \square$$

Proposition 2.2. *Let $\mathbb{F}$ be an arbitrary field and $\mathscr{A}$ be a finite-dimensional algebra with unity over $\mathbb{F}$. Assume that $\mathscr{S} = \{a_1, \ldots, a_k\}$ is a system of generators for this algebra such that $1_{\mathscr{A}} \notin \langle a_1, \ldots, a_k \rangle$. Then for any $\gamma_1, \ldots, \gamma_k \in \mathbb{F}$ the set*

$$\mathscr{S}_1 = \{a_1 + \gamma_1 1_{\mathscr{A}}, \ldots, a_k + \gamma_k 1_{\mathscr{A}}\}$$

is a generating system of the algebra $\mathscr{A}$ and $l(\mathscr{S}_1) = l(\mathscr{S})$.

Proof. As in the previous statement, we prove the equality $L_n(\mathscr{S}) = L_n(\mathscr{S}_1)$ for all n by the induction on n.

The base. Since $1_{\mathscr{A}} \in L_0(\mathscr{S}) = L_0(\mathscr{S}_1)$, we have $L_1(\mathscr{S}_1) = L_1(\mathscr{S})$.

The step. Let us assume that $n > 1$ and the statement is true for all $m < n$. Then by the equality

$$L_n(\mathscr{S}) = \langle L_1(\mathscr{S}) L_{n-1}(\mathscr{S}) \rangle = \langle L_1(\mathscr{S}_1) L_{n-1}(\mathscr{S}_1) \rangle = L_n(\mathscr{S}_1), \qquad \square$$

the result follows.

Proposition 2.3. *Let $\mathbb{F}$ be an arbitrary field, $\mathscr{A}$ be a finite-dimensional $\mathbb{F}$-algebra without unity and let $\mathscr{S} = \{a_1, \ldots, a_k\}$ be a system of generators for the algebra $\mathscr{A}$. Then there exists a generating system $\mathscr{S}'$ for $\mathscr{A}$ such that the following conditions hold*:

1. $\mathscr{S}' \subseteq \mathscr{S}$;
2. $\dim L_1(\mathscr{S}') = |\mathscr{S}'|$;
3. $l(\mathscr{S}') = l(\mathscr{S})$.

Proof. By definition for any generating system we have $L_0(\mathscr{S}) = 0$ and $L_1(\mathscr{S}) = \langle \mathscr{S} \rangle$. Consider those elements $a_{i_1}, \ldots, a_{i_m}$, in $\mathscr{S}$ that form a basis of $\langle \mathscr{S} \rangle$. Then we define $\mathscr{S}' = \{a_{i_1}, \ldots, a_{i_m}\}$. Therefore, conditions 1 and 2 hold.

As in the previous statement, we prove the equality $L_n(\mathscr{S}) = L_n(\mathscr{S}')$ for all n by the induction on n.

The base. Since $L_0(\mathscr{S}) = 0 = L_0(\mathscr{S}')$, then $L_1(\mathscr{S}) = \langle \mathscr{S} \rangle = \langle \mathscr{S}' \rangle = L_1(\mathscr{S}')$.

The step. Let us assume that $n > 1$ and the statement is true for all $r < n$. Then by the equality

$$L_n(\mathscr{S}) = \langle L_1(\mathscr{S}) L_{n-1}(\mathscr{S}) \rangle = \langle L_1(\mathscr{S}') L_{n-1}(\mathscr{S}') \rangle = L_n(\mathscr{S}'),$$

and the condition 3 follows. $\square$

Proposition 2.4. *Let* $\mathbb{F}$ *be an arbitrary field,* $\mathscr{A}$ *be a finite-dimensional* $\mathbb{F}$*-algebra with a unity* $1_{\mathscr{A}}$ *and let* $\mathscr{S} = \{a_1, \ldots, a_k\}$ *be a system of generators for the algebra* $\mathscr{A}$*. Then there exists a generating system* $\mathscr{S}'$ *for* $\mathscr{A}$ *such that the following conditions hold*:

1. $\mathscr{S}' \subseteq \mathscr{S}$;
2. $1_{\mathscr{A}} \notin \langle \mathscr{S}' \rangle$
3. $\dim L_1(\mathscr{S}') = |\mathscr{S}'| + 1$;
4. $l(\mathscr{S}') = l(\mathscr{S})$.

Proof. By definition for any generating system we have $L_0(\mathscr{S}) = \langle 1_{\mathscr{A}} \rangle$ and $L_1(\mathscr{S}) = \langle \mathscr{S} \cup \{1_{\mathscr{A}}\} \rangle$. Consider the set $\mathscr{S}_1 = \{1_{\mathscr{A}}, a_1, a_2, \ldots, a_k\}$. Successively removing those elements of $\mathscr{S}_1$ that are linearly dependent with the elements with smaller indices we obtain a set $\mathscr{S}_2 = \{1_{\mathscr{A}}, a_{j_1}, a_{j_2}, \ldots, a_{j_m}\}$. By construction $\mathscr{S}_2$ is a basis of $L_1(\mathscr{S})$. That is, let us define $\mathscr{S}' = \mathscr{S}_2 \setminus \{1_{\mathscr{A}}\} = \{a_{j_1}, a_{j_2}, \ldots, a_{j_m}\}$. Then $1_{\mathscr{A}} \notin \langle \mathscr{S}' \rangle$ and

$$L_1(\mathscr{S}) = \langle \mathscr{S} \cup \{1_{\mathscr{A}}\} \rangle = \langle \mathscr{S}' \cup \{1_{\mathscr{A}}\} \rangle = \langle \mathscr{S}' \rangle + \langle 1_{\mathscr{A}} \rangle = \langle \mathscr{S}' \rangle \oplus \langle 1_{\mathscr{A}} \rangle = L_1(\mathscr{S}'). \quad (2)$$

Hence, conditions 1–3 hold.

As in the previous statement, we prove the equality $L_n(\mathscr{S}) = L_n(\mathscr{S}')$ for all n by the induction on n.

The base follows from Eq. (2).

The step. Let us assume that $n > 1$ and the statement is true for all $r < n$. Then by the equality

$$L_n(\mathscr{S}) = \langle L_1(\mathscr{S}) L_{n-1}(\mathscr{S}) \rangle = \langle L_1(\mathscr{S}') L_{n-1}(\mathscr{S}') \rangle = L_n(\mathscr{S}'),$$

and the condition 4 follows. □

Corollary 2.5. *Let* $\mathbb{F}$ *be an arbitrary field and* $\mathscr{A}$ *be a finite-dimensional* $\mathbb{F}$*-algebra with or without a unity. Then we can always assume that a system of generators of the algebra* $\mathscr{A}$ *does not contain linearly dependent elements.*

3. Length behavior and the unity adjunction

Theorem 3.1. *Let* $\mathbb{F}$ *be an arbitrary field and* $\mathscr{A}$ *be a finite-dimensional algebra without unity over the field* $\mathbb{F}$*. We define the* $\mathbb{F}$*-algebra* $\mathscr{A}_1 = \mathscr{A} \oplus \mathbb{F}$ *with the following operations*:

$$(a_1, f_1) + (a_2, f_2) = (a_1 + a_2, f_1 + f_2),$$
$$f_1(a, f_2) = (f_1 a, f_1 f_2),$$
$$(a_1, f_1)(a_2, f_2) = (a_1 a_2 + f_2 a_1 + f_1 a_2, f_1 f_2),$$
$$a, \ a_1, \ a_2 \in \mathscr{A}, \ f_1, \ f_2 \in \mathbb{F}.$$

Then $\mathscr{A}_1$ *is a finite-dimensional* $\mathbb{F}$*-algebra with the unity* $(0, 1)$ *and* $l(\mathscr{A}) = l(\mathscr{A}_1)$.

Proof. It is straightforward to check that $\mathscr{A}_1$ is an $\mathbb{F}$-algebra. $\mathscr{A}_1$ is finite-dimensional, since $\dim_{\mathbb{F}} \mathscr{A}_1 = \dim_{\mathbb{F}} \mathscr{A} + 1$.

The element (0, 1) is unity in the algebra $\mathscr{A}_1$, since for any $a \in \mathscr{A},\ f \in \mathbb{F}$ we have

$$(a, f)(0, 1) = (a0 + f0 + 1a, f1) = (a, f),$$
$$(0, 1)(a, f) = (0a + 1a + f0, 1f) = (a, f).$$

Let us prove now the equality $l(\mathscr{A}) = l(\mathscr{A}_1)$.

1. At first we show that $l(\mathscr{A}_1) \leqslant l(\mathscr{A})$.

To do this we take a system of generators $\mathscr{S}_1 = \{b_1, \dots, b_k\}$ in $\mathscr{A}_1$ such that $l(\mathscr{S}_1) = l(\mathscr{A}_1)$. By the definition of $\mathscr{A}_1$ the elements $b_i = (a_i, f_i)$, where $a_i \in \mathscr{A}$ and $f_i \in \mathbb{F}$, $i = 1, \dots, k$. Without loss of generality by Proposition 2.4 we can assume that $1_{\mathscr{A}_1} \notin \langle \mathscr{S}_1 \rangle$ and $\dim L_1(\mathscr{S}_1) = |\mathscr{S}_1| + 1 = k + 1$.

Then by the proposition 2.2 the set $\mathscr{S}_0 = \{(a_1, 0), \dots, (a_k, 0)\}$ is a system of generators for $\mathscr{A}_1$, and for this set it holds that

$$l(\mathscr{S}_1) = l(\mathscr{S}_0) \tag{3}$$

and $\dim L_1(\mathscr{S}_0) = \dim L_1(\mathscr{S}_1) = k + 1$. It follows that all the elements $a_1, \dots, a_k$ are linearly independent in $\mathscr{A}$. Let us denote $\mathscr{S} = \{a_1, \dots, a_k\}$. We have that for any $m \geqslant 1$

$$L_m(\mathscr{S}) = \langle \mathscr{S}^m \rangle \cong \langle \{(v, 0), v \in \mathscr{S}^m\} \rangle = \langle \mathscr{S}_0^m \backslash \mathscr{S}_0^0 \rangle. \tag{4}$$

By the definition of $\mathscr{S}_0$ we have

$$L_m(\mathscr{S}_0) = \langle \mathscr{S}_0^m \backslash \mathscr{S}_0^0 \rangle \oplus \langle \mathscr{S}_0^0 \rangle. \tag{5}$$

Since $\langle \mathscr{S}_0^0 \rangle \cong \mathbb{F}$, it follows that for any $m \geqslant 1$

$$\dim L_m(\mathscr{S}_0) = \dim L_m(\mathscr{S}) + 1. \tag{6}$$

By Eq. (3) we have that for $m = l(\mathscr{S}_1)$ it holds that $\dim L_m(\mathscr{S}_0) = \dim \mathscr{A}_1$. Then for $m = l(\mathscr{S}_1)$ we have that

$$\dim L_m(\mathscr{S}) = \dim L_m(\mathscr{S}_0) - 1 = \dim \mathscr{A}_1 - 1 = \dim \mathscr{A}.$$

Therefore the set $\mathscr{S}$ is a generating system for $\mathscr{A}$ and

$$l(\mathscr{A}_1) = l(\mathscr{S}_0) = l(\mathscr{S}) \leqslant \max_{\mathscr{S}} l(\mathscr{S}) = l(\mathscr{A}).$$

2. Let us prove the opposite inequality, $l(\mathscr{A}) \leqslant l(\mathscr{A}_1)$. To do this we take a generating system $\mathscr{S} = \{a_1, \dots, a_k\}$ in $\mathscr{A}$ with the length $l(\mathscr{S}) = l(\mathscr{A})$. Proposition 2.3 allows us to assume that $\dim L_1(\mathscr{S}) = |\mathscr{S}| = k$. Let us consider the subset $\mathscr{S}_0 = \{(a_1, 0), \dots, (a_k, 0)\}$ in $\mathscr{A}_1$. By its construction, $1_{\mathscr{A}_1} = (0, 1) \notin \langle \mathscr{S}_0 \rangle$ and $\dim L_1(\mathscr{S}_0) = k + 1$. Moreover, for any $m \geqslant 1$ the generating systems $\mathscr{S}$ and $\mathscr{S}_0$ satisfy the conditions (4)–(6). Hence for $m = l(\mathscr{S})$ we have

$$\dim L_m(\mathscr{S}_0) = \dim L_m(\mathscr{S}) + 1 = \dim \mathscr{A} + 1 = \dim \mathscr{A}_1.$$

Therefore the set $\mathscr{S}_0$ is a generating system for $\mathscr{A}_1$ and

$$l(\mathscr{A}) = l(\mathscr{S}) = l(\mathscr{S}_0) \leqslant \max_{\mathscr{S}_{\mathscr{A}_1}} l(\mathscr{S}_{\mathscr{A}_1}) = l(\mathscr{A}_1). \qquad \square$$

Corollary 3.2. *Let $\mathbb{F}$ be an arbitrary field and $\mathscr{A}$ be a subalgebra in $M_n(\mathbb{F})$, such that $I \notin \mathscr{A}$. If $\mathscr{A}_1 = \langle \mathscr{A}, I \rangle$, then $l(\mathscr{A}_1) = l(\mathscr{A})$.*

Therefore, without loss of generality we can further assume that all algebras which we consider contain the unit element.

4. The length of the direct sum of algebras

Let $\mathscr{A}$ and $\mathscr{B}$ be finite-dimensional algebras over a field $\mathbb{F}$. By $\mathscr{A} \oplus \mathscr{B}$ we denote the algebra of pairs $(a, b),\ a \in \mathscr{A},\ b \in \mathscr{B}$ with the addition, multiplication and multiplication by scalars defined coordinate-wise in the following way: $\forall (a_1, b_1),\ (a_2, b_2),\ (a, b) \in \mathscr{A} \oplus \mathscr{B}$ and $\forall \alpha \in \mathbb{F}$ we have

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2),$$
$$\alpha(a, b) = (\alpha a, \alpha b),$$
$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2).$$

The following bounds for the direct sum of algebras are proved in [9, Theorem 2]. We plan to apply them several times to prove our results and therefore we include the proof here for completeness.

Theorem 4.1 [9, Theorem 2]. *Let $\mathscr{A}$ and $\mathscr{B}$ be two finite-dimensional algebras over a field $\mathbb{F}$ with the lengths $l_{\mathscr{A}}$ and $l_{\mathscr{B}}$, correspondingly. Then the following inequalities are true*:

$$\max\{l_{\mathscr{A}}, l_{\mathscr{B}}\} \leqslant l(\mathscr{A} \oplus \mathscr{B}) \leqslant l_{\mathscr{A}} + l_{\mathscr{B}} + 1. \tag{7}$$

Proof. Let us denote $p = l_{\mathscr{A}},\ q = l_{\mathscr{B}}$. To prove the lower bound we consider two generating systems $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_m\}$ for algebras $\mathscr{A}$ and $\mathscr{B}$, correspondingly, with the lengths p and q, correspondingly. Then the set $\{(a_1, 0), \ldots, (a_k, 0), (0, b_1), \ldots, (0, b_m)\}$ will be a generating system in $\mathscr{A} \oplus \mathscr{B}$ of the length $\max\{p, q\}$.

Let us consider an algebra $\mathscr{A} \oplus \mathscr{B}$ and take the arbitrary generating system

$$\mathscr{S} = \{(c_1, d_1), \ldots, (c_n, d_n)\}.$$

We are going to prove that any word in elements from $\mathscr{S}$ of the length $p + q + 2$ is reducible. We denote $N = p + q + 2$. Assume that $v = (c_{i_1}, d_{i_1}) \ldots (c_{i_N}, d_{i_N}) = (c_{i_1} \ldots c_{i_N}, d_{i_1} \ldots d_{i_N})$, $i_j \in \{1, \ldots, n\},\ j = 1, \ldots, N$. Since the length of $\mathscr{A}$ is equal to p, the word $c_{i_1} \ldots c_{i_{p+1}}$ is reducible, i.e., $c_{i_1} \ldots c_{i_{p+1}} = \alpha_1 c_{i_1} \ldots c_{i_p} + \cdots + \alpha_{M-1} c_n + \alpha_M 1_{\mathscr{A}}$. Since the length of $\mathscr{B}$ is equal to q, the word $d_{i_{p+2}} \ldots d_{i_N}$ is reducible, i.e., $d_{i_{p+2}} \ldots d_{i_N} = \beta_1 d_{i_{p+2}} \ldots d_{i_{N-1}} + \cdots + \beta_{K-1} d_n + \beta_K 1_{\mathscr{B}}$. If we substitute the representations for $c_{i_1} \ldots c_{i_{p+1}}$ and $d_{i_{p+2}} \ldots d_{i_N}$ into v, then we obtain the equality

$$\begin{aligned}&\big\{(c_{i_1} \ldots c_{i_{p+1}}, d_{i_1} \ldots d_{i_{p+1}}) - \alpha_1 (c_{i_1} \ldots c_{i_p}, d_{i_1} \ldots d_{i_p}) - \cdots \\ &\quad - \alpha_{M-1}(c_n, d_n) - \alpha_M (1_{\mathscr{A}}, 1_{\mathscr{B}})\big\}\big\{(c_{i_{p+2}} \ldots c_{i_N}, d_{i_{p+2}} \ldots d_{i_N}) - \beta_K (1_{\mathscr{A}}, 1_{\mathscr{B}}) \\ &\quad - \beta_{K-1}(c_n, d_n) - \cdots - \beta_1 (c_{i_{p+2}} \ldots c_{i_{N-1}}, d_{i_{p+2}} \ldots d_{i_{N-1}})\big\} = (0, x)(y, 0) = 0.\end{aligned}$$

Therefore the word v can be represented as a linear combination of the words of smaller length. Since v is chosen arbitrary, we get $l(\mathscr{A} \oplus \mathscr{B}) \leqslant p + q + 1$. $\square$

There are examples showing that both bounds are sharp, see [9, Examples 4, 5].

Corollary 4.2 [9, Corollary 3]. *Let $\mathscr{A}$ be a subalgebra in the algebra of block-triangular matrices, i.e., all matrices in $\mathscr{A}$ have the form*

$$A = \left(\begin{array}{c|c|c|c} A_{1,1} & A_{1,2} & \dots & A_{1,k} \\ \hline 0 & A_{2,2} & \dots & A_{2,k} \\ \hline & & \dots\dots & \\ \hline 0 & 0 & \dots & A_{k,k} \end{array}\right),$$

where $A_{i,j} \in M_{n_i,n_j}(\mathbb{F}),\ n_1 + n_2 + \cdots + n_k = n,$ *and all matrices* $A_{i,i}$ *form subalgebras in* $M_{n_i,n_i}(\mathbb{F})$ *of lengths* l_i, $i = 1, \dots, k$. *Then for* $l(\mathscr{A})$ *we have the following inequalities*:

$$\max\{l_1, \dots, l_k\} \leqslant l(\mathscr{A}) \leqslant \sum_{j=1}^{k} l_j + k - 1.$$

As a direct consequence we have the following result:

Corollary 4.3. *Let* $\mathbb{F}$ *be an arbitrary field and* $\mathscr{A}$ *be an arbitrary subalgebra in* $T_n(\mathbb{F})$. *Then* $l(\mathscr{A}) \leqslant n - 1$.

Proof. Any subalgebra of the triangular matrix algebra is a block-triangular algebra, where all blocks are 1 by 1. Since $l(\mathbb{F}) = 0$, we have an estimate $l(\mathscr{A}) \leqslant 0 + n - 1 = n - 1$. □

In [9, theorem 1] it was proved that if $\mathscr{A} = T_n(\mathbb{F})$, then this bound is achieved.

5. The length function and field extensions

Consider a field $\mathbb{F}$ which is not algebraically closed. Let $\mathscr{A}_{\mathbb{F}}$ be a finite-dimensional unitary $\mathbb{F}$-algebra. Let us define the algebra $\mathscr{A}_{\overline{\mathbb{F}}}$ in the following way: $\mathscr{A}_{\overline{\mathbb{F}}} = (\mathscr{A}_{\mathbb{F}} \otimes_{\mathbb{F}} \overline{\mathbb{F}})_{\overline{\mathbb{F}}}$. That is, the following rule defines multiplication by elements of $\overline{\mathbb{F}}$ in $\mathbb{F}$-algebra $\mathscr{B}_{\mathbb{F}} = A_{\mathbb{F}} \otimes_{\mathbb{F}} \overline{\mathbb{F}}$: for $a \in \mathscr{B}_{\mathbb{F}},\ f \in \overline{\mathbb{F}}$ the product is defined $f \cdot a = (1 \otimes f) \cdot a$. Here $\otimes$ denotes the tensor product, see [13, §9] for the details.

Proposition 5.1. $l(\mathscr{A}_{\mathbb{F}}) \leqslant l(\mathscr{A}_{\overline{\mathbb{F}}})$.

Proof. Denote $l = l(\mathscr{A}_{\mathbb{F}})$. By the definition of the length of an algebra there exists a generating system for $\mathscr{A}_{\mathbb{F}}$ of the length l, denote it $\mathscr{S} = \{a_1, \dots, a_k\}$. We then prove that $\mathscr{S}' = \{a_1 \otimes 1_{\overline{\mathbb{F}}}, \dots, a_k \otimes 1_{\overline{\mathbb{F}}}\}$ is a generating system for $\mathscr{A}_{\overline{\mathbb{F}}}$ of the length greater than or equal to l.

The set $\mathscr{S}'$ is a generating system for the algebra $\mathscr{A}_{\overline{\mathbb{F}}}$ by the construction of the coefficient field extension of an algebra. Let us prove, that $l(\mathscr{S}) \leqslant l(\mathscr{S}')$. Since $l(\mathscr{S}) = l$, then there exists a word $v = v(a_1, \dots, a_k)$ of length l, which is irreducible in $L_l(\mathscr{S})$. Suppose, that $v \otimes 1_{\overline{\mathbb{F}}}$ is reducible in $L_l(\mathscr{S}')$. Then there exist words $v'_1, \dots, v'_m \in L_{l-1}(\mathscr{S}')$ and elements $\alpha_1, \dots, \alpha_m \in \overline{\mathbb{F}}$, such that $v \otimes 1_{\bar{\mathbb{F}}} = \sum_{i=1}^{m} \alpha_i v'_i$. However, by the definition of the generating system $\mathscr{S}'$ any word

$$v'_i = (a_{i_1} \otimes 1_{\overline{\mathbb{F}}}) \cdots (a_{i_j} \otimes 1_{\overline{\mathbb{F}}}) = (a_{i_1} \cdots a_{i_j}) \otimes 1_{\overline{\mathbb{F}}},$$

i.e. any word v'_i is of the following form $v'_i = v_i \otimes 1_{\overline{\mathbb{F}}}$ for some $v_i \in L_{l-1}(\mathscr{S})$. Then $v \otimes 1_{\overline{\mathbb{F}}} = \sum_{i=1}^{m} v_i \otimes \alpha_i$. It follows that $\alpha_i \in \mathbb{F}$ and $v = \sum_{i=1}^{m} \alpha_i v_i$. Consequently, $\alpha_i = 0$, $i = 1, \dots, m$, since the word v was taken to be irreducible in $L_l(\mathscr{S})$. Therefore v' is irreducible in $L_l(\mathscr{S}')$. Hence, $l = l(\mathscr{S}) \leqslant l(\mathscr{S}')$, which means that $l(\mathscr{A}_{\overline{\mathbb{F}}}) \geqslant l(\mathscr{A})$. □

In order to show that this inequality may be sharp let us compute the length of the diagonal matrix algebra over different fields.

For further discussions we need the following class of matrices:

Definition 5.2. Let $\mathbb{F}$ be an arbitrary field. A matrix $C \in M_n(\mathbb{F})$ is called *nonderogatory* provided that

$$\dim_{\mathbb{F}}(\langle I, C, C^2, \ldots, C^{n-1}\rangle) = n.$$

Remark 5.3. The following easy observation is pointed out in [5, Section 4.4.17]: Let $\mathbb{F}$ be an arbitrary field. A matrix $C \in M_n(\mathbb{F})$ is nonderogatory if and only if its minimal polynomial coincides with its characteristic polynomial.

Theorem 5.4. *Let $\mathbb{F}$ be an arbitrary field. If $|\mathbb{F}| \geqslant n$, then $l(D_n(\mathbb{F})) = n - 1$.*

Proof. It follows from Corollary 4.3 that for any field $\mathbb{F}$ it holds that $l(D_n(\mathbb{F})) \leqslant n - 1$. Let us construct a generating system for the algebra $D_n(\mathbb{F})$ providing the lower bound. The condition of the theorem provides us with a matrix $A \in D_n(\mathbb{F})$ such that its diagonal elements are pairwise distinct. Thus the minimal polynomial of the matrix A coincides with its characteristic polynomial, i.e. A is a nonderogatory matrix. From the definition of a nonderogatory matrix we derive the linear independence of $I, A, A^2, \ldots, A^{n-1}$ and consequently these powers form a basis of $D_n(\mathbb{F})$. That is, if we consider a generating system $\mathscr{S} = \{A\}$, then $l(\mathscr{S}) = n - 1$. $\square$

Example 5.5. Let $\mathscr{A}_{\mathbb{F}} = D_n(\mathbb{F}_2),\ n > 2$. It is shown in [10, Theorem 6.1] that $l(\mathscr{A}) = [\log_2(n)]$, but $|\overline{\mathbb{F}}_2| = \infty$, and hence by Theorem 5.4 $l(\mathscr{A}_{\overline{\mathbb{F}}}) = n - 1$. However $[\log_2(n)] < n - 1$ for $n > 2$. Thus the length of an algebra can increase while passing from a field to its closure.

6. The upper bound for the length of commutative subalgebras in $M_n(\mathbb{F})$

In his paper [12] Paz obtained, that the length of any commutative subalgebra in $M_n(\mathbb{C})$ is less or equal to $n - 1$. In this section we generalize this bound to the case of an arbitrary field.

Theorem 6.1. *Let $\mathbb{F}$ be an arbitrary field and $\mathscr{A}$ be a commutative subalgebra in $M_n(\mathbb{F})$. Then $l(\mathscr{A}) \leqslant n - 1$.*

Proof. It can be checked directly that the proof of [12, Theorem 2] uses only the fact that the ground field is algebraically closed, but not any other properties of the field of complex numbers. Consequently, the theorem holds true for an arbitrary algebraically closed field. Hence, for any given field $\mathbb{F}$ it follows from Proposition 5.1 that $l(\mathscr{A}) \leqslant l(\mathscr{A}_{\overline{\mathbb{F}}}) \leqslant n - 1$. $\square$

Lemma 6.3 provides the sharpness of the bound in Theorem 6.1 and is based on the following result of Gerstenhaber.

Theorem 6.2 [18, Theorem 1, 3]. *Let $\mathbb{F}$ be an arbitrary field, and assume that $A,\ B \in M_n(\mathbb{F})$ are such matrices that $AB = BA$. Let $\mathscr{A}(A, B)$ denote the subalgebra in $M_n(\mathbb{F})$ generated by A and B. Then $\dim_{\mathbb{F}} \mathscr{A}(A, B) \leqslant n$.*

Lemma 6.3. *Let $\mathbb{F}$ be an arbitrary field and let $\mathscr{A}$ be a commutative subalgebra in $M_n(\mathbb{F})$. If there exists a nonderogatory matrix $A \in \mathscr{A}$ then $\mathscr{A}$ is a subalgebra generated by A, and $l(\mathscr{A}) = n - 1$.*

Proof. Let us first prove that A generates $\mathscr{A}$. Since matrices $I, A, A^2, \dots, A^{n-1}$ are linearly independent (by the definition of a nonderogatory matrix), then $\dim_{\mathbb{F}} \mathscr{A} \geqslant n$. Let us consider a matrix $B \in \mathscr{A}$. Then $AB = BA$. Let $L(A)$ denote the subalgebra of $M_n(\mathbb{F})$, generated by the matrix A and $L(A, B)$ denote a subalgebra of $M_n(\mathbb{F})$, generated by the matrices $A,\ B$. Evidently $L(A) \subseteq L(A, B) \subseteq \mathscr{A}$. From Gerstenhaber's Theorem (Theorem 6.2 of this paper) we obtain $\dim_{\mathbb{F}} L(A, B) \leqslant n$. Consequently, $L(A, B) = L(A)$. Therefore, $B \in \langle I, A, A^2, \dots, A^{n-1}\rangle$, i.e. B is a value of a polynomial in A, $\mathscr{A} = L(A)$ and $\mathscr{A}$ is a maximal commutative subalgebra. It also follows from the linear independence of matrices $I, A, A^2, \dots, A^{n-1}$ that $l(\mathscr{A}) = n - 1$. □

7. Commutative matrix subalgebras of the maximal length

The investigation of commutative subalgebras in matrix algebra is a classical subject of research and dates back to the work of Schur, [14]. Further results can be found in [14,1–3,6,7,17,18]. In this section we show, that the length of a commutative subalgebra is equal to the maximal possible its value, $n - 1$, if and only if this subalgebra is generated by a nonderogatory matrix. As a consequence in Corollary 7.10 we prove that commutative subalgebras of the maximal length in $M_n(\mathbb{F})$ are maximal under inclusion. Note that not all maximal commutative subalgebras are generated by nonderogatory matrices, i.e., are of the maximal length, consider for example, Schur's algebra (see Example 8.1).

Remark 7.1. Assume that C is a nonderogatory matrix. Then it can be checked directly that $C + \alpha I$ is also a nonderogatory matrix for any $\alpha \in \mathbb{F}$. Therefore, an algebra $\mathscr{A}$ without a unit is generated by a nonderogatory matrix if and only if the unitary algebra $\mathscr{A}_1 = \langle I,\ A \mid A \in \mathscr{A}\rangle$ is generated by a nonderogatory matrix. Thus the fact that an algebra is generated by a nonderogatory matrix does not depend on the existence of a unit element in this algebra, and without the loss of generality we can assume that $I \in \mathscr{A}$.

Lemma 7.2. *Let $\mathbb{F}$ be an arbitrary field and let $\mathscr{A} \subseteq M_n(\mathbb{F})$ be a commutative block-diagonal matrix subalgebra with blocks $\mathscr{A}_i \subseteq M_{n_i}(\mathbb{F}),\ i = 1, \dots, k$. If $l(\mathscr{A}) = n - 1$, then $l(\mathscr{A}_i) = n_i - 1$ for any $i = 1, \dots, k$.*

Proof. The commutativity of $\mathscr{A}$ implies that all $\mathscr{A}_i$ are also commutative. Hence for any $i = 1, \dots, k$ it holds that $l(\mathscr{A}_i) \leqslant n_i - 1$. Suppose there exist such a number $i_0 \in \{1, \dots, k\}$ that $l(\mathscr{A}_{i_0}) < n_{i_0} - 1$. Then it follows from Corollary 4.2 that

$$l(\mathscr{A}) \leqslant \sum_{j=1}^{k} l(\mathscr{A}_j) + k - 1 < \sum_{j=1}^{k} (n_j - 1) + k - 1 = n - 1.$$

This contradiction proves the lemma. □

Proposition 7.3. *Let $\mathbb{F}$ be an arbitrary field. A commutative subalgebra $\mathscr{A}$ in $M_2(\mathbb{F})$ is of the maximal length if and only if it is generated by a nonderogatory matrix.*

Proof. Let $A \in \mathscr{A}$ be an arbitrary matrix. If the matrix A is not nonderogatory (is derogatory), i.e. $\dim(\langle I, A\rangle) = 1$, then $A = aI,\ a \in \mathbb{F}$. That is, if there exists $A \in \mathscr{A},\ A \neq aI$, then $\mathscr{A}$ is generated by a nonderogatory matrix A. In this case $l(\mathscr{A}) = 1$. And subalgebra generated by the identity matrix is of the length 0. □

Proposition 7.4. *Let $\mathbb{F}$ be a field, containing at least n distinct nonzero elements, let V be a subspace in $\mathbb{F}^n$ and for any $i = 1, \ldots, n$ there exists such $v^i \in V$ that $v^i_i \neq 0$. Then there exists such $v \in V$ that $v_i \neq 0$ for all $i = 1, \ldots, n$.*

Proof. We use the induction on n.

The base. For $n = 1$ the proposition holds.

The step. Assume, that $n > 1$ and for subspaces of $\mathbb{F}^{n-1}$ the proposition is valid. It follows that there exists such $v^0 \in V$ that $v^0_i \neq 0,\ i = 1, \ldots, n-1$. If $v^0_n \neq 0$, then we assign $v = v^0$. Otherwise there exists $v^1 \in V,\ v^1_n \neq 0$. Consider $v^0 + av^1,\ a \in \mathbb{F},\ a \neq 0$. Let $y_i(x) = v^1_i x + v^0_i,\ i = 1, \ldots, n-1$. For every i it holds by the induction hypothesis that $v^0_i \neq 0$. Hence, the linear equation $y_i(x) = 0$ has at most one solution in $\mathbb{F}$ for any given i, $i = 1, \ldots, n-1$. Consequently, if $X = \{x \in \mathbb{F} \mid \exists i \in \{1, \ldots, n-1\} : y_i(x) = 0\}$, then $|X| \leqslant n-1$. It follows from the condition on the field $\mathbb{F}$, that there exists $a \in \mathbb{F},\ a \neq 0$ such that $v^0_i + av^1_i \neq 0,\ i = 1, \ldots, n-1$. Now, $v^0_n + av^1_n = av^1_n \neq 0$ by the choice of $a \neq 0$ and by the choice of v^1. Therefore, the vector $v^0 + av^1$ satisfies the required condition. □

The following two notions are actual for our further considerations, see for example [13, §4] for the more details.

Definition 7.5. An associative algebra is called *local*, if it has a unique maximal right ideal.

Definition 7.6. *The Jacobson's radical* of an associative ring is the intersection of all its maximal right ideals.

Remark 7.7. It is well known, that the set of all noninvertible elements of a local algebra coincides with its Jacobson's radical; the Jacobson's radical J of an Artinian ring is nilpotent, i.e. there exists such a number N, that $J^N = (0)$, but $J^{N-1} \neq (0)$, here J^N denotes the set of products of elements from J of length N, the number N is called *the index of nilpotency* of the ideal J, see for example [13, §4.4]. In particular, the Jacobson's radical J of a finite-dimensional algebra is nilpotent.

In order to deal with Theorem 7.9 the following lemma is important. Also this lemma will be helpful to deal with Examples 8.1–8.3 in Section 8.

Lemma 7.8. *Let $\mathbb{F}$ be an arbitrary field and let $\mathscr{A}$ be a finite-dimensional local $\mathbb{F}$-algebra. Let $J(\mathscr{A})$ denote the Jacobson's radical of $\mathscr{A}$, and N denote the index of nilpotency of $J(\mathscr{A})$. Then $l(\mathscr{A}) \leqslant N - 1$.*

Proof. Let us show that any word of length N is reducible in $\mathscr{A}$, which means that the length of any generating system is less than N.

Consider N elements $a_1, \ldots, a_N$ in the algebra $\mathscr{A}$. Up to the permutation of indices we would assume that the elements $a_1, \ldots, a_k$ are invertible, and elements $a_{k+1}, \ldots, a_N \in J(\mathscr{A})$. Let $1_{\mathscr{A}}$ denote the unit of the algebra $\mathscr{A}$. By the main theorem on equivalence of different definitions of local rings, see [13, §5.2], there exist such $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$, that $a_i - \alpha_i 1_{\mathscr{A}} \in J(A)$, $i = 1, \ldots, k$. Then

$$(a_1 - \alpha_1 1_{\mathscr{A}}) \cdots (a_k - \alpha_l 1_{\mathscr{A}}) a_{k+1} \cdots a_N = 0,$$

by the definition of the index of nilpotency of an ideal. Expanding this expression and carrying all summands, but first, to the right side of the equality, we obtain that either $a_1 \cdots a_N = 0$, or

$$a_1 \cdots a_N = \sum_{i=1}^{k} \alpha_i a_1 \cdots a_{i-1} a_{i+1} \cdots a_N + \text{summands of smaller length},$$

i.e. any word of length N is reducible in terms of Definition 1.9. Consequently, $l(\mathscr{A}) \leqslant N - 1$. □

Theorem 7.9. *Let $\mathbb{F}$ be an algebraically closed field. A commutative subalgebra $\mathscr{A}$ in $M_n(\mathbb{F})$ is of the maximal length if and only if it is generated by a nonderogatory matrix.*

Proof. Sufficiency follows from Lemma 6.3.

Assume $l(\mathscr{A}) = n - 1$. Let us show that in this case $\mathscr{A}$ is generated by a nonderogatory matrix. We use the induction on n to prove the theorem.

The base. For $n = 1$ any unitary subalgebra $\mathscr{A}$ of $M_1(\mathbb{F})$ coincides with $M_1(\mathbb{F}) = \mathbb{F}$ and is generated by the nonderogatory element 1. For $n = 2$ the statement follows from Proposition 7.3.

The step. Assume that for commutative subalgebras in $M_k(\mathbb{F}),\ k < n$, the theorem is valid. Our proof is divided into two cases:

1. Assume there exists a matrix A in $\mathscr{A}$, which has at least two different eigenvalues. Let λ be an eigenvalue of A of multiplicity s. Hence $0 < s < n$ by the assumptions of this case. We denote the minimal polynomial of A by $\mu_A(t) = (t - \lambda)^m g(t)$, where $g(\lambda) \neq 0,\ m \leqslant s < n$. Then it follows from the Theorem on Jordan normal form that there exists such an invertible matrix $V \in M_n(\mathbb{F})$ that $A_0 = V^{-1}AV = \begin{pmatrix} A_s & 0 \\ 0 & A_{n-s} \end{pmatrix}$, where $A_s \in M_s(\mathbb{F})$ is a Jordan matrix, corresponding to λ, and $A_{n-s} \in M_{n-s}(\mathbb{F})$ is a Jordan matrix, corresponding to other eigenvalues of A. By the Theorem on the general form of matrix commuting with a given matrix, see [8, §16.6], we obtain that all matrices commuting with A_0 are also block-diagonal and consist of two blocks.

We introduce the algebra $\mathscr{A}_V = V^{-1}\mathscr{A}V = \{V^{-1}AV | A \in \mathscr{A}\}$. It is straightforward that $l(\mathscr{A}) = l(\mathscr{A}_V)$, and $\mathscr{A}$ is generated by a nonderogatory matrix if and only if $\mathscr{A}_V$ is generated by a nonderogatory matrix. Let us show that the algebra $\mathscr{A}_V$ is a block-diagonal algebra, i.e., a direct sum of certain $s \times s$ and $(n - s) \times (n - s)$ matrix subalgebras.

Let us consider the matrices $B_j = (A_0 - \lambda)^j g(A_0) \in \mathscr{A}_V$. Hence there exist scalars $\beta_j \in \mathbb{F}$ such that $E_s = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix} = \sum_{j=0}^{m} \beta_j B_j$. Hence, $E_s \in \mathscr{A}_V$. Therefore, the matrix $E_{n-s} = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-s} \end{pmatrix} = I - E_s \in \mathscr{A}_V$. It follows that $\mathscr{A}_V = E_s\mathscr{A}_V \oplus E_{n-s}\mathscr{A}_V$.

From Lemma 7.2 we obtain that blocks of $\mathscr{A}_V$ are of lengths $s - 1$ and $n - s - 1$, correspondingly. Then by the induction hypothesis the blocks are generated by nonderogatory matrices. Hence, there are matrices $C_1 = \begin{pmatrix} C_s & 0 \\ 0 & 0 \end{pmatrix} \in \mathscr{A}_V$ and $C_2 = \begin{pmatrix} 0 & 0 \\ 0 & C_{n-s} \end{pmatrix} \in \mathscr{A}_V$, where $C_s \in M_s(\mathbb{F}),\ C_{n-s} \in M_{n-s}(\mathbb{F})$ are nonderogatory matrices generating the corresponding subalgebras.

Let $C(\alpha) = C_1 + C_2 + \alpha E_{n-s}$. Let us show that there exists $\alpha \in \mathbb{F}$ such that the eigenvalues of the matrices C_1 and $C_2 + \alpha E_{n-s}$ are distinct, consequently, the matrix $C_1 + C_2 + \alpha E_{n-s}$ is nonderogatory. Denote by $X \subseteq \mathbb{F}$ the set of all solutions x of one of the equations $x + \gamma_i^1 = \gamma_j^2,\ i = 1, \ldots, s,\ j = 1, \ldots, n - s$, where $\gamma_1^1, \ldots, \gamma_s^1$ are the eigenvalues of the matrix C_1 and $\gamma_1^2, \ldots, \gamma_{n-s}^2$ are the eigenvalues of the matrix C_2. Then $|X| \leqslant s(n - s) \leqslant \left[\frac{n^2}{4}\right]$. Since $\mathbb{F}$ is infinite, then there exists $\alpha_0 \in \mathbb{F},\ \alpha_0 \notin X$. In this case the set of the eigenvalues of the matrix C_1 does not intersect the set of the eigenvalues of the matrix $C_2 + \alpha_0 E_{n-s}$. Hence, $\chi_{C(\alpha_0)}(t) = \chi_{C_1}(t)\chi_{(C_2+\alpha_0 E_{n-s})}(t) = \mu_{C_1}(t)\mu_{(C_2+\alpha_0 E_{n-s})}(t) = \mu_{C(\alpha_0)}(t)$ and the matrix

$C(\alpha_0) \in \mathscr{A}_V$ is nonderogatory. Therefore, the algebra $\mathscr{A}$ is generated by a nonderogatory matrix $VC(\alpha_0)V^{-1}$.

2. Assume any matrix in $\mathscr{A}$ has a unique eigenvalue of multiplicity n. Since any finite set of commuting matrices over an algebraically closed field has a common eigenvector (see [4, Chapter 1, §1.3, Lemma 1.3.17]), then there exist such a nonsingular matrix $V \in M_n(\mathbb{F})$ that $\mathscr{A}_V = V^{-1}\mathscr{A}V \subseteq T_n(\mathbb{F})$. Note that any matrix $C \in \mathscr{A}_V$ has the form $C = \gamma I + C_N,\ C_N \in N_n(\mathbb{F})$. Hence $\mathscr{A}_V \bigcap D_n(\mathbb{F}) = \langle I \rangle,\ \mathscr{A}_V \bigcap N_n(\mathbb{F}) = J$, where J denotes Jacobson's radical of $\mathscr{A}_V$.

If for any $k = 1, \ldots, n$ there exists a matrix $C_k = \{c_{ij}^{(k)}\}$ in J with $c_{k,k+1}^{(k)} \neq 0$, then it follows from Proposition 7.4 that there exists a linear combination of matrices C_k – a matrix $C = \{c_{ij}\}$—such that all $c_{k,k+1} \neq 0$. The matrices $I,\ C, \ldots, C^{n-1}$ are linearly independent. Indeed, let $m < n$. We denote $C^m = \{c_{ij}^{(m)}\}$. In these notations it holds that $c_{ij}^{(m)} = 0$ if $j < i + m$ and $c_{i,i+m}^{(m)} \neq 0$ since $c_{i,i+m}^{(m)}$ are equal to products of some of $c_{i,i+1}^{(1)} = c_{i,i+1} \neq 0$. Thus no one of $I,\ C, \ldots, C^{n-1}$ is a linear combination of the others, i.e., they are linearly independent. Therefore the matrix C is nonderogatory, and both $\mathscr{A}_V$ and $\mathscr{A}$ are generated by nonderogatory matrices.

Suppose there exists $l \in \{1, \ldots, n-1\}$ such that for any $B = \{b_{ij}\} \in J$ it holds that $b_{l,l+1} = 0$. Let $B_k = \{b_{ij}^{(k)}\} \in J,\ k = 1, \ldots, n-1$. The matrix $B_1 B_2 \ldots B_{n-1}$ can possess with a nonzero entry only on position $(1, n)$. This entry is equal to $b_{12}^{(1)} b_{23}^{(2)} \ldots b_{n-1,n}^{(n-1)}$. By our assumption one of the factors is always equal to zero. It follows that the product of any $n-1$ matrices in J is equal to 0. Then by Lemma 7.8 $l(\mathscr{A}) = l(\mathscr{A}_V) \leqslant n-2$, which contradicts its maximality. □

Corollary 7.10. *Let $\mathbb{F}$ be an algebraically closed field. If the length of a commutative subalgebra $\mathscr{A}$ in $M_n(\mathbb{F})$ is equal to $n-1$, then $\mathscr{A}$ is maximal under inclusion.*

Proof. It follows from Theorem 7.9 that if $l(\mathscr{A}) = n-1$, then $\mathscr{A}$ is generated by a nonderogatory matrix. Consequently, from Gerstenhaber's Theorem (Theorem 6.2 of this paper) we obtain that $\mathscr{A}$ is a maximal under inclusion commutative subalgebra. □

Remark 7.11. Theorem 7.9 and Corollary 7.10 are valid over an arbitrary field. The detailed proof will appear elsewhere.

8. Examples

In this section we show that there are maximal commutative subalgebras of nonmaximal length, including some classical algebras.

Example 8.1. Let us consider a subalgebra $\mathscr{A}_1 \subseteq M_n(\mathbb{F})$, consisting of all matrices of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{11} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & a_{11} \end{pmatrix}.$$

It is proved in [1, Example 2] that $\mathscr{A}_1$ is a maximal commutative subalgebra in $M_n(\mathbb{F})$. It is local and the index of nilpotency of the radical is equal to 2. Then it follows from Lemma 7.8

that $l(\mathscr{A}_1) \leqslant 1$. Straightforward computations show that $\mathscr{A}_1$ is not isomorphic to a field, i.e., $l(\mathscr{A}_1) \neq 0$. Hence, $l(\mathscr{A}_1) = 1$.

Example 8.2 (Schur algebra [14]). This is a commutative subalgebra in $M_n(\mathbb{F})$ of the maximal possible dimension.

Let $n = k + m,\ k, m \in \mathbb{N}$ and $|k - m| \leqslant 1$. Let $\mathscr{A}_S \subseteq M_n(\mathbb{F})$ be the following commutative subalgebra:

$$\mathscr{A}_S = \left\{ \left(\begin{array}{c|c} xI_k & Z \\ \hline 0 & xI_m \end{array} \right) \middle| \, x \in \mathbb{F},\ Z \in M_{k\times m}(\mathbb{F}) \right\}$$

Its dimension $\dim_{\mathbb{F}}(\mathscr{A}_S) = 1 + km = [n^2/4] + 1$ is the highest possible for a commutative matrix subalgebra, consequently, $\mathscr{A}_S$ is a maximal commutative subalgebra of $M_{n\times n}(\mathbb{F})$. $\mathscr{A}_S$ is local and $J(\mathscr{A}_S)$ consists of those matrices in which $x = 0$. Therefore, $N(J(\mathscr{A}_S)) = 2$ and $l(\mathscr{A}_S) \leqslant 1$. Straightforward computations show that $\mathscr{A}_S$ is not isomorphic to a field. Thus $l(\mathscr{A}_S) \neq 0$. Hence, $l(\mathscr{A}_S) = 1$.

Example 8.3 (Courter's algebra [2]). This is a famous maximal commutative subalgebra in $M_{14}(\mathbb{F})$. The dimension $\dim(\mathscr{A}_C) = 13$. And it was found by Courter in order to disprove the Gerstenhaber's conjecture that the dimension of a maximal commutative subalgebra in $M_n(\mathbb{F})$ is always greater than or equal to n, see [1,2,3]. Let us note that it is the minimal such example.

Consider the set J consisting of all matrices of order 14 of the following form:

$$\left(\begin{array}{cc|cccccccccc|cc}
0 & 0 & \multicolumn{10}{c|}{O_{2\times 10}} & \multicolumn{2}{c}{O_2} \\
0 & 0 & & & & & & & & & & & & \\ \hline
x_{11} & 0 & & & & & & & & & & & & \\
0 & x_{11} & & & & & & & & & & & & \\
x_{12} & 0 & & & & & & & & & & & & \\
0 & x_{12} & & & & & & & & & & & & \\
x_{21} & 0 & \multicolumn{10}{c|}{O_{10}} & \multicolumn{2}{c}{O_{10\times 2}} \\
0 & x_{21} & & & & & & & & & & & & \\
x_{22} & 0 & & & & & & & & & & & & \\
0 & x_{22} & & & & & & & & & & & & \\
z_{11} & z_{12} & & & & & & & & & & & & \\
z_{21} & z_{22} & & & & & & & & & & & & \\ \hline
y_{11} & y_{12} & z_{11} & z_{12} & z_{21} & z_{22} & 0 & 0 & 0 & 0 & x_{11} & x_{12} & 0 & 0 \\
y_{21} & y_{22} & 0 & 0 & 0 & 0 & z_{11} & z_{12} & z_{21} & z_{22} & x_{21} & x_{22} & 0 & 0
\end{array} \right),$$

where $x_{ij},\ y_{ij}$ and z_{ij} are arbitrary elements of $\mathbb{F}$, $O_{k\times m}$ denotes the zero matrix of size $k \times m$, and O_p denotes the zero matrix of order p. It follows from the definition of the set J that it is closed under multiplication and consists of pairwise commuting matrices. Let matrix $\delta \in J$. It can be partitioned into blocks as follows:

$$\delta = \begin{pmatrix} O_2 & O & O \\ A & O_{10} & O \\ Y & B & O_2 \end{pmatrix}.$$

Let

$$\delta' = \begin{pmatrix} O_2 & O & O \\ A' & O_{10} & O \\ Y' & B' & O_2 \end{pmatrix}$$

also be a matrix from J. Then

$$\delta'\delta = \begin{pmatrix} O_2 & O & O \\ O & O_{10} & O \\ B'A & O & O_2 \end{pmatrix}.$$

Consequently, the product of any three matrices from J is zero. Consider $\mathscr{A}_C = \langle I, J \rangle$, here $I \in M_{14}(\mathbb{F})$ is a 14×14 identity matrix. Then $\mathscr{A}_C$ is a local commutative subalgebra of $M_{14}(\mathbb{F})$ of dimension 13 with the index of nilpotency of the radical equal to 3. It is shown in [1, Example 4] that $\mathscr{A}_C$ is a maximal commutative subalgebra. It follows from Lemma 7.8 that $l(\mathscr{A}_C) \leqslant 2$. Let us consider the subset $\mathscr{S}$ of $\mathscr{A}_C$ consisting of the following matrices $X_{11} = E_{3,1} + E_{4,2} + E_{13,11}$, $X_{12} = E_{5,1} + E_{6,2} + E_{13,12}$, $X_{21} = E_{7,1} + E_{8,2} + E_{14,11}$, $X_{22} = E_{9,1} + E_{10,2} + E_{14,12}$, $Z_{11} = E_{11,1} + E_{13,3} + E_{14,7}$, $Z_{12} = E_{11,2} + E_{13,4} + E_{14,8}$, $Z_{21} = E_{12,1} + E_{13,5} + E_{14,9}$, $Z_{22} = E_{12,2} + E_{13,6} + E_{14,10}$ and the identity matrix. Each of these matrices has zeros in the lower left corner (the elements from the intersection of the 13-th and 14-th rows with the first and second columns). Hence, $L_1(\mathscr{S}) \neq \mathscr{A}_C$. But those elements can be obtained on the second step for example as follows: $E_{13,1} = X_{11}Z_{11}$, $E_{13,2} = X_{11}Z_{12}$, $E_{14,1} = X_{21}Z_{11}$ and $E_{14,2} = X_{22}Z_{22}$. Consequently, $\mathscr{S}$ is a generating system for $\mathscr{A}_C$ and $l(\mathscr{S}) = 2$. Hence, $l(\mathscr{A}_C) = 2$.

Corollary 8.4. *These examples also show that the bound obtained in Lemma 7.8 is sharp.*

Acknowledgments

The authors would like to express their gratitude to Professor Thomas J. Laffey for the attention given to the work and for the useful comments.

References

[1] W.C. Brown, F.W. Call, Maximal commutative subalgebras of $n \times n$ matrices, Comm. Algebra 21 (12) (1993) 4439–4460.
[2] R.C. Courter, The dimension of maximal commutative subalgebras of K_n, Duke Math. J. 32 (1965) 225–232.
[3] M. Gerstenhaber, On dominance and varieties of commuting matrices, Ann. Math. 73 (2) (1961) 324–348.
[4] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[5] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
[6] T.J. Laffey, The minimal dimension of maximal commutative subalgebras of full matrix algebras, Linear Algebra Appl. 71 (1985) 199–212.
[7] T.J. Laffey, S. Lazarus, Two-generated commutative matrix subalgebras, Linear Algebra Appl. 147 (1991) 249–273.
[8] A.I. Mal'cev, Basics of Linear Algebra, Nauka, Moscow, 1975 (in Russian).
[9] O.V. Markova, On the length of upper-triangular matrix algebra, Uspekhi Matem. Nauk 60 (5) (2005) 177–178 (in Russian). English translation: Russian Mathematical Surveys 60 (5) (2005) 984–985.
[10] O.V. Markova, Length computation of matrix subalgebras of special type, Fundam. Appl. Math. 13 (4) (2007) 165–197 (in Russian).
[11] C.J. Pappacena, An upper bound for the length of a finite-dimensional algebra, J. Algebra 197 (1997) 535–545.
[12] A. Paz, An application of the Cayley–Hamilton theorem to matrix polynomials in several variables, Linear and Multilinear Algebra 15 (1984) 161–170.
[13] R. Pierce, Associative Algebras, Springer-Verlag, Berlin, 1982.
[14] I. Schur, Zur Theorie der Vertauschbären Matrizen, J. Reine Angew. Math. 130 (1905) 66–76.
[15] A.J.M. Spencer, R.S. Rivlin, The theory of matrix polynomials and its applications to the mechanics of isotropic continua, Arch. Rat. Mech. Anal. 2 (1959) 309–336.

[16] A.J.M. Spencer, R.S. Rivlin, Further results in the theory of matrix polynomials, Arch. Rat. Mech. Anal. 4 (1960) 214–230.
[17] D.A. Suprunenko, R.I. Tyshkevich, Commutative Matrices, Academic Press, New York, NY, 1968.
[18] A. Wadsworth, The algebra generated by two commuting matrices, Linear and Multilinear Algebra 27 (1990) 159–162.