On Gerstenhaber’s theorem for spaces of nilpotent matrices over a skew field

Clément de Seguins Pazzis∗†

May 3, 2014

Abstract

Let $K$ be a skew field, and $K_0$ be a subfield of the central subfield of $K$ such that $K$ has finite dimension $q$ over $K_0$. Let $V$ be a $K_0$-linear subspace of $n \times n$ nilpotent matrices with entries in $K$. We show that the dimension of $V$ is bounded above by $q \frac{n^2}{2}$, and that equality occurs if and only if $V$ is similar to the space of all $n \times n$ strictly upper-triangular matrices over $K$. This generalizes famous theorems of Gerstenhaber and Serezhkin, which cover the special case $K = K_0$.

AMS MSC: 15A03, 15A30

Keywords: nilpotent matrices, Gerstenhaber theorem, skew fields.

1 Introduction

In this article, we let $K$ be an arbitrary skew field, and $K_0$ be a subfield of the central subfield of $K$ over which $K$ has finite dimension $q$. The set $K^n$ is always endowed with its canonical structure of right-$K$-vector space. We denote by $M_{n,p}(K)$ the set of all $n \times p$ matrices with entries in $K$, endowed with its canonical structure of vector space over $K_0$. We set $M_n(K) := M_{n,n}(K)$, and denote by $\text{GL}_n(K)$ its group of invertible elements. We denote by $\text{NT}_n(K)$ the set of all strictly upper-triangular matrices of $M_n(K)$.

∗Université de Versailles Saint-Quentin-en-Yvelines, Laboratoire de Mathématiques de Versailles, 45 avenue des Etats-Unis, 78035 Versailles cedex, France
†e-mail address: dsp.prof@gmail.com
The transpose of a matrix $M$ is denoted by $M^T$, and its trace by $\text{tr}(M)$. The relation of similarity between matrices is denoted by $\simeq$ and is naturally extended to subsets of $M_n(\mathbb{K})$.

A linear subspace $\mathcal{V}$ of $M_n(\mathbb{K})$ (over $\mathbb{K}_0$) is called nilpotent when all its elements are nilpotent matrices. In that case, we note that, for every $P \in \text{GL}_n(\mathbb{K})$, the set $P \mathcal{V} P^{-1}$ is a nilpotent linear subspace of $M_n(\mathbb{K})$ with the same dimension as $\mathcal{V}$.

In his first entry in a series of four landmark papers [1], Murray Gerstenhaber studied the structure of such nilpotent subspaces. Here is his most famous result:

**Theorem 1** (Gerstenhaber, Serezhkin). Assume that $\mathbb{K}$ is commutative, and let $\mathcal{V}$ be a nilpotent linear subspace of the $\mathbb{K}$-vector space $M_n(\mathbb{K})$. Then $\dim_\mathbb{K} \mathcal{V} \leq \binom{n}{2}$, and equality occurs if and only if $\mathcal{V}$ is similar to $\text{NT}_n(\mathbb{K})$.

Our main aim here is to prove the following generalization to skew fields:

**Theorem 2.** Let $\mathcal{V}$ be a nilpotent linear subspace of $M_n(\mathbb{K})$ (over $\mathbb{K}_0$). Then:

(a) $\dim_{\mathbb{K}_0} \mathcal{V} \leq q \binom{n}{2}$.

(b) If $\dim_{\mathbb{K}_0} \mathcal{V} = q \binom{n}{2}$, then $\mathcal{V}$ is similar to $\text{NT}_n(\mathbb{K})$.

If $\mathbb{K}$ is finite (and therefore commutative), choosing $\mathbb{K}_0$ as its prime subfield yields the following corollary:

**Corollary 3.** Assume $\mathbb{K}$ is finite with cardinality $p$. Let $\mathcal{V}$ be a subgroup of $(M_n(\mathbb{K}), +)$ in which every matrix is nilpotent. Then $\# \mathcal{V} \leq p^{\binom{n}{2}}$, and equality occurs only if $\mathcal{V}$ is similar to $\text{NT}_n(\mathbb{K})$.

At the time of [1], Gerstenhaber was actually able to prove Theorem 1 only for fields with at least $n$ elements, mostly because his methods relied on the use of polynomials. A lot of progress has been made since then: we now have elementary and elegant proofs of the inequality statement that are valid for every field [3, 2], and the case of equality has been obtained for an arbitrary field by V.N. Serezhkin [7] (for fields with more than two elements, we now have a shorter proof based upon Jacobson’s generalization of Engels’s theorem, see [3]).
Recent progress on the topic must be signaled here: in [5], the inequality statement of Theorem 1 has been extended to linear subspaces of $M_n(K)$ with a trivial spectrum, i.e., which consist solely of matrices with no non-zero eigenvalue in $K$. The study of such spaces is motivated by its connection with the affine subspaces of matrices with a rank bounded below by some fixed integer. More recently [4], a classification of the linear subspaces of $M_n(K)$ with a trivial spectrum and the maximal dimension $\binom{n}{2}$ has been discovered for fields with more than two elements: for such fields, Theorem 1 appears as an easy consequence of it (see Section 5 of [4]). Finally, in [6], we have been able to prove a theorem similar to Gerstenhaber’s for linear subspaces of matrices with exactly one eigenvalue in an algebraic closure of $K$.

Both [4] and [6] are based upon a new technique which we will call the diagonal-compatibility method. The purpose of this paper is to demonstrate how this strategy can be used to obtain Theorem 2 with essentially no prior knowledge on the topic. In particular, this will yield an alternative proof of Theorem 1 (in the course of the proof, we will point out to some shortcuts for the case $K = K_0$). Note that in some cases (e.g., $K$ is commutative and separable over $K_0$), the line of reasoning of [3] may be adapted with some effort by using the trace of $K$ over $K_0$; this however fails to yield our more general theorem, so we will not use this strategy.

Our key lemma, which is proven in Section 2, is a variation of Proposition 10 of [5]. It will help us prove both points in Theorem 2 first, point (a) in Section 3 and then point (b) in the longer Section 4.

For to simplify the case $K = K_0$, we recall the following classical result, which is proven in [2, 3]. We give a simple proof of it.

**Lemma 4.** Assume that $K$ is commutative, and let $A$ and $B$ be two nilpotent matrices of $M_n(K)$ such that $A + B$ is nilpotent. Then $\text{tr}(AB) = 0$.

**Proof.** For $M = (m_{i,j})_{1 \leq i, j \leq n}$, we denote by $c_2(M)$ the coefficient in front of $t^{n-2}$ in the characteristic polynomial of $M$. Using $c_2(M) = \sum_{1 \leq i < j \leq n} m_{i,i} m_{i,j} m_{j,i} m_{j,j}$, one finds the formula

$$\forall (M, N) \in M_n(K)^2, \quad c_2(M + N) - c_2(M) - c_2(N) = \text{tr}(M) \text{tr}(N) - \text{tr}(MN). \quad (1)$$

As $A$, $B$ and $A + B$ are nilpotent, we find $\text{tr}(A) = \text{tr}(B) = 0$ and $c_2(A) = c_2(B) = c_2(A + B) = 0$, which yields $\text{tr}(AB) = 0$. \qed
2 The key lemma

Definition 1. Let \( V \) be a subset of \( M_n(\mathbb{K}) \). A vector \( X \in \mathbb{K}^n \) is called \( V \)-adapted if it is non-zero and no matrix of \( V \) has \( X \mathbb{K} \) as its column space.

Lemma 5. Let \( V \) be a subset of \( M_n(\mathbb{K}) \) which is closed under addition and contains only nilpotent matrices, and denote by \( (e_1, \ldots, e_n) \) the canonical basis of the \( \mathbb{K} \)-vector space \( \mathbb{K}^n \). Then one of the vectors \( e_1, \ldots, e_n \) is \( V \)-adapted.

The proof is largely similar to that of Proposition 10 in [5].

Proof. The result is trivial for \( n = 1 \). We use an induction, assuming, given an integer \( n \geq 2 \), that the result holds for the integer \( n - 1 \). Let \( V \) be a subset of \( M_n(\mathbb{K}) \) which is closed under addition and contains only nilpotent matrices. We assume that none of \( e_1, \ldots, e_n \) is \( V \)-adapted.

For \( (i, j) \in [1, n]^2 \), we denote by \( E_{i,j} \) the matrix of \( M_n(\mathbb{K}) \) with a zero entry everywhere except at the \( (i, j) \)-spot where the entry is 1. Denote by \( W \) the subset of \( V \) consisting of its matrices with a zero \( n \)-th row. Every \( M \in W \) may be written as

\[
M = \begin{bmatrix}
K(M) & \mathbb{I}_{(n-1)\times 1} \\
[0]_{1 \times (n-1)} & 0
\end{bmatrix}
\]

with \( K(M) \in M_{n-1}(\mathbb{K}) \), so that \( K(W) \) consists of nilpotent matrices and is obviously closed under addition. By induction, we know that there is some \( i \in \{1, n-1 \} \) such that \( e_i \) is \( K(W) \)-adapted (identifying \( \mathbb{K}^{n-1} \) with the subspace \( \mathbb{K}^{n-1} \times \{0\} \) of \( \mathbb{K}^n \) in the usual way). However, we have assumed that \( e_i \) is not \( V \)-adapted, therefore some matrix \( M \) of \( V \) has all rows zero except the \( i \)-th. Then \( M \in W \), and as \( e_i \) is \( K(W) \)-adapted, we find that \( K(M) = 0 \). Thus, \( M = a E_{i,n} \) for some \( a \in \mathbb{K} \setminus \{0\} \).

Now, the same argument may be applied to \( PVP^{-1} \) for any \( n \times n \) permutation matrix \( P \). By doing so, we find a map \( f : [1, n] \to [1, n] \) and a list \( (a_1, \ldots, a_n) \in (\mathbb{K} \setminus \{0\})^n \) such that \( V \) contains \( a_k E_{f(k),k} \) for all \( k \in [1, n] \). Let us choose a cycle for \( f \), i.e. a list \( (i_1, \ldots, i_p) \) of pairwise distinct elements of \( [1, n] \) such that \( f(i_1) = i_2, \ldots, f(i_{p-1}) = i_p \) and \( f(i_p) = i_1 \). To obtain such a cycle, one notes that some element in the sequence \( (f^i(1))_{i \geq 0} \) appears several times, to the effect that one may choose non-negative integers \( i < j \), with \( j - i \) minimal, such that \( f^i(1) = f^j(1) \); then \( (i_1, \ldots, i_p) := (f^i(1), \ldots, f^{j-1}(1)) \) is a cycle for \( f \).

Then, the matrix \( M := \sum_{k=1}^p a_{i_k} E_{f(i_k),i_k} \) belongs to \( V \) and satisfies \( M^{p} e_{i_1} = e_{i_1} \left( \prod_{k=1}^p a_{i_{p+1-k}} \right) \). This shows that \( M \) is non-nilpotent, which is a contradiction.
This *reductio ad absurdum* yields that some $e_j$ is $V$-adapted, which concludes the proof by induction.

### 3 Proving the inequality statement

Now, we use Lemma 3 to obtain point (a) of Theorem 2 just as Proposition 10 was used to obtain Theorem 9 in [5].

Again, we use an induction on $n$. The case $n = 1$ is trivial. Let $V$ be a nilpotent linear subspace of the $\mathbb{K}_0$-vector space $M_n(\mathbb{K})$. First of all, we know that some $e_i$ is $V$-adapted. Replacing $V$ with $PV^{-1}$ for a well-chosen permutation matrix $P$, we may assume that $e_n$ is $V$-adapted. In that case, we write every matrix of $V$ as

$$M = \begin{bmatrix} K(M) & C(M) \\ L(M) & a(M) \end{bmatrix},$$

where $K(M), C(M), L(M)$ are respectively $(n-1) \times (n-1), (n-1) \times 1, 1 \times (n-1)$ matrices, and $a(M) \in \mathbb{K}$. Set

$$W_1 := \{ M \in V : C(M) = 0 \}.$$

Any $M \in W_1$ is nilpotent, which yields that $a(M) = 0$ and $K(M)$ is nilpotent. Moreover, that $e_n$ is $V$-adapted yields:

$$\forall M \in W_1, K(M) = 0 \Rightarrow M = 0.$$

Using the rank theorem, one finds

$$\dim_{\mathbb{K}_0} V = \dim_{\mathbb{K}_0} K(W_1) + \dim_{\mathbb{K}_0} C(V).$$

As $K(W_1)$ is a nilpotent $\mathbb{K}_0$-linear subspace of $M_n(\mathbb{K})$ and $C(V) \subset \mathbb{K}^{n-1}$, the induction hypothesis yields

$$\dim_{\mathbb{K}_0} V \leq q \left( \frac{n-1}{2} \right) + q(n-1) = q \left( \frac{n}{2} \right).$$

Thus, point (a) of Theorem 2 is proven by induction on $n$.

### 4 Solving the case of equality

Here, we prove point (b) of Theorem 2 by induction on $n$. The case $n = 1$ is trivial.
4.1 The case $n = 2$

This case is trivial if $K = K_0$ but otherwise needs an explanation. Let $A, B$ be non-zero nilpotent matrices of $M_2(K)$ such that $A + B$ is nilpotent. Assume that $\text{Ker} \ A \neq \text{Ker} \ B$. Then $K^2 = \text{Ker} \ A \oplus \text{Ker} \ B$, and we may therefore find a basis $(f_1, f_2)$ of the $K$-vector space $K^2$ such that $f_1 \in \text{Ker} \ A$ and $f_2 \in \text{Ker} \ B$. This yields some $P \in \text{GL}_2(K)$ and some $(a, b) \in (K \setminus \{0\})^2$ such that

$$PAP^{-1} = \begin{bmatrix} 0 & a \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad PBP^{-1} = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix}.$$ 

Therefore $P(A + B)P^{-1} = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$, which is a non-singular matrix. This is a contradiction.

Now, let $V$ be a $q$-dimensional linear subspace of the $K_0$-vector space $M_2(K)$ in which every matrix is nilpotent. Choose $A \in V \setminus \{0\}$. Then we have just shown that every non-zero matrix of $V$ vanishes on $\text{Ker} \ A$. Choosing a basis $(g_1, g_2)$ of the $K$-vector space $K^2$ with $g_1 \in \text{Ker} \ A$, we find a non-singular matrix $P \in \text{GL}_2(K)$ such that every matrix of $PVP^{-1}$ has a zero first column. As $PVP^{-1}$ is nilpotent, we deduce that $PVP^{-1} \subset \text{NT}_2(K)$, and the equality of dimensions yields $PVP^{-1} = \text{NT}_2(K)$.

4.2 Setting things up for $n \geq 3$

In the rest of the proof, we assume that $n \geq 3$ and that point (b) of Theorem 2 holds for any nilpotent linear subspace of the $K_0$-vector space $M_{n-1}(K)$.

Let $V$ be a nilpotent $K_0$-linear subspace of $M_n(K)$ with dimension $q \binom{n}{2}$. Seing $V$ as a set of linear endomorphisms of the right-$K$-vector space $K^n$, what we need is to find a basis $(e'_1, \ldots, e'_{n})$ of the $K$-vector space $K^n$ in which the operators in $V$ are represented exactly by the strictly upper-triangular $n \times n$ matrices. Our method is to construct such a basis step-by-step. Equivalently, we will replace successively $V$ with similar linear subspace of matrices in order to simplify $V$ more and more, until we finally find the space $\text{NT}_n(K)$. Let us quickly lay out the sequence of choices that we will make:

- We will start by choosing the last vector $e'_n$ among the vectors that are $V$-adapted. Then we will choose a basis $(e'_1, \ldots, e'_{n-1})$ of the quotient space $K^n/(e'_nK)$ that is well-suited to $V$. Those first two operations will be done within the current section.
• At this point, each one of the vectors $e'_1, \ldots, e'_{n-1}$ will be well determined up to addition of a vector of $e'_n \mathbb{K}$.

• A reasonable choice of $e'_2, \ldots, e'_{n-1}$ will then be obtained (Section 4.3).

• A reasonable choice of $e'_1$ will come last, after a more extensive inquiry (in the end of Section 4.4).

In the rest of the proof, we denote by $(e_1, \ldots, e_n)$ the canonical basis of the $\mathbb{K}$-vector space $\mathbb{K}^n$. As in Section 3, we lose no generality in assuming that $e_n$ is $\mathcal{V}$-adapted. With the same notation as in Section 3, we deduce from the equality

\[
\dim_{\mathbb{K}_0} \mathcal{V} = q \binom{n}{2} \quad \text{and} \quad \dim_{\mathbb{K}_0} C(\mathcal{V}) = q(n-1).
\]

Set

\[\mathcal{V}_{ul} := K(\mathcal{W}_1)\]

(the subscript “ul” stands for “upper left”). Using the induction hypothesis, we deduce that:

(A) There exists $Q \in \text{GL}_{n-1}(\mathbb{K})$ such that $Q \mathcal{V}_{ul} Q^{-1} = \text{NT}_{n-1}(\mathbb{K})$.

(B) $C(\mathcal{V}) = \mathbb{K}^{n-1}$.

Setting $P_1 := Q \oplus 1$ and replacing $\mathcal{V}$ with $P_1 \mathcal{V} P_1^{-1}$ leaves conditions (A) and (B) unchanged and does not modify the assumption that $e_n$ is adapted to the space under consideration. Therefore, we may now assume, in addition to those properties:

(A’) $\mathcal{V}_{ul} = \text{NT}_{n-1}(\mathbb{K})$.

4.3 Corner-compatibility and special matrices in $\mathcal{V}$

Here, we will repeat part of the strategy of Section 4.2. Let $M \in \mathcal{V}$ and assume that $M$ vanishes on $e_2, \ldots, e_n$. Then $M \in \mathcal{W}_1$. Using $K(M) \in \text{NT}_{n-1}(\mathbb{K})$, we find $K(M) = 0$ and therefore $M = 0$. It follows that $e_1$ is $\mathcal{V}^T$-adapted.

For any $M \in \mathcal{V}$, we now write:

\[
M = \begin{bmatrix}
  b(M) & R(M) \\
  [?]_{(n-1) \times 1} & I(M)
\end{bmatrix},
\]
where $R(M)$ and $I(M)$ are respectively $1 \times (n-1)$ and $(n-1) \times (n-1)$ matrices, and $b(M) \in \mathbb{K}$. We set
\[
W_2 := \{ M \in \mathcal{V} : R(M) = 0 \},
\]
which is a nilpotent linear subspace of the $\mathbb{K}_0$-vector space $M_n(\mathbb{K})$. Thus $b(M) = 0$ for every $M \in W_2$, and $\mathcal{V}_{lr} := I(W_2)$ is a nilpotent linear subspace of the $\mathbb{K}_0$-vector space $M_{n-1}(\mathbb{K})$ (the subscript "lr" stands for "lower-right"). Finally, as $\epsilon_1$ is $\mathcal{V}^T$-adapted, we find that
\[
\forall M \in W_2, \ I(M) = 0 \Rightarrow M = 0.
\]
Using the rank theorem, we deduce that
\[
dim_{\mathbb{K}_0} \mathcal{V} = dim_{\mathbb{K}_0} \mathcal{V}_{lr} + dim_{\mathbb{K}_0} R(\mathcal{V}).
\]
As in Section 4.2, equality $dim_{\mathbb{K}_0} \mathcal{V} = q \binom{n}{2}$ and the induction hypothesis yield:

(C) There exists $Q' \in GL_{n-1}(\mathbb{K})$ such that $\mathcal{V}_{lr} = Q' \mathcal{V}_{n-1}(\mathbb{K}) (Q')^{-1}$.

We aim at modifying $\mathcal{V}$ once more so as to keep (A’) and (B) while sharpening (C).

Remark 1. In the rest of the proof, every matrix of $M_n(\mathbb{K})$ will be written as a block matrix with the following shape:
\[
\begin{bmatrix}
? & ? & ? \\
? & ? & ? \\
? & ? & ? \\
\end{bmatrix},
\]
where the question marks in the corners represent scalars.

Let us find some special matrices in $\mathcal{V}$. First of all, (A’) yields:

(D) There are $\mathbb{K}_0$-linear mappings $\varphi : M_{1,n-2}(\mathbb{K}) \to M_{1,n-2}(\mathbb{K})$ and $f : M_{1,n-2}(\mathbb{K}) \to \mathbb{K}$ such that, for every $L \in M_{1,n-2}(\mathbb{K})$, the space $\mathcal{V}$ contains
\[
A_L := \begin{bmatrix}
0 & L & 0 \\
0 & 0 & 0 \\
f(L) & \varphi(L) & 0 \\
\end{bmatrix}.
\]
Let \( C \in M_{n-2,1}(\mathbb{K}) \). By (B), we know that \( V \) contains a matrix of the form

\[
\begin{bmatrix}
? & ? & 0 \\
? & ? & C \\
? & ? & ?
\end{bmatrix}.
\]

By summing it with a matrix of type \( A_L \), we may assume furthermore that its first row has the form \([? \ 0 \ \cdots \ 0]\); in that case this row is zero as explained above. Therefore, \( V \) contains a matrix of the following form:

\[
\begin{bmatrix}
0 & 0 & 0 \\
? & ? & C' \\
? & ? & ?
\end{bmatrix}.
\]

On the other hand, we know from (A') that, for every \( U \in NT_{n-2}(\mathbb{K}) \), the subspace \( V \) contains a matrix of the form

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & U & 0 \\
? & ? & 0
\end{bmatrix}.
\]

We shall now use those observations to prove the following:

**Claim 1.** There exists a row matrix \( L \in M_{1,n-2}(\mathbb{K}) \) such that, for \( Q_1 := \begin{bmatrix} I_{n-2} & [0]_{(n-2)\times 1} \\ L & 1 \end{bmatrix} \), one has \( Q_1 V_L Q_1^{-1} = NT_{n-1}(\mathbb{K}) \).

**Proof.** Let us consider a matrix \( Q' \) given by property (C). Denote by \((e_1, \ldots, e_{n-1})\) the canonical basis of the \( \mathbb{K} \)-vector space \( \mathbb{K}^{n-1} \). Then \( V_L x \subseteq Q' \text{span}_\mathbb{K}(e_1, \ldots, e_{n-2}) \) for every \( x \in \mathbb{K}^{n-1} \). Using the matrices of type (2), we find that \( V_L e_{n-1} \) contains a \( q(n-2) \)-dimensional subspace of the \( \mathbb{K}^q \)-vector space \( \mathbb{K}^{n-1} \). Therefore \( V_L e_{n-1} = Q' \text{span}_\mathbb{K}(e_1, \ldots, e_{n-2}) \), and in particular \( V_L e_{n-1} \) is an \((n-2)\)-dimensional \( \mathbb{K} \)-linear subspace of \( \mathbb{K}^{n-1} \). Moreover, \( V_L e_{n-1} \) has a trivial intersection with \( e_{n-1} \mathbb{K} \) since every matrix of \( V \) is nilpotent. This yields a \( \mathbb{K} \)-linear

map \( u : \mathbb{K}^{n-2} \to \mathbb{K} \) such that \( V_L e_{n-1} = \{(y, u(y)) \mid y \in \mathbb{K}^{n-2}\} \). Writing \( u \) as \((y_1, \ldots, y_{n-2}) \mapsto a_1 y_1 + \cdots + a_{n-2} y_{n-2}\) for some \((a_1, \ldots, a_{n-2}) \in \mathbb{K}^{n-2}\), we set \( L := [-a_1 \ \cdots \ -a_{n-2}] \) and \( Q_1 := \begin{bmatrix} I_{n-2} & [0]_{(n-2)\times 1} \\ L & 1 \end{bmatrix} \). As \( V_L x \subseteq V_L e_{n-1} \) for every \( x \in \mathbb{K}^{n-1} \), we deduce that the last row of every matrix of \( U := Q_1 V_L Q_1^{-1} \) is zero.

We now wish to prove that \( U = NT_{n-1}(\mathbb{K}) \). First of all, any matrix \( N \) of \( U \) may be written as

\[
N = \begin{bmatrix}
T(N) & [?]_{(n-2)\times 1} \\
[0]_{1 \times (n-2)} & 0
\end{bmatrix}
\]

where \( T(N) \) is an \((n-2) \times (n-2)\)-matrix.
Then \( T(U) \) is a nilpotent linear subspace of the \( \mathbb{K}_0 \)-vector space \( M_{n-2}(\mathbb{K}) \).

With the shape of \( Q_1 \) and the matrices of type \( [3] \), we find that \( T(U) \) contains \( NT_{n-2}(\mathbb{K}) \). As \( \dim_{\mathbb{K}_0} T(U) \leq q \left( \binom{n-2}{2} \right) = \dim_{\mathbb{K}_0} NT_{n-2}(\mathbb{K}) \) by point (a) in Theorem \( [2] \), we deduce that \( T(U) = NT_{n-2}(\mathbb{K}) \). It follows that \( U \subset NT_{n-1}(\mathbb{K}) \), and the equality of dimensions over \( \mathbb{K}_0 \) then yields \( U = NT_{n-1}(\mathbb{K}) \), which finishes the proof.

With \( Q_1 \) given by Claim \( [1] \) we set \( P_2 := 1 \oplus Q_1 \) and replace \( V \) with \( P_2 V P_2^{-1} \).

Then all the preceding properties are unchanged, but we now have the improved:

\((C')\) \( V_{1r} = NT_{n-1}(\mathbb{K}) \).

Applying that property to the matrices of type \( [2] \) and \( [3] \), we find the following properties:

\((E)\) There is a \( \mathbb{K}_0 \)-linear map \( h : NT_{n-2}(\mathbb{K}) \rightarrow \mathbb{K} \) such that, for every \( U \in NT_{n-2}(\mathbb{K}), \) the space \( V \) contains the matrix

\[
E_U := \begin{bmatrix}
0 & 0 & 0 \\
0 & U & 0 \\
h(U) & 0 & 0
\end{bmatrix}.
\]

\((F)\) There are two \( \mathbb{K}_0 \)-linear maps \( \psi : M_{n-2,1}(\mathbb{K}) \rightarrow M_{n-2,1}(\mathbb{K}) \) and \( g : M_{n-2,1}(\mathbb{K}) \rightarrow \mathbb{K} \) such that, for every \( C \in M_{n-2,1}(\mathbb{K}), \) the space \( V \) contains the matrix

\[
B_C := \begin{bmatrix}
0 & 0 & 0 \\
\psi(C) & 0 & C \\
g(C) & 0 & 0
\end{bmatrix}.
\]

Finally, for every \( a \in \mathbb{K}, \) property (B) yields that \( V \) contains a matrix with entry \( a \) at the \( (1, n) \)-spot: subtracting matrices of type \( A_L \) and \( B_C \) from such a matrix yields that \( V \) contains a matrix of the form

\[
J_a = \begin{bmatrix}
? & 0 & a \\
? & ? & 0 \\
? & ? & ?
\end{bmatrix}.
\]
4.4 Analyzing \( \varphi, \psi \), and performing the last change of basis

**Claim 2.** For every \( L \in M_{1,n-2}(K) \), there exists \( a_L \in K \) such that \( \varphi(L) = a_L L \). For every \( C \in M_{n-2,1}(K) \), there exists \( b_C \in K \) such that \( \psi(C) = C b_C \).

**Proof.** Let \( (L, C) \in M_{1,n-2}(K) \times M_{n-2,1}(K) \) be such that \( LC = 0 \).

Setting \( M := A_L + B_C \), we compute
\[
M^2 = \begin{bmatrix}
L\psi(C) & 0 & 0 \\
? & ? & 0 \\
? & ? & \varphi(L)C
\end{bmatrix}.
\]

As \( M \in V \), we know that \( M^2 \) is nilpotent and therefore
\[
\varphi(L)C = 0 \quad \text{and} \quad L\psi(C) = 0.
\]

If we fix \( L \in M_{1,n-2}(K) \), varying \( C \) yields that the annihilator of the row matrix \( \varphi(L) \) contains that of \( L \), and therefore \( \varphi(L) = a_L L \) for some \( a_L \in K \). The same line of reasoning yields the second part of Claim 2. \( \square \)

**Claim 3.** There is a scalar \( \lambda \in K \) such that
\[
\forall (L, C) \in M_{1,n-2}(K) \times M_{n-2,1}(K), \quad \varphi(L) = \lambda L \quad \text{and} \quad \psi(C) = -C \lambda.
\]

**Proof.** By Claim 2 there are endomorphisms \( \varphi_1, \ldots, \varphi_{n-2} \) of the \( K_0 \)-vector space \( K \) such that
\[
\forall L = \begin{bmatrix} l_1 & \cdots & l_{n-2} \end{bmatrix} \in M_{1,n-2}(K), \quad \varphi(L) = \begin{bmatrix} \varphi_1(l_1) & \cdots & \varphi_{n-2}(l_{n-2}) \end{bmatrix}.
\]

Applying Claim 2 to the row matrices in which all the entries are equal, we find \( \varphi_1 = \cdots = \varphi_{n-2} \). As the same line of reasoning applies to \( \psi \), we obtain two endomorphisms \( u \) and \( v \) of the \( K_0 \)-vector space \( K \) such that
\[
\forall L = \begin{bmatrix} l_1 & \cdots & l_{n-2} \end{bmatrix} \in M_{1,n-2}(K), \quad \varphi(L) = \begin{bmatrix} u(l_1) & \cdots & u(l_{n-2}) \end{bmatrix}
\]
and
\[
\forall C = \begin{bmatrix} c_1 & \cdots & c_{n-2} \end{bmatrix}^T \in M_{n-2,1}(K), \quad \psi(C) = \begin{bmatrix} v(c_1) & \cdots & v(c_{n-2}) \end{bmatrix}^T.
\]

Let \( (a, b) \in K^2 \), and set \( L_0 := \begin{bmatrix} a & 0 & \cdots & 0 \end{bmatrix} \in M_{1,n-2}(K) \) and \( C_0 := \begin{bmatrix} b & 0 & \cdots & 0 \end{bmatrix}^T \in M_{n-2,1}(K) \). We notice that \( M := A_{L_0} + B_{C_0} \) stabilizes the
\(\mathbb{K}\)-subspace \(\text{span}_\mathbb{K}(e_1, e_2, e_n)\) and induces an endomorphism of it represented by \(N = \begin{bmatrix} 0 & a & 0 \\ v(b) & 0 & b \\ ? & u(a) & 0 \end{bmatrix}\). Then \(N\) is a \(3 \times 3\) nilpotent matrix, and therefore \(N^3 = 0\).

One computes that the entry of \(N^3\) at the \((1, 2)\)-spot is \(a(v(b)a + bu(a))\). For \(a \neq 0\), this yields
\[v(b) a + b u(a) = 0,\]which is also obviously true for \(a = 0\).

Set now \(\lambda := u(1)\). Taking \(a = 1\) in (4) yields:
\[v(b) = -b \lambda \]for all \(b \in \mathbb{K}\). Thus, \(v(1) = -\lambda\), and taking \(b = 1\) in (4) yields \(u(a) = \lambda a\) for all \(a \in \mathbb{K}\). This finishes the proof of Claim 3.

Remark 2. In the case \(\mathbb{K} = \mathbb{K}_0\), Claim 3 has a far more simple proof. Indeed, Claim 2 then readily yields a pair \((\lambda, \mu) \in \mathbb{K}^2\) such that \(\forall (L, C) \in M_{1, n-2}(\mathbb{K}) \times M_{n-2, 1}(\mathbb{K})\), \(\varphi(L) = \lambda L\) and \(\psi(C) = \mu C\); as \(\mathbb{K}\) is commutative, we find \(\text{tr}(A_L B_C) = 0\) for every \((L, C) \in M_{1, n-2}(\mathbb{K}) \times M_{n-2, 1}(\mathbb{K})\), and hence \(\mu + \lambda = 0\).

Now, we perform one last change of basis. We set \(P := \begin{bmatrix} 1 & 0 & 0 \\ 0 & I_{n-2} & 0 \\ -\lambda & 0 & 1 \end{bmatrix} \in \text{GL}_n(\mathbb{K})\) and we replace \(V\) with \(P V P^{-1}\). Note then that all properties (A'), (B), (C'), (D), (E) and (F) still hold, but we now have a simplified form for the matrices of type \(A_L\) and \(B_C\):
\[
\forall (L, C) \in M_{1, n-2}(\mathbb{K}) \times M_{n-2, 1}(\mathbb{K})\), \(A_L = \begin{bmatrix} 0 & L & 0 \\ 0 & 0 & 0 \\ f(L) & 0 & 0 \end{bmatrix}\) and \(B_C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & C \\ g(C) & 0 & 0 \end{bmatrix}\).

From there, our aim is to prove that \(\mathcal{V} = \mathcal{N} T_n(\mathbb{K})\). In order to do so, we will show that all the matrices of type \(A_L\), \(B_C\), \(E_U\) and \(J_a\) are strictly upper-triangular. This will prove the inclusion \(\mathcal{N} T_n(\mathbb{K}) \subset \mathcal{V}\), and the equality of dimensions over \(\mathbb{K}_0\) will help us complete the proof. We start by showing that \(f\) and \(g\) vanish everywhere.

### 4.5 The vanishing of \(f\) and \(g\)

**Claim 4.** One has \(f = 0\) and \(g = 0\).
Proof. We claim that
\[
\forall (L, C) \in M_{1,n-2}(K) \times M_{n-2,1}(K), \quad LC \neq 0 \Rightarrow f(L) + g(C) = 0.
\] (5)
Let indeed \((L, C) \in M_{1,n-2}(K) \times M_{n-2,1}(K)\) be such that \(LC \neq 0\); setting \(M := A_L + B_C\), we compute \(M^3 e_1 = e_1 (LC(f(L) + g(C)))\) and [5] follows as \(M^3\) is nilpotent.

- Assume that \(n \geq 4\). Let \(L \in M_{1,n-2}(K)\). As \(n - 2 \geq 2\), we may choose \(C \in M_{n-2,1}(K) \setminus \{0\}\) such that \(LC = 0\), and then we may choose \(L_1 \in M_{1,n-2}(K)\) such that \(L_1 C = 1\). Then \((L + L_1)C = 1\), which yields \(f(L + L_1) = -g(C) = f(L_1)\). Thus, \(f(L) = 0\). The same line of reasoning yields \(g = 0\).

- Assume that \(n = 3\) and \(\#K > 2\). Let \(x \in K\). Then we may choose \(y \in K \setminus \{0, -x\}\), so that \(y \neq 0\) and \(x + y \neq 0\). Therefore, \(f(x + y) = -g(1) = f(y)\), and hence \(f(x) = 0\). The same line of reasoning yields \(g = 0\).

- Assume finally that \(n = 3\) and \(\#K = 2\), so that \(K_0 = K \simeq F_2\). Then, \(f(1) = g(1)\). Assume that \(f(1) = 1\). Then \(V\) contains the matrices

\[
A := \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}
\]

and a matrix of the form

\[
J = \begin{bmatrix} a & 0 & 1 \\ b & c & 0 \\ d & e & f \end{bmatrix}.
\]

Note that \(K\) is commutative, thus Lemma 4 yields \(\text{tr}(AJ) = \text{tr}(BJ) = 0\), and hence \(b = e = 1\). As \(J\) is nilpotent, we also have \(\text{tr}(J) = 0\), and hence \(f = a + c\). Using \(\forall t \in K, \ t^2 = t\) and \(2t = 0\), we finally compute:

\[
\forall (x, y) \in K^2, \quad 0 = \det(J + xA + yB) = 1 + cd + (a + c) y + a x + d xy.
\]

This yields both \(cd = 1\) and \(d = 0\), a contradiction.

Therefore, \(f(1) = g(1) = 0\), and so \(f = 0\) and \(g = 0\), as claimed. \(\square\)
4.6 The vanishing of \( h \)

**Claim 5.** One has \( h = 0 \).

**Proof.** Let \( U \in \text{NT}_{n-2}(\mathbb{K}) \) be such that \( U^2 = 0 \). Set \( L_0 := \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \in \text{M}_{1,n-2}(\mathbb{K}) \) and \( C_0 := L_0^T \), so that \( L_0 U C_0 = 0 \) and \( L_0 C_0 = 1 \). Setting \( M := A L_0 + B C_0 + EU \), one checks that \( M^3 e_n = e_n h(U) \), and therefore \( h(U) = 0 \).

In particular, \( h(E_{i,j} a) = 0 \) for every \( a \in \mathbb{K} \) and every \((i, j) \in [1, n-2]^2\) with \( j > i \) (where \( E_{i,j} \) is the matrix with all entries zero except at the \((i, j)\)-spot where the entry is 1). As \( h \) is additive, we deduce that \( h \) vanishes everywhere on \( \text{NT}_{n-2}(\mathbb{K}) \). \( \square \)

4.7 The matrices of type \( J_a \)

4.7.1 Simplifying the \( J_a \) matrices

Let us sum up. For every triple \((L, C, U) \in \text{M}_{1,n-2}(\mathbb{K}) \times \text{M}_{n-2,1}(\mathbb{K}) \times \text{NT}_{n-2}(\mathbb{K})\), the space \( \mathcal{V} \) contains the matrices

\[
A_L = \begin{bmatrix} 0 & L & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & C \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad E_U = \begin{bmatrix} 0 & 0 & 0 \\ 0 & U & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\]

Adding an appropriate \( E_U \) to each matrix of type \( J_a \), one finds \( \mathbb{K}_0 \)-linear maps \( \alpha : \mathbb{K} \rightarrow \mathbb{K}, \beta : \mathbb{K} \rightarrow \mathbb{K}, \gamma : \mathbb{K} \rightarrow \mathbb{K}, L_1 : \mathbb{K} \rightarrow \text{M}_{1,n-2}(\mathbb{K}), C_1 : \mathbb{K} \rightarrow \text{M}_{n-2,1}(\mathbb{K}), T : \mathbb{K} \rightarrow \text{LT}_{n-2}(\mathbb{K}) \) (where \( \text{LT}_{n-2}(\mathbb{K}) \) denotes the set of lower-triangular matrices of \( \text{M}_{n-2}(\mathbb{K}) \)) such that, for every \( a \in \mathbb{K} \), the subspace \( \mathcal{V} \) contains

\[
J_a := \begin{bmatrix} \alpha(a) & 0 & a \\ C_1(a) & T(a) & 0 \\ \beta(a) & L_1(a) & \gamma(a) \end{bmatrix}.
\]

Our aim in what follows is to prove:

**Claim 6.** All the maps \( \alpha, \beta, \gamma, L_1, C_1 \) and \( T \) vanish everywhere on \( \mathbb{K} \).

We have to distinguish between two cases, the main problem being the handling of fields with two elements.

14
4.7.2 Proof of Claim 6: the case $\mathbb{K} = \mathbb{K}_0$

We assume $\mathbb{K} = \mathbb{K}_0$. In particular, $\mathbb{K}$ is commutative, which allows us to use Lemma 4 to obtain $\text{tr}(J_1A_L) = 0$, $\text{tr}(J_1B_C) = 0$ and $\text{tr}(J_1E_U) = 0$ for all $(L, C, U) \in M_{1,n-2}(\mathbb{K}) \times M_{n-2,1}(\mathbb{K}) \times NT_{n-2}(\mathbb{K})$. Therefore, $L_1(1) = 0$, $C_1(1) = 0$ and $T(1)$ is a diagonal matrix. Every diagonal entry of $T(1)$ is an eigenvalue of $J_1$, and hence $T(1) = 0$. Then $J_1$ induces an endomorphism of $\text{span}_\mathbb{K}(e_1, e_n)$ whose matrix in $(e_1, e_n)$ is $N = \begin{bmatrix} \alpha(1) & 1 \\ \beta(1) & \gamma(1) \end{bmatrix}$. This last matrix must be nilpotent, and hence $\alpha(1) = -\gamma(1)$ and $\beta(1) = -\gamma(1)^2$ (as $\text{tr} N = 0$ and $\text{det} N = 0$). Choose finally $(L, C) \in M_{1,n-2}(\mathbb{K}) \times M_{n-2,1}(\mathbb{K})$ such that $LC \neq 0$, and set $M := J_1 + A_L + B_C$. One checks that $M^3e_1 = -\gamma(1)^2LC e_1$, and hence $\gamma(1) = 0$. Therefore, the maps $\alpha$, $\beta$, $\gamma$, $L_1$, $C_1$ and $T$ all vanish on 1; since they are $\mathbb{K}$-linear, Claim 6 is proven in the case $\mathbb{K} = \mathbb{K}_0$.

4.7.3 Proof of Claim 6: the case $\#\mathbb{K} > 2$

We assume here that $\#\mathbb{K} > 2$, which holds whenever $\mathbb{K}_0 \subseteq \mathbb{K}$.

Fix $a \in \mathbb{K}$. Let $C_0 \in M_{n-2,1}(\mathbb{K}) \setminus \{0\}$. Let $x \in \mathbb{K}$. We consider the non-zero vector $X := \begin{bmatrix} x \\ C_0 \\ 0 \end{bmatrix}$ of $\mathbb{K}^n$. The $\mathbb{K}_0$-vector space $\mathcal{V}X$ must intersect $X \mathbb{K}$ trivially as all the elements of $\mathcal{V}$ are nilpotent. Thus $\dim_{\mathbb{K}_0} \mathcal{V}X \leq (n-1)q$. However, for every $(L, C) \in M_{1,n-2}(\mathbb{K}) \times M_{n-2,1}(\mathbb{K})$, we have

$$A_LX = \begin{bmatrix} LC_0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad B_CX = \begin{bmatrix} 0 \\ C \\ 0 \end{bmatrix}$$

Varying $L$ and $C$ then yields the inclusion $\mathbb{K}^{n-1} \times \{0\} \subset \mathcal{V}X$. Since $\dim_{\mathbb{K}_0} \mathcal{V}X \leq (n-1)q = \dim_{\mathbb{K}_0}(\mathbb{K}^{n-1} \times \{0\})$, we deduce that $\mathcal{V}X = \mathbb{K}^{n-1} \times \{0\}$. However, the last entry of $J_nX$ is $\beta(a)x + L_1(a)C_0 + \gamma(a)$, and therefore:

$$\forall x \in \mathbb{K}, \quad \beta(a)x + L_1(a)C_0 + \gamma(a) = 0.$$

We deduce that $L_1(a)C_0 + \gamma(a) = 0$ and $\beta(a) = 0$, which yields:

$$\forall C \in M_{n-2,1}(\mathbb{K}) \setminus \{0\}, \forall y \in \mathbb{K} \setminus \{0\}, \quad L_1(a)Cy + \gamma(a) = 0.$$

As $\#\mathbb{K} > 2$, we deduce that $\gamma(a) = 0$ and

$$\forall C \in M_{n-2,1}(\mathbb{K}) \setminus \{0\}, \quad L_1(a)C = 0.$$
Varying $C$ then yields $L_1(a) = 0$.

Let again $C_0 \in M_{1,n-2}(\mathbb{K}) \setminus \{0\}$, and set $Y := \begin{bmatrix} 1 \\ C_0 \\ 0 \end{bmatrix}$. For every $(L,C) \in M_{1,n-2}(\mathbb{K}) \times M_{n-2,1}(\mathbb{K})$, we have

$$A_L^T Y = \begin{bmatrix} 0 \\ L^T \\ 0 \end{bmatrix} \quad \text{and} \quad B_C^T X = \begin{bmatrix} 0 \\ 0 \\ C^T C_0 \end{bmatrix}.$$

As above, varying $C$ and $L$ yields $V^T Y = \{0\} \times \mathbb{K}^{n-1}$. The first entry of $J_a^T Y$ is $\alpha(a) + C_1(a)^T C_0$ and it must be 0. Again, varying $C_0$ yields both $\alpha(a) = 0$ and $C_1(a) = 0$.

Let $U \in \text{NT}_{n-2}(\mathbb{K})$. For every $t \in \mathbb{K}_0$, the matrix $E_U + tJ_a$ is nilpotent and stabilizes the $\mathbb{K}$-vector space span$_{\mathbb{K}}(e_2,\ldots,e_{n-1})$, with an induced endomorphism represented in $(e_2,\ldots,e_{n-1})$ by $U + tT(a)$. It follows that $\text{NT}_{n-2}(\mathbb{K}) + \mathbb{K}_0 T(a)$ is a nilpotent $\mathbb{K}_0$-linear subspace of $M_{n-2}(\mathbb{K})$. If $T(a) \neq 0$, then we have a contradiction with point (a) of Theorem 2. Therefore $T(a) = 0$, and Claim 6 is proven.

### 4.8 Conclusion

We have shown that, for every list $(L,C,U,a) \in M_{1,n-2}(\mathbb{K}) \times M_{n-2,1}(\mathbb{K}) \times \text{NT}_{n-2}(\mathbb{K}) \times \mathbb{K}$, the additive group $V$ contains all four matrices

$$\begin{bmatrix} 0 & L & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & C \\ 0 & U & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

It follows that $V$ contains $\text{NT}_{n}(\mathbb{K})$. As $\dim_{\mathbb{K}_0} V = q \binom{n}{2} = \dim_{\mathbb{K}_0} \text{NT}_{n}(\mathbb{K})$, we conclude that $V = \text{NT}_{n}(\mathbb{K})$. This completes our proof of point (b) of Theorem 2.

### Acknowledgement

The author would like to thank Alexander Guterman for his outstanding effort in helping him understand the essence of Serezhkin’s proof.
References

[1] M. Gerstenhaber, On nilalgebras and linear varieties of nilpotent matrices (I), *Amer. J. Math.* **80** (1958) 614-622.

[2] G.W. MacDonald, J.A. MacDougall, L.G. Sweet, On the dimension of linear spaces of nilpotent matrices, *Linear Algebra Appl.* **437** (2012) 2210-2230.

[3] B. Mathes, M. Omladič, H. Radjavi, Linear spaces of nilpotent matrices, *Linear Algebra Appl.* **149** (1991) 215-225.

[4] C. de Seguins Pazzis, Large affine spaces of non-singular matrices, *Trans. Amer. Math. Soc.* **365** (2013) 2569-2596.

[5] C. de Seguins Pazzis, On the matrices of given rank in a large subspace, *Linear Algebra Appl.* **435-1** (2011) 147-151.

[6] C. de Seguins Pazzis, Spaces of matrices with a sole eigenvalue, *Lin. Multilin. Alg.* **60-10** (2012) 1165-1190.

[7] V.N. Serezhkin, Linear transformations preserving nilpotency (in Russian), *Izv. Akad. Nauk BSSR, Ser. Fiz.-Mat. Nauk* **125** (1985) 46-50.