Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

COMMUTATORS FROM A HYPERPLANE OF MATRICES*

CLÉMENT DE SEGUINS PAZZIS†

Abstract. Denote by $\mathrm{M}_n(\mathbb{K})$ the algebra of n by n matrices with entries in the field $\mathbb{K}$. A theorem of Albert and Muckenhoupt states that every trace zero matrix of $\mathrm{M}_n(\mathbb{K})$ can be expressed as $AB - BA$ for some pair $(A, B) \in \mathrm{M}_n(\mathbb{K})^2$. Assuming that $n > 2$ and that $\mathbb{K}$ has more than 3 elements, it is proved that the matrices A and B can be required to belong to an arbitrary given hyperplane of $\mathrm{M}_n(\mathbb{K})$.

Key words. Commutator, Trace, Hyperplane, Matrices.

AMS subject classifications. 15A24, 15A30.

1. Introduction.

1.1. The problem. In this article, we let $\mathbb{K}$ be an arbitrary field. We denote by $\mathrm{M}_n(\mathbb{K})$ the algebra of square matrices with n rows and entries in $\mathbb{K}$, and by $\mathfrak{sl}_n(\mathbb{K})$ its hyperplane of trace zero matrices. The trace of a matrix $M \in \mathrm{M}_n(\mathbb{K})$ is denoted by $\mathrm{tr} M$. Given two matrices A and B of $\mathrm{M}_n(\mathbb{K})$, one sets

$$[A, B] := AB - BA,$$

known as the commutator, or Lie bracket, of A and B. Obviously, $[A, B]$ belongs to $\mathfrak{sl}_n(\mathbb{K})$. Although it is easy to see that the linear subspace spanned by the commutators is $\mathfrak{sl}_n(\mathbb{K})$, it is more difficult to prove that every trace zero matrix is actually a commutator, a theorem which was first proved by Shoda [9] for fields of characteristic 0, and later generalized to all fields by Albert and Muckenhoupt [1]. Recently, exciting new developments on this topic have appeared; most notably, the long-standing conjecture that the result holds for all principal ideal domains has just been solved by Stasinski [10] (the case of integers had been worked out earlier by Laffey and Reams [5]).

Here, we shall consider the following variation of the above problem:

> Given a (linear) hyperplane $\mathcal{H}$ of $\mathrm{M}_n(\mathbb{K})$, is it true that every trace zero matrix is the commutator of two matrices of $\mathcal{H}$?

*Received by the editors on July 7, 2013. Accepted for publication on February 2, 2014. Handling Editor: Raphael Loewy.

†Université de Versailles Saint-Quentin-en-Yvelines, Laboratoire de Mathématiques de Versailles, 45 avenue des Etats-Unis, 78035 Versailles cedex, France (dsp.prof@gmail.com).

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Our first motivation is that this constitutes a natural generalization of the following result of Thompson:

THEOREM 1.1 (Thompson, Theorem 5 of [11]). *Assume that $n \geq 3$. Then,* $[\mathfrak{sl}_n(\mathbb{K}), \mathfrak{sl}_n(\mathbb{K})] = \mathfrak{sl}_n(\mathbb{K})$.

Another motivation stems from the following known theorem:

THEOREM 1.2 (Proposition 4 of [8]). *Let $\mathcal{V}$ be a linear subspace of $\mathrm{M}_n(\mathbb{K})$ with* $\operatorname{codim}\mathcal{V} < n-1$. *Then,* $\mathfrak{sl}_n(\mathbb{K}) = \operatorname{span}\bigl\{[A,B] \mid (A,B) \in \mathcal{V}^2\bigr\}$.

Thus, a natural question to ask is whether, in the above situation, every trace zero matrix is a commutator of two matrices of $\mathcal{V}$. Studying the case of hyperplanes is an obvious first step in that direction (and a rather non-trivial one, as we shall see).

An additional motivation is the corresponding result for products (instead of commutators) that we have obtained in [8]:

THEOREM 1.3 (Theorem 3 of [8]). *Let $\mathcal{H}$ be a (linear) hyperplane of $\mathrm{M}_n(\mathbb{K})$, with $n > 2$. Then, every matrix of $\mathrm{M}_n(\mathbb{K})$ splits up as AB for some $(A,B) \in \mathcal{H}^2$.*

1.2. Main result. In the present paper, we shall prove the following theorem:

THEOREM 1.4. *Assume that $\#\mathbb{K} > 3$ and $n > 2$. Let $\mathcal{H}$ be an arbitrary hyperplane of $\mathrm{M}_n(\mathbb{K})$. Then, every trace zero matrix of $\mathrm{M}_n(\mathbb{K})$ splits up as $AB - BA$ for some $(A,B) \in \mathcal{H}^2$.*

Let us immediately discard an easy case. Assume that $\mathcal{H}$ does not contain the identity matrix I_n. Then, given $(A,B) \in \mathrm{M}_n(\mathbb{K})^2$, we have

$$[\lambda I_n + A, \mu I_n + B] = [A,B]$$

for all $(\lambda,\mu) \in \mathbb{K}^2$, and obviously there is a unique pair $(\lambda,\mu) \in \mathbb{K}^2$ such that $\lambda I_n + A$ and $\mu I_n + B$ belong to $\mathcal{H}$. In that case, it follows from the Albert-Muckenhoupt theorem that every matrix of $\mathfrak{sl}_n(\mathbb{K})$ is a commutator of matrices of $\mathcal{H}$. Thus, the only case left to consider is the one when $I_n \in \mathcal{H}$. As we shall see, this is a highly non-trivial problem. Our proof will broadly consist in refining Albert and Muckenhoupt's method.

The case $n = 2$ can be easily described over any field:

PROPOSITION 1.5. *Let $\mathcal{H}$ be a hyperplane of $\mathrm{M}_2(\mathbb{K})$.*

(a) If $\mathcal{H}$ contains I_2, then $[\mathcal{H},\mathcal{H}]$ is a 1-dimensional linear subspace of $\mathrm{M}_2(\mathbb{K})$.
(b) If $\mathcal{H}$ does not contain I_2, then $[\mathcal{H},\mathcal{H}] = \mathfrak{sl}_2(\mathbb{K})$.

Proof. Point (b) has just been explained. Assume now that $I_2 \in \mathcal{H}$. Then, there

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

are matrices A and B such that (I_2, A, B) is a basis of $\mathcal{H}$. For all $(a, b, c, a', b', c') \in \mathbb{K}^6$, one finds

$$[aI_2 + bA + cB\,,\, a'I_2 + b'A + c'B] = (bc' - b'c)[A, B].$$

Moreover, as A is a 2×2 matrix and not a scalar multiple of the identity, it is similar to a companion matrix, whence the space of all matrices which commute with A is $\operatorname{span}(I_2, A)$. This yields $[A, B] \neq 0$. As obviously $\mathbb{K} = \{bc' - b'c \mid (b, c, b', c') \in \mathbb{K}^4\}$, we deduce that $[\mathcal{H}, \mathcal{H}] = \mathbb{K}\,[A, B]$ with $[A, B] \neq 0$. □

1.3. Additional definitions and notation.

- Given a subset $\mathcal{X}$ of $\mathrm{M}_n(\mathbb{K})$, we set

$$[\mathcal{X}, \mathcal{X}] := \{[A, B] \mid (A, B) \in \mathcal{X}^2\}.$$

- The canonical basis of $\mathbb{K}^n$ is denoted by $(e_1, \dots, e_n)$.
- Given a basis $\mathcal{B}$ of $\mathbb{K}^n$, the matrix of coordinates of $\mathcal{B}$ in the canonical basis of $\mathbb{K}^n$ is denoted by $P_{\mathcal{B}}$.
- Given i and j in $[\![1, n]\!]$, one denotes by $E_{i,j}$ the matrix of $\mathrm{M}_n(\mathbb{K})$ with all entries zero except the one at the (i, j)-spot, which equals 1.
- A matrix of $\mathrm{M}_n(\mathbb{K})$ is *cyclic* when its minimal polynomial has degree n or, equivalently, when it is similar to a companion matrix.
- The n by n nilpotent Jordan matrix is denoted by

$$J_n = \begin{bmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{bmatrix}.$$

- A Hessenberg matrix is a square matrix $A = (a_{i,j}) \in \mathrm{M}_n(\mathbb{K})$ in which $a_{i,j} = 0$ whenever $i > j + 1$. In that case, we set

$$\ell(A) := \{j \in [\![1, n-1]\!] :\ a_{j+1,j} \neq 0\}.$$

- One equips $\mathrm{M}_n(\mathbb{K})$ with the non-degenerate symmetric bilinear form

$$b : (M, N) \mapsto \operatorname{tr}(MN),$$

to which orthogonality refers in the rest of the article.

Given $A \in \mathrm{M}_n(\mathbb{K})$, one sets

$$\mathrm{ad}_A : M \in \mathrm{M}_n(\mathbb{K}) \mapsto [A, M] \in \mathrm{M}_n(\mathbb{K}),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

which is an endomorphism of the vector space $\mathrm{M}_n(\mathbb{K})$; its kernel is the centralizer

$$\mathcal{C}(A) := \{M \in \mathrm{M}_n(\mathbb{K}) : AM = MA\}$$

of the matrix A. Recall the following nice description of the range of ad_A, which follows from the rank theorem and the basic observation that ad_A is skew-symmetric for the bilinear form $(M, N) \mapsto \mathrm{tr}(MN)$:

Lemma 1.6. *Let $A \in \mathrm{M}_n(\mathbb{K})$. The range of ad_A is the orthogonal of $\mathcal{C}(A)$, that is the set of all $N \in \mathrm{M}_n(\mathbb{K})$ for which*

$$\forall B \in \mathcal{C}(A),\ \mathrm{tr}(B\,N) = 0.$$

In particular, if A is cyclic then its centralizer is $\mathbb{K}[A] = \mathrm{span}(I_n, A, \dots, A^{n-1})$, whence $\mathrm{im}(\mathrm{ad}_A)$ is defined by a set of n linear equations:

Lemma 1.7. *Let $A \in \mathrm{M}_n(\mathbb{K})$ be a cyclic matrix. The range of ad_A is the set of all $N \in \mathrm{M}_n(\mathbb{K})$ for which*

$$\forall k \in [\![0, n-1]\!],\ \mathrm{tr}(A^k\,N) = 0.$$

Remark 1. Interestingly, the two special cases below yield the strategy for Shoda's approach and Albert and Muckenhoupt's, respectively:

(i) Let D be a diagonal matrix of $\mathrm{M}_n(\mathbb{K})$ with distinct diagonal entries. Then, the centralizer of D is the space $\mathcal{D}_n(\mathbb{K})$ of all diagonal matrices, and hence, $\mathrm{im}\,\mathrm{ad}_D$ is the space of all matrices with diagonal zero. As every trace zero matrix that is not a scalar multiple of the identity is similar to a matrix with diagonal zero [4], Shoda's theorem of [9] follows easily.

(ii) Consider the case of the Jordan matrix J_n. As J_n is cyclic, Lemma 1.7 yields that $\mathrm{im}(\mathrm{ad}_{J_n})$ is the set of all matrices $A = (a_{i,j}) \in \mathrm{M}_n(\mathbb{K})$ for which $\sum_{k=1}^{n-\ell} a_{k+\ell,k} = 0$ for all $\ell \in [\![0, n-1]\!]$. In particular, if $A = (a_{i,j}) \in \mathrm{M}_n(\mathbb{K})$ is Hessenberg, then this condition is satisfied whenever $\ell > 1$, and hence, $A \in \mathrm{im}(\mathrm{ad}_{J_n})$ if and only if $\mathrm{tr}A = 0$ and $\sum_{k=1}^{n-1} a_{k+1,k} = 0$. Albert and Muckenhoupt's proof is based upon the fact that, except for a few special cases, the similarity class of a matrix must contain a Hessenberg matrix A that satisfies the extra equation $\sum_{k=1}^{n-1} a_{k+1,k} = 0$.

2. Proof of the main theorem.

2.1. Proof strategy. Let $\mathcal{H}$ be a hyperplane of $\mathrm{M}_n(\mathbb{K})$. We already know that $[\mathcal{H}, \mathcal{H}] = \mathfrak{sl}_n(\mathbb{K})$ if $I_n \notin \mathcal{H}$. Thus, in the rest of the article, we will only consider the case when $I_n \in \mathcal{H}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Our proof will use three basic but potent principles:

(1) Given $A \in \mathfrak{sl}_n(\mathbb{K})$, if some $A_1 \in \mathcal{H}$ satisfies $A \in \operatorname{im}(\operatorname{ad}_{A_1})$ and $\mathcal{C}(A_1) \not\subset \mathcal{H}$, then $A \in [\mathcal{H}, \mathcal{H}]$. Indeed, in that situation, we find $A_2 \in \mathrm{M}_n(\mathbb{K})$ such that $A = [A_1, A_2]$, together with some $A_3 \in \mathcal{C}(A_1)$ for which $A_3 \notin \mathcal{H}$. Then, the affine line $A_2 + \mathbb{K}A_3$ is included in the inverse image of $\{A\}$ by ad_{A_1} and it has exactly one common point with $\mathcal{H}$.
(2) Let $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$ and $\lambda \in \mathbb{K}$. If there are matrices A_1 and A_2 such that $A = [A_1, A_2]$ and $\operatorname{tr}(B\,A_1) = \operatorname{tr}(B\,A_2) = 0$, then we also have $\operatorname{tr}((B - \lambda\,A)A_1) = \operatorname{tr}((B - \lambda\,A)A_2) = 0$. Indeed, equality $A = [A_1, A_2]$ ensures that $\operatorname{tr}(A\,A_1) = \operatorname{tr}(A\,A_2) = 0$ (see Lemma 1.6).
(3) Let $(A, B) \in \mathrm{M}_n(\mathbb{K})^2$ and $P \in \mathrm{GL}_n(\mathbb{K})$. Setting $\mathcal{G} := \{B\}^\perp$, we see that the assumption $A \in [\mathcal{G}, \mathcal{G}]$ implies $PAP^{-1} \in [P\mathcal{G}P^{-1}, P\mathcal{G}P^{-1}]$, while $P\mathcal{G}P^{-1} = \{PBP^{-1}\}^\perp$.

Now, let us give a rough idea of the proof strategy. One fixes $A \in \mathfrak{sl}_n(\mathbb{K})$ and aims at proving that $A \in [\mathcal{H}, \mathcal{H}]$. We fix a non-zero matrix B such that $\mathcal{H} = \{B\}^\perp$.

Our basic strategy is the Albert-Muckenhoupt method: We try to find a cyclic matrix M in $\mathcal{H}$ such that $A \in \operatorname{im}(\operatorname{ad}_M)$; if $A \notin \operatorname{ad}_M(\mathcal{H})$, then we learn that $\mathcal{C}(M) \subset \mathcal{H}$ (see principle (1) above), which yields additional information on B. Most of the time, we will search for such a cyclic matrix M among the nilpotent matrices with rank $n-1$. The most favorable situation is the one where A is either upper-triangular or Hessenberg with enough non-zero sub-diagonal entries: In these cases, we search for a good matrix M among the strictly upper-triangular matrices with rank $n-1$ (see Lemma 2.2). If this method yields no solution, then we learn precious information on the simultaneous reduction of the endomorphisms $X \mapsto AX$ and $X \mapsto BX$. Using changes of bases, we shall see that either the above method delivers a solution for a pair (A', B') that is simultaneously similar to (A, B), in which case Principle (3) shows that we have a solution for (A, B), or (I_n, A, B) is locally linearly dependent (see the definition below), or else $n = 3$ and A is similar to $\lambda I_3 + E_{2,3}$ for some $\lambda \in \mathbb{K}$. When (I_n, A, B) is locally linearly dependent and A is not of that special type, one uses the classification of locally linearly dependent triples to reduce the situation to the one where $B = I_n$, that is $\mathcal{H} = \mathfrak{sl}_n(\mathbb{K})$, and in that case the proof is completed by invoking Theorem 1.1. Finally, the case when A is similar to $\lambda I_3 + E_{2,3}$ for some $\lambda \in \mathbb{K}$ will be dealt with independently (Section 2.5) by applying Albert and Muckenhoupt's method for well-chosen companion matrices instead of a Jordan nilpotent matrix.

Let us finish these strategic considerations by recalling the notion of local linear dependence:

Definition 2.1. Given vector spaces U and V, linear maps $f_1, \dots, f_n$ from U to V are called *locally linearly dependent* (in abbreviated form: *LLD*) when the vectors

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

$f_1(x), \dots, f_n(x)$ are linearly dependent for all $x \in U$.

We adopt a similar definition for matrices by referring to the linear maps that are canonically associated with these matrices.

2.2. The basic lemma.

LEMMA 2.2. *Let* $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$ *be with* $B = (b_{i,j}) \neq 0$, *and set* $\mathcal{H} := \{B\}^\perp$. *In each one of the following cases,* A *belongs to* $[\mathcal{H}, \mathcal{H}]$:

(a) $\#\mathbb{K} > 2$, A *is upper-triangular and* B *is not Hessenberg.*
(b) $\#\mathbb{K} > 3$, A *is Hessenberg and there exist* $i \in [\![2, n-1]\!]$ *and* $j \in [\![3, n]\!] \setminus \{i\}$ *such that* $\{1, i\} \subset \ell(A)$ *and* $b_{j,1} \neq 0$.

Proof. We use a *reductio ad absurdum*, assuming that $A \notin [\mathcal{H}, \mathcal{H}]$. We write $A = (a_{i,j})$.

(a) Assume that $\#\mathbb{K} > 2$, that A is upper-triangular and that B is not Hessenberg. We choose a pair $(l, l') \in [\![1, n]\!]^2$ such that $b_{l,l'} \neq 0$, with $l - l'$ maximal for such pairs. Thus, $l - l' > 1$. Let $(x_1, \dots, x_{n-1}) \in (\mathbb{K}^*)^{n-1}$, and set

$$\beta := \frac{\sum_{k=1}^{n-1} b_{k+1,k}\, x_k}{b_{l,l'}} \quad \text{and} \quad M := \sum_{k=1}^{n-1} x_k\, E_{k,k+1} - \beta\, E_{l',l}.$$

We see that M is nilpotent of rank $n-1$, and hence, it is cyclic. One notes that $M \in \mathcal{H}$. Moreover, $\operatorname{tr}(AM^k) = 0$ for all $k \geq 1$, because A is upper-triangular and M is strictly upper-triangular, whereas $\operatorname{tr}(A) = 0$ by assumption. Thus, $A \in \operatorname{im}(\operatorname{ad}_M)$. As it is assumed that $A \notin \operatorname{ad}_M(\mathcal{H})$, one deduces from principle (1) in Section 2.1 that $\mathcal{C}(M) \subset \mathcal{H}$; in particular $\operatorname{tr}(M^{l-l'}B) = 0$, which, as $b_{i,j} = 0$ whenever $i - j > l - l'$, reads

$$b_{l-l'+1,1}\, x_1 x_2 \cdots x_{l-l'} + b_{l-l'+2,2}\, x_2 x_3 \cdots x_{l-l'+1} + \cdots + b_{n,n-l+l'}\, x_{n-l+l'} \cdots x_{n-1} = 0.$$

Here, we have a polynomial with degree at most 1 in each variable x_i, and this polynomial vanishes at every $(x_1, \dots, x_{n-1}) \in (\mathbb{K}^*)^{n-1}$, with $\#\mathbb{K}^* \geq 2$. It follows that $b_{i,j} = 0$ for all $(i, j) \in [\![1, n]\!]^2$ with $i - j = l - l'$, and the special case $(i, j) = (l, l')$ yields a contradiction.

(b) Now, we assume that $\#\mathbb{K} > 3$, that A is Hessenberg and that there exist $i \in [\![2, n]\!]$ and $j \in [\![3, n]\!] \setminus \{i\}$ such that $\{1, i\} \subset \ell(A)$ and $b_{j,1} \neq 0$. The proof strategy is similar to the one of case (a), with additional technicalities. One chooses a pair $(l, l') \in [\![1, n]\!]^2$ such that $b_{l,l'} \neq 0$, with $l - l'$ maximal for such pairs (again, the assumptions yield $l - l' \geq j - 1 > 1$). As $a_{2,1} \neq 0$, no generality is lost in assuming that $a_{2,1} = 1$. We introduce the formal polynomial

$$\mathbf{p} := \sum_{k=1}^{n-2} a_{k+2,k+1}\, \mathbf{x}_k \in \mathbb{K}[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-2}].$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Let $(x_1,\dots,x_{n-2}) \in (\mathbb{K}^*)^{n-2}$, and set

$$\alpha := \mathbf{p}(x_1,\dots,x_{n-2}) \quad \text{and} \quad \beta := \frac{\alpha\, b_{2,1} - \sum_{k=1}^{n-2} x_k\, b_{k+2,k+1}}{b_{l,l'}}.$$

Finally, set

$$M := -\alpha\, E_{1,2} + \sum_{k=1}^{n-2} x_k\, E_{k+1,k+2} + \beta\, E_{l',l}.$$

The definition of M shows that $\operatorname{tr}(MA) = \operatorname{tr}(MB) = 0$, and in particular $M \in \mathcal{H}$. Assume now that $\mathbf{p}(x_1,\dots,x_{n-2}) \neq 0$. Then, M is cyclic as it is nilpotent with rank $n-1$. As A is Hessenberg, we also see that $\operatorname{tr}(M^k A) = 0$ for all $k \geq 2$. Thus, $\operatorname{tr}(M^k A) = 0$ for every non-negative integer k, and hence, Lemma 1.7 yields $A \in \operatorname{im}(\operatorname{ad}_M)$. It ensues that $\mathcal{C}(M) \subset \mathcal{H}$, and in particular $\operatorname{tr}(M^{j-1}B) = 0$. As $l - l' > 1$, we see that, for all $(a,b) \in [\![1,n]\!]^2$ with $b - a \leq l - l'$, and every integer $c > 1$, the matrices M^c and $\left(-\alpha\, E_{1,2} + \sum_{k=1}^{n-2} x_k\, E_{k+1,k+2}\right)^c$ have the same entry at the (a,b)-spot; in particular, for all $k \in [\![2, n-j+1]\!]$, the entry of M^{j-1} at the $(k, j+k-1)$-spot is $x_{k-1} x_k \cdots x_{k-3+j}$, and the entry of M^{j-1} at the $(1,j)$-spot is $-\alpha\, x_1 \cdots x_{j-2}$; moreover, for all $(a,b) \in [\![1,n]\!]^2$ with $b - a \leq \ell - \ell'$ and $b - a \neq j-1$, the entry of M^{j-1} at the (a,b)-spot is 0. Therefore, equality $\operatorname{tr}(M^{j-1}B) = 0$ yields

$$-b_{j,1}\, \alpha\, x_1 \cdots x_{j-2} + b_{j+1,2}\, x_1 \cdots x_{j-1} + b_{j+2,3}\, x_2 \cdots x_j + \cdots + b_{n,n-j+1}\, x_{n-j} \cdots x_{n-2} = 0.$$

We conclude that we have established the following identity: For the polynomial

$$\mathbf{q} := \mathbf{p} \times \Big(-b_{j,1}\, \mathbf{p}\, \mathbf{x}_1 \cdots \mathbf{x}_{j-2} + b_{j+1,2}\, \mathbf{x}_1 \cdots \mathbf{x}_{j-1} + \cdots + b_{n,n-j+1}\, \mathbf{x}_{n-j} \cdots \mathbf{x}_{n-2}\Big),$$

we have

$$\forall (x_1,\dots,x_{n-2}) \in (\mathbb{K}^*)^{n-2}, \quad \mathbf{q}(x_1,\dots,x_{n-2}) = 0.$$

Noting that $\mathbf{q}$ has degree at most 3 in each variable, we split the discussion into two main cases.

Case 1. $\#\mathbb{K} > 4$.
Then, $\#\mathbb{K}^* > 3$ and hence $\mathbf{q} = 0$. As $\mathbf{p} \neq 0$ (remember that $a_{i+1,i} \neq 0$), it follows that

$$-b_{j,1}\, \mathbf{p}\, \mathbf{x}_1 \cdots \mathbf{x}_{j-2} + b_{j+1,2}\, \mathbf{x}_1 \cdots \mathbf{x}_{j-1} + b_{j+2,3}\, \mathbf{x}_2 \cdots \mathbf{x}_j + \cdots + b_{n,n-j+1}\, \mathbf{x}_{n-j} \cdots \mathbf{x}_{n-2} = 0.$$

As $b_{j,1} \neq 0$, identifying the coefficients of the monomials of type $\mathbf{x}_1 \cdots \mathbf{x}_{j-2} \mathbf{x}_k$ with $k \in [\![1, n-2]\!] \setminus \{j-1\}$ leads to $a_{k+2,k+1} = 0$ for all such k. This contradicts the assumption that $a_{i+1,i} \neq 0$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Case 2. $\#\mathbb{K} = 4$.

A polynomial of $\mathbb{K}[t]$ which vanishes at every non-zero element of $\mathbb{K}$ must be a multiple of $t^3 - 1$. In particular, if such a polynomial has degree at most 3, we may write it as $\alpha_3 t^3 + \alpha_2 t^2 + \alpha_1 t + \alpha_0$, and we obtain $\alpha_3 = -\alpha_0$. From there, we split the discussion into two subcases.

Subcase 2.1. $i > j$.

Then, $\mathbf{q}$ has degree at most 2 in $\mathbf{x}_{i-1}$. Thus, if we see $\mathbf{q}$ as a polynomial in the sole variable $\mathbf{x}_{i-1}$, the coefficients of this polynomial must vanish for every specialization of $\mathbf{x}_1, \dots, \mathbf{x}_{i-2}, \mathbf{x}_i, \dots, \mathbf{x}_{n-2}$ in $\mathbb{K}^*$; extracting the coefficients of $(\mathbf{x}_{i-1})^2$ leads to the identity

$$\forall (x_1, \dots, x_{i-2}, x_i, \dots, x_{n-2}) \in (\mathbb{K}^*)^{n-3}, \;\; -b_{j,1}(a_{i+1,i})^2\, x_1 \cdots x_{j-2} + \mathbf{r}(x_1, \dots, x_{n-2}) = 0,$$

where $\mathbf{r} = \sum_{k=i-j+1}^{n-j} a_{i+1,i}\, b_{j+k,k+1}\, \mathbf{x}_k \cdots \mathbf{x}_{i-2}\mathbf{x}_i \cdots \mathbf{x}_{j-2+k}$. Noting that the degree of $-b_{j,1}(a_{i+1,i})^2\, \mathbf{x}_1 \cdots \mathbf{x}_{j-2} + \mathbf{r}$ is at most 1 in each variable, we deduce that this polynomial is zero. This contradicts the fact that the coefficient of $\mathbf{x}_1 \cdots \mathbf{x}_{j-2}$ is $-b_{j,1}(a_{i+1,i})^2$, which is non-zero according to our assumptions.

Subcase 2.2. $i < j$.

Let us fix $x_1, \dots, x_{i-2}, x_i, \dots, x_{n-2}$ in $\mathbb{K}^*$. The coefficient of $\mathbf{q}(x_1, \dots, x_{i-2}, \mathbf{x}_{i-1}, x_i, \dots, x_{n-2})$ with respect to $(\mathbf{x}_{i-1})^3$ is

$$-b_{j,1}(a_{i+1,i})^2\, x_1 \cdots x_{i-2} x_i \cdots x_{j-2}.$$

One the other hand, with

$$\mathbf{s} := \sum_{i \le k \le n-j} b_{j+k,k+1} \prod_{\ell=k}^{j-2+k} \mathbf{x}_\ell,$$

the coefficient of $\mathbf{q}(x_1, \dots, x_{i-2}, \mathbf{x}_{i-1}, x_i, \dots, x_{n-2})$ with respect to $(\mathbf{x}_{i-1})^0$ is

$$\mathbf{s}(x_1, \dots, x_{i-2}, x_i, \dots, x_{n-2}) \sum_{k \in [\![1,n-2]\!] \setminus \{i-1\}} a_{k+2,k+1}\, x_k.$$

Therefore, for all $(x_1, \dots, x_{n-2}) \in (\mathbb{K}^*)^{n-2}$,

$$\begin{aligned} b_{j,1}(a_{i+1,i})^2\, x_1 \cdots x_{i-2} x_i \cdots x_{j-2} \;&=\; \mathbf{s}(x_1, \dots, x_{i-2}, x_i, \dots, x_{n-2}) \\ &\quad \times \sum_{k \in [\![1,n-2]\!] \setminus \{i-1\}} a_{k+2,k+1}\, x_k. \end{aligned}$$

On both sides of this equality, we have polynomials of degree at most 2 in each variable. As $\#(\mathbb{K}^*) > 2$, we deduce the identity

$$b_{j,1}(a_{i+1,i})^2\, \mathbf{x}_1 \dots \mathbf{x}_{i-2}\mathbf{x}_i \cdots \mathbf{x}_{j-2} \;=\; \mathbf{s} \times \sum_{k \in [\![1,n-2]\!] \setminus \{i-1\}} a_{k+2,k+1}\, \mathbf{x}_k.$$

However, on the left-hand side of this identity is a non-zero homogeneous polynomial of degree $j-3$, whereas its right-hand side is a homogeneous polynomial of degree j. There lies a final contradiction. □

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

2.3. Reduction to the case when I_n, A, B are locally linearly dependent. In this section, we use Lemma 2.2 to prove the following result:

LEMMA 2.3. *Assume that $\#\mathbb{K} > 3$, let $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$ be such that $B \neq 0$, and set $\mathcal{H} := \{B\}^\perp$. Then, either $A \in [\mathcal{H}, \mathcal{H}]$, or (I_n, A, B) is LLD, or A is similar to $\lambda I_3 + E_{2,3}$ for some $\lambda \in \mathbb{K}$.*

In order to prove Lemma 2.3, one needs two preliminary results. The first one is a basic result in the theory of matrix spaces with rank bounded above.

LEMMA 2.4 (Lemma 2.4 of [6]). *Let m, n, p, q be positive integers, and $\mathcal{V}$ be a linear subspace of $\mathrm{M}_{m+p,n+q}(\mathbb{K})$ in which every matrix splits up as*

$$M = \begin{bmatrix} A(M) & ? \\ 0 & B(M) \end{bmatrix},$$

where $A(M) \in \mathrm{M}_{m,n}(\mathbb{K})$ and $B(M) \in \mathrm{M}_{p,q}(\mathbb{K})$. Assume that there is an integer r such that $\forall M \in \mathcal{V}$, $\mathrm{rk}(M) \leq r < \#\mathbb{K}$, and set $s := \max\{\mathrm{rk}(A(M)) \mid M \in \mathcal{V}\}$ and $t := \max\{\mathrm{rk}(B(M)) \mid M \in \mathcal{V}\}$. Then, $s + t \leq r$.

LEMMA 2.5. *Assume that $\#\mathbb{K} \geq 3$. Let V be a vector space over $\mathbb{K}$ and u be an endomorphism of V that is not a scalar multiple of the identity. Then, there are two linearly independent non-eigenvectors of u.*

Proof. As u is not a scalar multiple of the identity, some vector $x \in V \setminus \{0\}$ is not an eigenvector of u. Then, the 2-dimensional subspace $P := \mathrm{span}(x, u(x))$ contains x. As $u_{|P}$ is not a scalar multiple of the identity, u stabilizes at most two 1-dimensional subspaces of P. As $\#\mathbb{K} > 2$, there are at least four 1-dimensional subspaces of P, whence at least two of them are not stable under u. This proves our claim. ◻

Now, we are ready to prove Lemma 2.3.

Proof of Lemma 2.3. Throughout the proof, we assume that $A \notin [\mathcal{H}, \mathcal{H}]$ and that there is no scalar λ such that A is similar to $\lambda I_3 + E_{2,3}$. Our aim is to show that (I_n, A, B) is LLD.

Note that, for all $P \in \mathrm{GL}_n(\mathbb{K})$, no pair $(M, N) \in \mathrm{M}_n(\mathbb{K})^2$ satisfies both $[M, N] = P^{-1}AP$ and $\mathrm{tr}((P^{-1}BP)M) = \mathrm{tr}((P^{-1}BP)N) = 0$.

Let us say that a vector $x \in \mathbb{K}^n$ has *order* 3 when $\mathrm{rk}(x, Ax, A^2x) = 3$. Let $x \in \mathbb{K}^n$ be of order 3. Then, (x, Ax, A^2x) may be extended into a basis $\mathbf{B} = (x_1, x_2, x_3, x_4, \dots, x_n)$ of $\mathbb{K}^n$ such that $A' := P_{\mathbf{B}}^{-1}AP_{\mathbf{B}}$ is Hessenberg[1]. Moreover, one sees that $\{1, 2\} \subset \ell(A')$. Applying point (a) of Lemma 2.2, one obtains that the

[1]One finds such a basis by induction as follows: One sets $(x_1, x_2, x_3) := (x, Ax, A^2x)$ and, given $k \in [\![4, n]\!]$ such that $x_1, \dots, x_{k-1}$ are defined, one sets $x_k := Ax_{k-1}$ if $Ax_{k-1} \notin \mathrm{span}(x_1, \dots, x_{k-1})$, otherwise one chooses an arbitrary vector $x_k \in \mathbb{K}^n \setminus \mathrm{span}(x_1, \dots, x_{k-1})$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

entries in the first column of $P_{\mathbf{B}}^{-1}\,B\,P_{\mathbf{B}}$ are all zero starting from the third one, which means that $Bx \in \operatorname{span}(x, Ax)$.

Let now $x \in \mathbb{K}^n$ be a vector that is not of order 3. If x and Ax are linearly dependent, then x, Ax, Bx are linearly dependent. Thus, we may assume that $\operatorname{rk}(x, Ax) = 2$ and $A^2x \in \operatorname{span}(x, Ax)$. We split $\mathbb{K}^n = \operatorname{span}(x, Ax) \oplus F$ and we choose a basis $(f_3, \dots, f_n)$ of F. For $\mathbf{B} := (x, Ax, f_3, \dots, f_n)$, we now have, for some $(\alpha, \beta) \in \mathbb{K}^2$ and some $N \in \mathrm{M}_{n-2}(\mathbb{K})$,

$$P_{\mathbf{B}}^{-1}\,A\,P_{\mathbf{B}} = \begin{bmatrix} K & ? \\ 0 & N \end{bmatrix}, \quad \text{where } K = \begin{bmatrix} 0 & \alpha \\ 1 & \beta \end{bmatrix}.$$

From there, we split the discussion into several cases, depending on the form of N and its relationship with K.

Case 1. $N \notin \mathbb{K}I_{n-2}$.
Then, there is a vector $y \in \mathbb{K}^{n-2}$ for which y and Ny are linearly independent. Denoting by z the vector of F with coordinate list y in $(f_3, \dots, f_n)$, one obtains $\operatorname{rk}(x, Ax, z, Az) = 4$, and hence, one may extend (x, Ax, z, Az) into a basis $\mathbf{B}'$ of $\mathbb{K}^n$ such that $A' := P_{\mathbf{B}'}^{-1} A P_{\mathbf{B}'}$ is Hessenberg with $\{1, 3\} \subset \ell(A')$. Point (b) of Lemma 2.2 shows that, in the first column of $P_{\mathbf{B}'}^{-1} B P_{\mathbf{B}'}$, all the entries must be zero starting from the fourth one, yielding $Bx \in \operatorname{span}(x, Ax, z)$. As $N \notin \mathbb{K}I_{n-2}$, we know from Lemma 2.5 that we may find another vector $z' \in F \setminus \mathbb{K}z$ such that $\operatorname{rk}(x, Ax, z', Az') = 4$, which yields $Bx \in \operatorname{span}(x, Ax, z')$. Thus, $Bx \in \operatorname{span}(x, Ax, z) \cap \operatorname{span}(x, Ax, z') = \operatorname{span}(x, Ax)$.

Case 2. $N = \lambda\, I_{n-2}$ for some $\lambda \in \mathbb{K}$.

Subcase 2.1. λ is not an eigenvalue of K.
Then, $G := \operatorname{Ker}(A - \lambda I_n)$ has dimension $n-2$. For $z \in \mathbb{K}^n$, denote by p_z the monic generator of the ideal $\{q \in \mathbb{K}[t] : \; q(A)z = 0\}$. Recall that, given y and z in $\mathbb{K}^n$ for which p_y and p_z are mutually prime, one has $p_{y+z} = p_y p_z$. In particular, as p_x has degree 2, p_z has degree 3 for every $z \in (\mathbb{K}x \oplus G) \setminus (\mathbb{K}x \cup G)$, that is every z in $(\mathbb{K}x \oplus G) \setminus (\mathbb{K}x \cup G)$ has order 3; thus, $\operatorname{rk}(z, Az, Bz) \leq 2$ for all such z. Moreover, it is obvious that $\operatorname{rk}(z, Az, Bz) \leq 2$ for all $z \in G$.

Let us choose a non-zero linear form φ on $\mathbb{K}x \oplus G$ such that $\varphi(x) = 0$. For every $z \in \mathbb{K}x \oplus G$, set

$$M(z) = \begin{bmatrix} \varphi(z) & 0 & 0 & 0 \\ 0 & z & Az & Bz \end{bmatrix} \in \mathrm{M}_{n+1,4}(\mathbb{K}).$$

Then, with the above results, we know that $\operatorname{rk}(M(z)) \leq 3$ for all $z \in \mathbb{K}x \oplus G$. On the other hand, $\max\{\operatorname{rk}(\varphi(z)) \mid z \in (\mathbb{K}x \oplus G)\} = 1$. Using Lemma 2.4, we deduce that $\operatorname{rk}(z, Az, Bz) \leq 2$ for all $z \in \mathbb{K}x \oplus G$. In particular, $\operatorname{rk}(x, Ax, Bx) \leq 2$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Subcase 2.2. λ is an eigenvalue of K with multiplicity 1.
Then, there are eigenvectors y and z of A, with distinct corresponding eigenvalues, such that $x = y + z$. Thus, (y, z) may be extended into a basis $\mathbf{B}'$ of $\mathbb{K}^n$ such that $P_{\mathbf{B}'}^{-1} A P_{\mathbf{B}'}$ is upper-triangular. It follows from point (a) of Lemma 2.2 that $P_{\mathbf{B}'}^{-1} B P_{\mathbf{B}'}$ is Hessenberg, and in particular $By \in \operatorname{span}(y, z)$. Starting from (z, y) instead of (y, z), one finds $Bz \in \operatorname{span}(y, z)$. Therefore, all the vectors $y + z$, $A(y + z)$ and $B(y + z)$ belong to the 2-dimensional space $\operatorname{span}(y, z)$, which yields $\operatorname{rk}(x, Ax, Bx) \leq 2$.

Subcase 2.3. λ is an eigenvalue of K with multiplicity 2.
Then, the characteristic polynomial of A is $(t - \lambda)^n$.

- Assume that $n \geq 4$. One chooses an eigenvector y of A in $\operatorname{span}(x, Ax)$, so that (y, x) is a basis of $\operatorname{span}(x, Ax)$. Then, one chooses an arbitrary non-zero vector $u \in F$, and one extends (y, x, u) into a basis $\mathbf{B}'$ of $\mathbb{K}^n$ such that $P_{\mathbf{B}'}^{-1} A P_{\mathbf{B}'}$ is upper-triangular. Applying point (a) of Lemma 2.2 once more yields $Bx \in \operatorname{span}(y, x, u) = \operatorname{span}(x, Ax, u)$. As $n \geq 4$, we can choose another vector $v \in F \setminus \mathbb{K}u$, and the above method yields $Bx \in \operatorname{span}(x, Ax, v)$, while x, Ax, u, v are linearly independent. Therefore, $Bx \in \operatorname{span}(x, Ax, u) \cap \operatorname{span}(x, Ax, v) = \operatorname{span}(x, Ax)$.
- Finally, assume that $n = 3$. As A is not similar to $\lambda I_3 + E_{2,3}$, the only remaining option is that $\operatorname{rk}(A - \lambda I_3) = 2$. Then, we can find a linear form φ on $\mathbb{K}^3$ with kernel $\operatorname{Ker}(A - \lambda I_3)^2$. Every vector $z \in \mathbb{K}^3 \setminus \operatorname{Ker}(A - \lambda I_3)^2$ has order 3. Therefore, for every $z \in \mathbb{K}^3$, either $\varphi(z) = 0$ or $\operatorname{rk}(z, Az, Bz) \leq 2$. With the same line of reasoning as in Subcase 2.1, we obtain $\operatorname{rk}(x, Ax, Bx) \leq 2$. This completes the proof. ∎

Thus, only two situations are left to consider: The one where (I_n, A, B) is LLD, and the one where A is similar to $\lambda I_3 + E_{2,3}$ for some $\lambda \in \mathbb{K}$. They are dealt with separately in the next two sections.

2.4. The case when (I_n, A, B) is locally linearly dependent. In order to analyze the situation where (I_n, A, B) is LLD, we use the classification of LLD triples over fields with more than 2 elements (this result is found in [7]; prior to that, the result was known for infinite fields [2] and for fields with more than 4 elements [3]).

THEOREM 2.6 (Classification theorem for LLD triples). *Let (f, g, h) be an LLD triple of linear operators from a vector space U to a vector space V, where the underlying field has more than 2 elements. Assume that f, g, h are linearly independent and that* $\operatorname{Ker}(f) \cap \operatorname{Ker}(g) \cap \operatorname{Ker}(h) = \{0\}$ *and* $\operatorname{im}(f) + \operatorname{im}(g) + \operatorname{im}(h) = V$. *Then:*

(a) Either there is a 2-dimensional subspace $\mathcal{P}$ of $\operatorname{span}(f, g, h)$ *and a 1-dimensional subspace $\mathcal{D}$ of V such that* $\operatorname{im}(u) \subset \mathcal{D}$ *for all $u \in \mathcal{P}$;*
(b) Or $\dim V \leq 2$;

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

(c) Or $\dim U = \dim V = 3$ *and there are bases of* U *and* V *in which the operator space* $\operatorname{span}(f, g, h)$ *is represented by the space* $\mathrm{A}_3(\mathbb{K})$ *of all* 3×3 *alternating matrices.*

Corollary 2.7. *Assume that* $\#\mathbb{K} > 2$*, and let* A *and* B *be matrices of* $\mathrm{M}_n(\mathbb{K})$*, with* $n \geq 3$*, such that* (I_n, A, B) *is LLD. Then, either* I_n, A, B *are linearly dependent, or there is a* 1*-dimensional subspace* $\mathcal{D}$ *of* $\mathbb{K}^n$ *and scalars* λ *and* μ *such that* $\operatorname{im}(A - \lambda I_n) = \mathcal{D} = \operatorname{im}(B - \mu I_n)$*.*

Proof. Assume that I_n, A, B are linearly independent. As $\operatorname{Ker} I_n = \{0\}$ and $\operatorname{im} I_n = \mathbb{K}^n$, we are in the position to use Theorem 2.6. Moreover, $\operatorname{rk} I_n > 2$ discards Cases (b) and (c) altogether (as no 3×3 alternating matrix is invertible). Therefore, we have a 2-dimensional subspace $\mathcal{P}$ of $\operatorname{span}(I_n, A, B)$ and a 1-dimensional subspace $\mathcal{D}$ of $\mathbb{K}^n$ such that $\operatorname{im} M \subset \mathcal{D}$ for all $M \in \mathcal{P}$. In particular $I_n \notin \mathcal{P}$, whence $\operatorname{span}(I_n, A, B) = \mathbb{K} I_n \oplus \mathcal{P}$. This yields a pair $(\lambda, M_1) \in \mathbb{K} \times \mathcal{P}$ such that $A = \lambda I_n + M_1$, and hence, $\operatorname{im}(A - \lambda I_n) \subset \mathcal{D}$. As $A - \lambda I_n \neq 0$ (we have assumed that I_n, A, B are linearly independent), we deduce that $\operatorname{im}(A - \lambda I_n) = \mathcal{D}$. Similarly, one finds a scalar μ such that $\operatorname{im}(B - \mu I_n) = \mathcal{D}$. ∎

From there, we can prove the following result as a consequence of Theorem 1.1:

Lemma 2.8. *Assume that* $\#\mathbb{K} > 3$ *and* $n \geq 3$*. Let* $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$ *be with* $B \neq 0$*, and set* $\mathcal{H} := \{B\}^\perp$*. Assume that* (I_n, A, B) *is LLD and that* A *is not similar to* $\lambda I_3 + E_{2,3}$ *for some* $\lambda \in \mathbb{K}$*. Then,* $A \in [\mathcal{H}, \mathcal{H}]$*.*

Proof. We use a *reductio ad absurdum* by assuming that $A \notin [\mathcal{H}, \mathcal{H}]$. By Corollary 2.7, we can split the discussion into two main cases.

Case 1. I_n, A, B are linearly dependent.
Assume first that $A \in \mathbb{K} I_n$. Then, $P^{-1}AP$ is upper-triangular for every $P \in \mathrm{GL}_n(\mathbb{K})$, and hence, Lemma 2.2 yields that $P^{-1}BP$ is Hessenberg for every such P. In particular, let $x \in \mathbb{K}^n \setminus \{0\}$. For every $y \in \mathbb{K}^n \setminus \mathbb{K}x$, we can extend (x, y) into a basis $(x, y, y_3, \dots, y_n)$ of $\mathbb{K}^n$, and hence, we learn that $Bx \in \operatorname{span}(x, y)$. Using the basis $(x, y_3, y, y_4, \dots, y_n)$, we also find $Bx \in \operatorname{span}(x, y_3)$, whence $Bx \in \mathbb{K}x$. Varying x, we deduce that $B \in \mathbb{K} I_n$, whence $\mathcal{H} = \mathfrak{sl}_n(\mathbb{K})$. Theorem 1.1 then yields $A \in [\mathcal{H}, \mathcal{H}]$, contradicting our assumptions.

Assume now that $A \notin \mathbb{K} I_n$. Then, there are scalars λ and μ such that $B = \lambda A + \mu I_n$. By Theorem 1.1, there are trace zero matrices M and N such that $A = [M, N]$. Thus, $\operatorname{tr}((B - \lambda A)M) = \operatorname{tr}((B - \lambda A)N) = 0$. Using principle (2) of Section 2.1, we deduce that $(M, N) \in \mathcal{H}^2$, whence $A \in [\mathcal{H}, \mathcal{H}]$.

Case 2. I_n, A, B are linearly independent.
By Corollary 2.7, there are scalars λ and μ together with a 1-dimensional subspace $\mathcal{D}$ of $\mathbb{K}^n$ such that $\operatorname{im}(A - \lambda I_n) = \operatorname{im}(B - \mu I_n) = \mathcal{D}$. In particular, $A - \lambda I_n$ has rank

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

1, and hence, it is diagonalisable or nilpotent. In any case, A is triangularizable; in the second case, the assumption that A is not similar to $\lambda I_3 + E_{2,3}$ leads to $n \geq 4$.

Let x be an eigenvector of A. Then, we can extend x into a triple (x, y, z) of linearly independent eigenvectors of A (this uses $n \geq 4$ in the case when $A - \lambda I_n$ is nilpotent). Then, we further extend this triple into a basis $(x, y, z, y_4, \dots, y_n)$ in which $v \mapsto Av$ is upper-triangular. Point (a) in Lemma 2.2 yields $Bx \in \operatorname{span}(x, y)$. With the same line of reasoning, $Bx \in \operatorname{span}(x, z)$, and hence, $Bx \in \operatorname{span}(x, y) \cap \operatorname{span}(x, z) = \mathbb{K}x$. Thus, we have proved that every eigenvector of A is an eigenvector of B. In particular, $\operatorname{Ker}(A - \lambda I_n)$ is stable under $v \mapsto Bv$, and the resulting endomorphism is a scalar multiple of the identity. This provides us with some $\alpha \in \mathbb{K}$ such that $(B - \alpha I_n)z = 0$ for all $z \in \operatorname{Ker}(A - \lambda I_n)$. In particular, α is an eigenvalue of B with multiplicity at least $n - 1$, and since μ shares this property and $n < 2(n - 1)$, we deduce that $\alpha = \mu$. As $\operatorname{rk}(A - \lambda I_n) = \operatorname{rk}(B - \mu I_n) = 1$, we deduce that $\operatorname{Ker}(A - \lambda I_n) = \operatorname{Ker}(B - \mu I_n)$. Thus, $A - \lambda I_n$ and $B - \mu I_n$ are two rank 1 matrices with the same kernel and the same range, and hence, they are linearly dependent. This contradicts the assumption that I_n, A, B be linearly independent, thereby completing the proof. □

2.5. The case when $A = \lambda I_3 + E_{2,3}$.

LEMMA 2.9. *Assume that $\#\mathbb{K} > 2$. Let $\lambda \in \mathbb{K}$. Assume that $A := \lambda I_3 + E_{2,3}$ has trace zero. Let $B \in \mathfrak{sl}_3(\mathbb{K}) \setminus \{0\}$, and set $\mathcal{H} := \{B\}^\perp$. Then, $A \in [\mathcal{H}, \mathcal{H}]$.*

Proof. We assume that $A \notin [\mathcal{H}, \mathcal{H}]$ and search for a contradiction. By point (a) in Lemma 2.2, for every basis $\mathbf{B} = (x, y, z)$ of $\mathbb{K}^3$ for which $P_{\mathbf{B}}^{-1} A\, P_{\mathbf{B}}$ is upper-triangular, we find $Bx \in \operatorname{span}(x, y)$. In particular, for every basis (x, y) of $\operatorname{span}(e_1, e_2)$, the triple (x, y, e_3) qualifies, whence $Bx \in \operatorname{span}(x, y) = \operatorname{span}(e_1, e_2)$. It follows that $\operatorname{span}(e_1, e_2)$ is stable under B. As $z \mapsto Az$ is also represented by an upper-triangular matrix in the basis (e_2, e_3, e_1), one finds $Be_2 \in \operatorname{span}(e_2, e_3)$, whence $Be_2 \in \mathbb{K}e_2$. Thus, B has the following shape:

$$B = \begin{bmatrix} a & 0 & d \\ b & c & e \\ 0 & 0 & f \end{bmatrix}.$$

From there, we split the discussion into two main cases.

Case 1. $\lambda = 0$.
Using (e_2, e_1, e_3) as our new basis, we are reduced to the case when

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} ? & ? & ? \\ 0 & ? & ? \\ 0 & 0 & ? \end{bmatrix}.$$

Then, one checks that $[J_2, E_{2,3}] = A$, and $\operatorname{tr}(J_2 B) = 0 = \operatorname{tr}(E_{2,3} B)$. This yields

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

$A \in [\mathcal{H}, \mathcal{H}]$, contradicting our assumptions.

Case 2. $\lambda \neq 0$.
As we can replace A with $\lambda^{-1}A$, which is similar to $I_3 + E_{2,3}$, no generality is lost in assuming that $\lambda = 1$. According to principle (2) of Section 2.1, no further generality is lost in subtracting a scalar multiple of A from B, to the effect that we may assume that $f = 0$ and $B \neq 0$ (if B is a scalar multiple of A, then the same principle combined with the Albert-Muckenhoupt theorem shows that $A \in [\mathcal{H}, \mathcal{H}]$). As $\mathrm{tr} B = 0$, we find that

$$B = \begin{bmatrix} a & 0 & d \\ b & -a & e \\ 0 & 0 & 0 \end{bmatrix}.$$

Note finally that $\mathbb{K}$ must have characteristic 3 since $\mathrm{tr} A = 0$.

Subcase 2.1. $b \neq 0$.
As the problem is unchanged in multiplying B with a non-zero scalar, we can assume that $b = 1$. Assume furthermore that $d \neq 0$. Let $(\alpha, \beta) \in \mathbb{K}^2$, and set

$$C := \begin{bmatrix} 0 & 1 & 0 \\ \alpha & 0 & 1 \\ \beta & 0 & 0 \end{bmatrix}.$$

Note that C is a cyclic matrix and

$$C^2 = \begin{bmatrix} \alpha & 0 & 1 \\ \beta & \alpha & 0 \\ 0 & \beta & 0 \end{bmatrix}.$$

Thus, $\mathrm{tr}(AC) = 0$, $\mathrm{tr}(BC) = \beta d + 1$, $\mathrm{tr}(AC^2) = 2\alpha + \beta = \beta - \alpha$ and $\mathrm{tr}(BC^2) = e\beta$. As $d \neq 0$, we can set $\beta := -d^{-1}$ and $\alpha := \beta$, so that $\beta \neq 0$ and $\mathrm{tr}(A) = \mathrm{tr}(AC) = \mathrm{tr}(AC^2) = 0$. Thus, $A \in \mathrm{im}(\mathrm{ad}_C)$ by Lemma 1.7, and on the other hand $C \in \mathcal{H}$. As $A \notin [\mathcal{H}, \mathcal{H}]$, it follows that $\mathcal{C}(C) \subset \mathcal{H}$, and hence, $\mathrm{tr}(BC^2) = 0$. As $\beta \neq 0$, this yields $e = 0$.

From there, we can find a non-zero scalar t such that $d + t\,a \neq 0$ (because $\#\mathbb{K} > 2$). In the basis $(e_1, e_2, e_3 + t\,e_1)$, the respective matrices of $z \mapsto Az$ and $z \mapsto Bz$ are $I_3 + E_{2,3}$ and

$$\begin{bmatrix} a & 0 & d + t\,a \\ 1 & -a & t \\ 0 & 0 & 0 \end{bmatrix}.$$

As $d + t\,a \neq 0$ and $t \neq 0$, we find a contradiction with the above line of reasoning.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

Therefore, $d = 0$. Then, the matrices of $z \mapsto Az$ and $z \mapsto Bz$ in the basis $(e_1, e_2, e_3 + e_1)$ are, respectively, $I_3 + E_{2,3}$ and $\begin{bmatrix} a & 0 & a \\ 1 & -a & e+1 \\ 0 & 0 & 0 \end{bmatrix}$. Applying the above proof in that new situation yields $a = 0$. Therefore,

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & e \\ 0 & 0 & 0 \end{bmatrix}.$$

With $(e_3 - e\, e_1, e_1, e_2)$ as our new basis, we are finally left with the case when

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Set

$$C := \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

and note that C is cyclic and

$$C^2 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

One sees that $\operatorname{tr}(A) = \operatorname{tr}(AC) = \operatorname{tr}(AC^2) = 0$, and hence, $A \in \operatorname{im}(\operatorname{ad}_C)$ by Lemma 1.7. On the other hand, $\operatorname{tr}(BC) = 0$. As $A \notin [\mathcal{H}, \mathcal{H}]$, one should find $\operatorname{tr}(BC^2) = 0$, which is obviously false. Thus, we have a final contradiction in that case.

Subcase 2.2. $b = 0$.
Assume furthermore that $a \neq 0$. Then, in the basis $(e_1 + e_2, e_2, e_3)$, the respective matrices of $z \mapsto Az$ and $z \mapsto Bz$ are $I_3 + E_{2,3}$ and $\begin{bmatrix} a & 0 & d \\ -2a & -a & e-d \\ 0 & 0 & 0 \end{bmatrix}$. This sends us back to Subcase 2.1, which leads to another contradiction. Therefore, $a = 0$.

If $d = 0$, then we see that $B \in \operatorname{span}(I_n, A)$, and hence, principle (2) from Section 2.1 combined with Theorem 1.1 shows that $A \in [\mathcal{H}, \mathcal{H}]$, contradicting our assumptions. Thus, $d \neq 0$. Replacing the basis (e_1, e_2, e_3) with $(d\, e_1 + e\, e_2, e_2, e_3)$, we are reduced to the case when

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 27, pp. 39-54, February 2014

ELA

In that case, we set

$$C := \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

which is a cyclic matrix with

$$C^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix},$$

so that $\operatorname{tr}(A) = \operatorname{tr}(AC) = \operatorname{tr}(AC^2) = 0$ and $\operatorname{tr}(BC) = 0$. As $\operatorname{tr}(BC^2) \neq 0$, this contradicts again the assumption that $A \notin [\mathcal{H}, \mathcal{H}]$. This final contradiction shows that the initial assumption $A \notin [\mathcal{H}, \mathcal{H}]$ was wrong. $\square$

2.6. Conclusion. Let $A \in \mathrm{M}_n(\mathbb{K})$ and $B \in \mathrm{M}_n(\mathbb{K}) \setminus \{0\}$, where $n \geq 3$ and $\#\mathbb{K} \geq 4$. Set $\mathcal{H} := \{B\}^\perp$ and assume that $\operatorname{tr}(A) = 0$ and $\operatorname{tr}(B) = 0$. If A is similar to $\lambda I_3 + E_{2,3}$, then we know from Lemma 2.9 and principle (3) of Section 2.1 that $A \in [\mathcal{H}, \mathcal{H}]$. Otherwise, if (I_n, A, B) is LLD then we know from Lemma 2.8 that $A \in [\mathcal{H}, \mathcal{H}]$. Using Lemma 2.3, we conclude that $A \in [\mathcal{H}, \mathcal{H}]$ in every possible situation. This completes the proof of Theorem 1.4.

REFERENCES

[1] A.A. Albert and B. Muckenhoupt. On matrices of trace zero. *Michigan Math. J.*, 4:1–3, 1957.
[2] M. Brešar and P. Šemrl. On locally linearly dependent operators and derivations. *Trans. Amer. Math. Soc.*, 351:1257–1275, 1999.
[3] M. Chebotar and P. Šemrl. Minimal locally linearly dependent spaces of operators. *Linear Algebra Appl.*, 429:887–900, 2008.
[4] P.A. Fillmore. On similarity and the diagonal of a matrix. *Amer. Math. Monthly*, 76:167–169, 1969.
[5] T. Laffey and R. Reams. Integral similarity and commutators of integral matrices. *Linear Algebra Appl.*, 197/198:671–689, 1994.
[6] C. de Seguins Pazzis. Local linear dependence seen through duality I. Preprint, arXiv: http://arxiv.org/abs/1306.1845, 2013.
[7] C. de Seguins Pazzis. Local linear dependence seen through duality II. Preprint, arXiv: http://arxiv.org/abs/1306.5722, 2013.
[8] C. de Seguins Pazzis. To what extent is a large space of matrices not closed under product? *Linear Algebra Appl.*, 435:2708–2721, 2011.
[9] K. Shoda. Einige Sätze über Matrizen. *Jap. J. Math.*, 13:361–365, 1936.
[10] A. Stasinski. Similarity and commutators of matrices over principal ideal rings. Preprint, arXiv: http://arxiv.org/abs/1211.6872, 2012.
[11] R.C. Thompson. Matrices with zero trace. *Israel J. Math.*, 4:33–42, 1966.