ASSOCIATIVE NIL-ALGEBRAS OVER FINITE FIELDS.

ARTEM A. LOPATIN AND IVAN P. SHESTAKOV

ABSTRACT. The nilpotency degree of a relatively free finitely generated associative algebra with the identity $x^n = 0$ is studied over finite fields.

Keywords: Nil-algebras, nilpotency degree, Nagata–Higman Theorem, PI-algebras, associative algebras, finite fields.
2010 MSC: 16R10, 16R40, 16N40, 11T06.

1. INTRODUCTION

We denote by $\langle X \rangle_d \subset \langle X \rangle$ the semigroups (without unity) freely generated by *letters* x_1, \ldots, x_d and x_1, x_2, \ldots , respectively. Elements of $\langle X \rangle$ are called *words*. Let $\mathbb{F}\langle X \rangle_d$ and $\mathbb{F}\langle X \rangle$ be the vector spaces over the field \mathbb{F} with the bases $\langle X \rangle_d$ and $\langle X \rangle$, respectively. Note that elements of $\mathbb{F}\langle X \rangle_d$ and $\mathbb{F}\langle X \rangle$ are finite linear combinations of words from $\langle X \rangle_d$ and $\langle X \rangle$, respectively. Denote by

$$N_{n,d} = \frac{\mathbb{F}\langle X \rangle_d}{\mathrm{id}\{x^n \mid x \in \mathbb{F}\langle X \rangle_d\}}$$

the relatively free finitely generated associative algebra with the identity $x^n = 0$. The connection between this algebra and analogues of the Burnside problems for associative algebras suggested by Kurosh and Levitzky is discussed in recent survey [19] by Zelmanov. The algebra $N_{n,d}$ also plays a crucial role in the construction of minimal systems of generators for algebras of polynomial invariants of several matrices (see [3], [9], [10], [11], [12]). It is well-known that $N_{n,d}$ is a nilpotent algebra. For example, it follows from the Shirshov Height Theorem [17] and the fact that

(1)
$$L_{1^n}(a_1,\ldots,a_n) = \sum_{\sigma \in S_n} a_{\sigma(1)} \cdots a_{\sigma(n)},$$

the complete linearization of x^n , is equal to zero in $N_{n,d}$ (see below for the details). We write

$$C_{n,d} = \min\{c > 0 \mid a_1 \cdots a_c = 0 \text{ for all } a_1, \dots, a_c \in N_{n,d}\}$$

for the *nilpotency* degree of $N_{n,d}$. Since $C_{1,d} = 1$ and $C_{n,1} = n$, we assume that $n, d \geq 2$ unless otherwise stated. To specify the field \mathbb{F} , we write $N_{n,d}^{\mathbb{F}}$, $C_{n,d}^{\mathbb{F}}$ for $N_{n,d}$, $C_{n,d}$, respectively. We say that an element f of $\mathbb{F}\langle X \rangle_d$ is a *relation* for $N_{n,d}$ and write f = 0 in $N_{n,d}$ if the image of f in $N_{n,d}$ is zero.

In characteristic zero case of the field \mathbb{F} we have $\frac{1}{2}n(n+1) \leq C_{n,d} \leq n^2$, where the lower bound was established by Kuzmin [7] and the upper bound was given by Razmyslov [14]. Kuzmin also conjectured that $C_{n,d} = \frac{1}{2}n(n+1)$, which was shown to be true for $n \leq 4$ by Vaughan–Lee [18]. The case of n = 5 and d = 2 was considered by Shestakov and Zhukavets [15]. An English translation of Kuzmin's result can be found in books [1] and [5].

The difference of the case of positive characteristic of \mathbb{F} from the case of zero characteristic is due to the fact that in the latter case any relation for $N_{n,d}$ belongs to the ideal generated by L_{1^n} , which in general is not the case in positive characteristic. As the result, in case $0 < \operatorname{char} \mathbb{F} \leq n$ we have $C_{n,d} \to \infty$ as $d \to \infty$ by [3]. Recently, Belov and Kharitonov [2] established the that $C_{n,d} \leq 2^{18} \cdot n^{12\log_3(n)+28} d$. In the case of an infinite field \mathbb{F} with char $\mathbb{F} > \frac{n}{2}$ Lopatin [13] proved that $C_{n,d} < 4 \cdot 2^{\frac{n}{2}} d$. See Remark 4.8 of [13] for the comparison of these two upper bounds.

Given a field \mathbb{F} of characteristic $p \geq 0$, the nilpotency degree $C_{n,d}$ is known for n = 2 for an arbitrary \mathbb{F} (for example, see [3]) and for n = 3 for an infinite \mathbb{F} (see [8]) and [9]):

$$C_{2,d} = \begin{cases} 3, & \text{if } p = 0 \text{ or } p > 2\\ d+1, & \text{if } p = 2 \end{cases} \quad \text{and} \quad C_{3,d} = \begin{cases} 6, & \text{if } p = 0 \text{ or } p > 3\\ 6, & \text{if } p = 2 \text{ and } d = 2\\ d+3, & \text{if } p = 2 \text{ and } d > 2\\ 3d+1, & \text{if } p = 3. \end{cases}$$

Moreover, in case \mathbb{F} is infinite Lopatin explicitly described an \mathbb{F} -basis for $N_{3,d}$ (see [9]) and calculated $C_{4,d}$ with deviation three for all d under assumption that $p \neq 2$ (see [13]).

In the case of an infinite field \mathbb{F} all partial linearizations of x^n are relations for $N_{n,d}$, which in general does not hold in the case of a finite field. For an arbitrary field of characteristic p Eick [6] obtained the following results:

- if p > n, then $C_{n,d}$ is the same as in the case of an infinite field of the same characteristic;
- $C_{3,d}$ is computed for p = 2, 3 and $d \le 4$;
- $C_{4,2}$ is calculated for p = 2, 3, 5 and $C_{5,2}$ is calculated for p = 0, 2, 3, 5, 7.

In this paper we extend the mentioned results as follows. Let \mathbb{F} be an arbitrary field and \mathbb{L} be an infinite field with char $\mathbb{F} = \operatorname{char} \mathbb{L}$. Then

- in case $\#\mathbb{F} \ge n$ we have $C_{n,d}^{\mathbb{F}} = C_{n,d}^{\mathbb{L}}$ (see Corollary 2.9);
- we completed the description of $C_{3,d}^{\mathbb{F}}$ for all d by proving that $C_{3,d}^{\mathbb{F}} = C_{3,d}^{\mathbb{L}}$ (see Section 3). We also explicitly described an \mathbb{F} -basis for $N_{3,d}^{\mathbb{F}}$ (see Remark 3.7). Note that

;

$$\dim N_{3,d}^{\mathbb{F}} = \begin{cases} \dim N_{3,d}^{\mathbb{L}} + d(d-1)/2, & \text{if } \#\mathbb{F} = 2\\ \dim N_{3,d}^{\mathbb{L}}, & \text{otherwise} \end{cases}$$

- in case #F = n − 1 we have C^L_{n,d} ≤ C^F_{n,d} ≤ C^L_{n,d} + 1 (see Theorem 4.1);
 as a consequence, C_{4,d} is described with deviation four for all d under assumption that $p \neq 2$ (see Corollary 4.5).

The following conjecture holds for n = 2, 3 and all d, for n = 4, 5 and d = 2.

Conjecture 1.1. For a finite field \mathbb{F} and an infinite field \mathbb{L} with char $\mathbb{F} = \operatorname{char} \mathbb{L}$ the equality $C_{n,d}^{\mathbb{F}} = C_{n,d}^{\mathbb{L}}$ holds.

2. General Case

We start with some notations. Let $\mathbb{N} = \{1, 2, ...\}, \mathbb{N}_0 = \mathbb{N} \sqcup \{0\}, \mathbb{Z}$ be the ring of integers, and $\mathbb{F}^* = \mathbb{F} \setminus \{0\}$. For an $a \in \langle X \rangle_d$ and a letter x we denote by $\deg_x(a)$ the degree of a in the letter x, by $mdeg(a) = (deg_{x_1}(a), \ldots, deg_{x_d}(a))$ the multidegree of a, and by $\deg(a) = \deg_{x_1}(a) + \cdots + \deg_{x_d}(a)$ the *degree* of a. As usually, the degree of $f = \sum_{i} \alpha_{i} a_{i} \in \mathbb{F}\langle X \rangle_{d}$ is the maximum of degrees of words a_{i} . For short, we write 1^n for multidegree $(1, \ldots, 1)$ (*n* times). Given $\underline{\alpha} = (\alpha_1, \ldots, \alpha_r) \in \mathbb{N}_0^r$, we set $\#\underline{\alpha} = r$, $|\underline{\alpha}| = \alpha_1 + \cdots + \alpha_r$. Given a prime *p*, denote by \mathbb{F}_p the field with *p* entries. In case p = 0 we set $\mathbb{F}_p = \mathbb{Z}$.

We define $\langle X \rangle_d^{\#} = \langle X \rangle_d \sqcup \{1\}$ and $\langle X \rangle^{\#} = \langle X \rangle \sqcup \{1\}$. Given $\underline{\theta} \in \mathbb{N}_0^r$ with $|\underline{\theta}| = n$, denote by $L_{\underline{\theta}}(\underline{x}) = L_{\underline{\theta}}(x_1, \ldots, x_r) \in \mathbb{F}\langle X \rangle$ the coefficient of $\alpha_1^{\theta_1} \cdots \alpha_r^{\theta_r}$ in $(\alpha_1 x_1 + \cdots + \alpha_r x_r)^n$, where $\alpha_i \in \mathbb{F}$. Note that $L_{\underline{\theta}}(\underline{x})$ is a non-zero element of $\mathbb{F}\langle X \rangle$. We say that $L_{\underline{\theta}}(\underline{x})$ is the *partial linearization* of x^n of multidegree $\underline{\theta}$. As an example, see above formula (1) for L_{1^n} .

Remark 2.1. If \mathbb{F} is infinite, then by the standard Vandermonde arguments (or see the proof of part (a) of Theorem 2.5 below) we obtain that the ideal of relations for $N_{n,d}$ is generated by $L_{\underline{\theta}}(\underline{a})$ for $|\underline{\theta}| = n$ and $a_i \in \langle X \rangle_d$ for all i.

Remark 2.2. Assume that the characteristic of \mathbb{F} is $p \ge 0$. Let $A = \mathbb{F}\langle X \rangle_d / I$ for the ideal I generated by some polynomials from $\mathbb{F}_p \langle X \rangle_d$. Consider a basis B of A consisting of words. Then in the algebra A we have

- (a) every $w \in \langle X \rangle_d$ belongs to the \mathbb{F}_p -span of B in A;
- (b) if $w_1, \ldots, w_r \in \langle X \rangle_d$ are linearly dependent in A, then $\sum_i \alpha_i w_i = 0$ in A for such $\alpha_i \in \mathbb{F}_p$ that not all of them are equal to zero.

Remarks 2.1 and 2.2 imply that if \mathbb{F} and \mathbb{L} are infinite fields and char $\mathbb{F} = \operatorname{char} \mathbb{L}$, then

- every basis for $N_{n,d}^{\mathbb{F}}$ consisting of words is a basis for $N_{n,d}^{\mathbb{L}}$;
- $C_{n,d}^{\mathbb{F}} = C_{n,d}^{\mathbb{L}}$.

Definition 2.3 (of Frobenius partial linearization). Assume that \mathbb{F} is finite and $q = \#\mathbb{F}$. Given $\underline{\delta} \in \mathbb{N}^r$ with $|\underline{\delta}| = n$, we say that the Frobenius partial linearization of x^n of multidegree $\underline{\delta}$ is

$$F_{\underline{\delta}}(\underline{x}) = F_{\underline{\delta}}(x_1, \dots, x_r) = \sum L_{\underline{\theta}}(x_1, \dots, x_r) \in \mathbb{F}\langle X \rangle,$$

where $\underline{\theta}$ ranges over vectors from \mathbb{N}^r satisfying

- $|\underline{\theta}| = n;$
- $\theta_i \equiv \delta_i \pmod{q-1}$ for all *i*.

As an example, if n = 5 and $\#\mathbb{F} = 3$, then $F_{41}(\underline{x}) = L_{41}(\underline{x}) + L_{23}(\underline{x})$, $F_{32}(\underline{x}) = L_{32}(\underline{x}) + L_{14}(\underline{x})$, and $F_{311}(\underline{x}) = L_{311}(\underline{x}) + L_{131}(\underline{x}) + L_{113}(\underline{x})$. Note that $L_{\underline{\theta}}(x_1, \ldots, x_r) = L_{\sigma\underline{\theta}}(x_{\sigma(1)}, \ldots, x_{\sigma(r)})$ for all permutations $\sigma \in S_r$, where $\sigma\underline{\theta}$ stands for $(\theta_{\sigma(1)}, \ldots, \theta_{\sigma(r)})$. Similar remark holds for $F_{\underline{\delta}}(\underline{x})$. We will use these remarks without references to them.

A vector $\underline{\theta} \in \mathbb{N}^r$ is called *ordered* if $\theta_1 \geq \cdots \geq \theta_r$. A vector $\underline{\delta} \in \mathbb{N}^r$ is called *q*-maximal if $\underline{\delta} \geq \underline{\theta}$ for every $\underline{\theta} \in \mathbb{N}^r$ with $|\underline{\theta}| = |\underline{\delta}|$ and $\theta_i \equiv \delta_i \pmod{q-1}$ for all *i*, where \geq stands for the usual lexicographical order on \mathbb{N}^r .

Remark 2.4. For $q = \#\mathbb{F}$ we have that

- (a) if $q \ge n$, then $F_{\underline{\delta}}(\underline{x}) = L_{\underline{\delta}}(\underline{x})$;
- (b) every word $a \in \langle X \rangle_r$ of degree n, satisfying $\deg_{x_i}(a) > 0$ for all $1 \le i \le r$, is a summand of $F_{\underline{\delta}}(\underline{x})$ for one and only one q-maximal vector $\underline{\delta} \in \mathbb{N}^r$ with $|\underline{\delta}| = n$;

(c) for every $\underline{\delta} \in \mathbb{N}^r$ there exists a *q*-maximal $\underline{\nu} \in \mathbb{N}^r$ satisfying $F_{\underline{\delta}}(\underline{x}) = F_{\underline{\nu}}(\underline{x})$ and $|\underline{\delta}| = |\underline{\nu}|$.

Theorem 2.5. Assume \mathbb{F} is finite and $q = \#\mathbb{F}$. Then the ideal of relations for $N_{n,d}$ is generated as a vector space over \mathbb{F} by $uf(a_1, \ldots, a_r)v$ for all $a_1, \ldots, a_r \in \langle X \rangle_d$, $u, v \in \langle X \rangle_d^{\#}$ and $f \in S$, where the set S is defined as follows:

- (a) if $q \ge n$, then S consists of $L_{\underline{\theta}}(\underline{x})$ for all ordered $\underline{\theta}$ with $|\underline{\theta}| = n$;
- (b) if q < n, then S consists of $F_{\delta}(\underline{x})$ for all ordered q-maximal $\underline{\delta}$ with $|\underline{\delta}| = n$.

We split the proof of the theorem into several lemmas. Given $1 \le r \le n$, denote

$$P_r(\underline{x}) = P_r(x_1, \dots, x_r) = \sum L_{\underline{\theta}}(x_1, \dots, x_r),$$

where the sum ranges over all $\underline{\theta} \in \mathbb{N}^r$ with $|\underline{\theta}| = n$.

Lemma 2.6. We have $P_r(a_1, \ldots, a_r) = 0$ in $N_{n,d}$ for all $1 \le r \le n$ and $a_1, \ldots, a_r \in \mathbb{F}\langle X \rangle_d$.

Proof. We prove by the inducton on r. If r = 1, then $P_1(a_1) = a_1^n = 0$ in $N_{n,d}$ for all $a_1 \in \mathbb{F}\langle X \rangle_d$.

Assume that for all $1 \leq k < r$ we have $P_k(a_1, \ldots, a_k) = 0$ in $N_{n,d}$ for all $a_1, \ldots, a_k \in \mathbb{F}\langle X \rangle_d$. Since $(a_1 + \cdots + a_r)^n$ is equal to

$$P_r(a_1,...,a_r) + \sum_{1 \le k \le r-1} \sum P_k(a_{i_1},...,a_{i_k}),$$

where the second sum ranges over all $1 \le i_1 < \cdots < i_k \le r$, the induction hypothesis completes the proof.

Lemma 2.7. Let \mathbb{F} be finite and $q = \#\mathbb{F}$. Then $F_{\underline{\delta}}(\underline{a}) = 0$ in $N_{n,d}$ for every $\underline{\delta}$ with $|\underline{\delta}| = n$ and $a_i \in \mathbb{F}\langle X \rangle_d$ for all i. In particular, in case $q \ge n$ we have $L_{\underline{\theta}}(\underline{a}) = 0$ in $N_{n,d}$ for all $\underline{\theta}$ with $|\underline{\theta}| = n$ and $a_i \in \mathbb{F}\langle X \rangle_d$.

Proof. By part (c) of Remark 2.4, without loss of generality we can assume that $\underline{\delta} \in \mathbb{N}^r$ is a *q*-maximal vector, where r > 0. Let $\{\underline{\delta} = \underline{\delta}^1, \ldots, \underline{\delta}^k\}$ be the set of all pairwise different *q*-maximal vectors of \mathbb{N}^r with $|\underline{\delta}^i| = n$. By part (b) of Remark 2.4, we obtain

$$P_r(\underline{x}) = F_{\underline{\delta}^1}(\underline{x}) + \dots + F_{\underline{\delta}^k}(\underline{x}).$$

Consider vectors $\underline{\nu}^1, \ldots, \underline{\nu}^k$ from \mathbb{N}_0^r such that $0 \leq \nu_i^j \leq q-2$ and $\nu_i^j \equiv \delta_i^j \pmod{q-1}$ for all *i* and *j*. By the equality $\alpha^{q-1} = \alpha^0$ in \mathbb{F} , for $\alpha_1, \ldots, \alpha_r$ from \mathbb{F} we have

$$P_r(\alpha_1 a_1, \dots, \alpha_r a_r) = \underline{\alpha}^{\underline{\nu}^1} F_{\underline{\delta}^1}(a_1, \dots, a_r) + \dots + \underline{\alpha}^{\underline{\nu}^k} F_{\underline{\delta}^k}(a_1, \dots, a_r)$$

for all $a_1, \ldots, a_r \in \mathbb{F}\langle X \rangle_d$, where $\underline{\alpha}^{\underline{\nu}^j}$ stands for $\alpha_1^{\nu_1^j} \cdots \alpha_r^{\nu_r^j}$. Denote by $(E_{\underline{\alpha}})$ the following linear equation in variables y_1, \ldots, y_k :

$$(E_{\underline{\alpha}}): \underline{\alpha}^{\underline{\nu}^1} y_1 + \dots + \underline{\alpha}^{\underline{\nu}^k} y_k = 0.$$

Consider the system of all linear equations $(E_{\underline{\alpha}})$, where $\alpha_1, \ldots, \alpha_r$ range over the set of non-zero elements of \mathbb{F} .

Consider some non-zero elements $\alpha_2, \ldots, \alpha_r$ from \mathbb{F} . For $0 \leq s \leq q-2$ denote

$$z_s = z_s(\alpha_2, \dots, \alpha_r) = \sum_j \alpha_2^{\nu_2^j} \cdots \alpha_r^{\nu_r^j} y_j,$$

where the sum ranges over all $1 \le j \le k$ with $\nu_1^j = s$. In case $\nu_1^j \ne s$ for all j we set $z_s = 0$ and say that z_s is trivial. Since $0 \le \nu_1^j \le q - 2$ for all j, the equation $z_0 + \alpha_1 z_1 + \alpha_1^2 z_2 + \cdots + \alpha_1^{q-2} z_{q-2} = 0$ holds for all $\alpha_1 \in \mathbb{F}$. Denote non-zero elements of \mathbb{F} by ξ_1, \ldots, ξ_{q-1} . Therefore, $B \cdot (z_0, \ldots, z_{q-2})^T = 0$ for the Vandermonde matrix

$$B = \begin{pmatrix} 1 & \xi_1 & \xi_1^2 & \cdots & \xi_1^{q-2} \\ 1 & \xi_2 & \xi_2^2 & \cdots & \xi_2^{q-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \xi_{q-1} & \xi_{q-1}^2 & \cdots & \xi_{q-1}^{q-2} \end{pmatrix}$$

Then det(B) = $\prod_{1 \le i \le j \le q} (\xi_j - \xi_i)$ is not zero and $z_0 = \cdots = z_{q-2} = 0$ for all $\alpha_2,\ldots,\alpha_r.$

Then for every non-trivial z_s we consider the system of linear equations $z_s(\alpha_2,\ldots,\alpha_r)=0$, where α_2,\ldots,α_r range over the set of non-zero elements of F. Note that for every $1 \leq j \leq k$ the variable y_j appears in one and only one such system of linear equations. Repeating the above reasoning for α_2 , α_3 and so on, we finally obtain that $y_1 = \cdots = y_k = 0$. Lemma 2.6 concludes the proof of the first part of the lemma. The second part follows from part (a) of Remark 2.4. \square

Now we can proof Theorem 2.5.

Proof. By part (b) of Remark 2.4, for a_1, \ldots, a_k from $\langle X \rangle_d$ and $\alpha_1, \ldots, \alpha_k$ from \mathbb{F} we have

$$(\alpha_1 a_1 + \dots + \alpha_k a_k)^n = \sum_{r=1}^{\min\{n,k\}} \sum_{\underline{\delta}} \sum_{\varphi} F_{\underline{\delta}}(\alpha_{\varphi(1)} a_{\varphi(1)}, \dots, \alpha_{\varphi(k)} a_{\varphi(k)}),$$

where $\underline{\delta}$ ranges over q-maximal vectors from \mathbb{N}^r with $|\underline{\delta}| = n$ and φ ranges over strictly increasing maps $\overline{1, r} \to \overline{1, k}$. Since

$$F_{\underline{\delta}}(\alpha_1 a_1, \dots, \alpha_r a_r) = \alpha_1^{\delta_1} \cdots \alpha_r^{\delta_r} F_{\underline{\delta}}(a_1, \dots, a_r),$$

Lemma 2.7 implies that the ideal of relations for $N_{n,d}$ is generated by all $F_{\delta}(\underline{a})$ with q-maximal $\underline{\delta}$ satisfying $|\underline{\delta}| = n$ and $a_i \in \langle X \rangle_d$. Parts (a) and (c) of Remark 2.4 complete the proof of parts (a) and (b) of the theorem, repectively.

Example 2.8. In the formulation of Theorem 2.5 the set S is the following one:

- if $\#\mathbb{F} = n 1$, then S consists of $L_{n-1,1} + L_{1,n-1}$ and $L_{\underline{\theta}}$ for all ordered $\underline{\theta}$ with $|\underline{\theta}| = n$ and $\underline{\theta} \neq (n - 1, 1)$;
- if $\#\mathbb{F} = 2$, then $S = \{P_1, \dots, P_n\};$
- if n = 5 and $\#\mathbb{F} = 3$, then S consists of L_5 , $L_{41} + L_{23}$, $L_{311} + L_{131} + L_{113}$, $L_{221}, L_{2111}, L_{15}.$

Corollary 2.9. Assume \mathbb{F} is a finite field and \mathbb{L} is a finite or infinite field with $\operatorname{char} \mathbb{F} = \operatorname{char} \mathbb{L} = p \text{ and } \# \mathbb{F} < \# \mathbb{L}.$

- 1. Let $\#\mathbb{F} < n$ and $\#\mathbb{L} < n$. If l|t for $\#\mathbb{F} = p^l$ and $\#\mathbb{L} = p^t$, then $C_{n,d}^{\mathbb{F}} \ge C_{n,d}^{\mathbb{L}}$.
- 2. If $\#\mathbb{F} < n \leq \#\mathbb{L}$, then $C_{n,d}^{\mathbb{F}} \geq C_{n,d}^{\mathbb{L}}$. 3. If $\#\mathbb{F} \geq n$ and $\#\mathbb{L} \geq n$, then $C_{n,d}^{\mathbb{F}} = C_{n,d}^{\mathbb{L}}$.

Proof. In part 1 we can assume that $\mathbb{F} \subset \mathbb{L}$, which implies the required. Consider parts 2 and 3 of the theorem. Theorem 2.5 and Remark 2.1 imply that the ideal of relations for $N_{n,d}^{\mathbb{F}}$ ($N_{n,d}^{\mathbb{L}}$, respectively) is generated by some set $W^{\mathbb{F}} \subset \mathbb{F}_p\langle X \rangle_d$ (a set $W^{\mathbb{L}} \subset \mathbb{F}_p\langle X \rangle_d$, respectively). Moreover, $W^{\mathbb{F}}$ belongs to \mathbb{F}_p -span of $W^{\mathbb{L}}$ in part 2 and $W^{\mathbb{F}} = W^{\mathbb{L}}$ in part 3. The proof is completed.

3. The case of
$$x^3 = 0$$

In this section we investigate $N_{3,d}^{\mathbb{F}}$ for any field \mathbb{F} . In the case of an infinite field \mathbb{F} , the nilpotency degree and a basis for $N_{3,d}^{\mathbb{F}}$ were established in [9] developing ideas from [8]. By Theorem 2.5 and Remarks 2.1 and 2.2, in the case of $\#\mathbb{F} > 2$, the nilpotency degree and a basis consisting of words are the same as in the case of infinite field of the same characteristic. So the only case that is left to consider is the case of $\mathbb{F} = \mathbb{F}_2$.

Theorem 3.1. Let $\mathbb{F} = \mathbb{F}_2$. Then for every infinite field \mathbb{L} of characteristic 2 we have

 $\begin{array}{ll} \text{(a)} & C_{3,d}^{\mathbb{F}} = C_{3,d}^{\mathbb{L}} = \left\{ \begin{array}{ll} 6, & \textit{if } d = 2 \\ d + 3, & \textit{if } d \geq 3 \end{array} \right. ; \\ \text{(b)} & \dim N_{3,d}^{\mathbb{F}} = \dim N_{3,d}^{\mathbb{L}} + d(d-1)/2. \end{array}$

The proof of this theorem is given at the end of the section.

Definition 3.2 $(of \approx)$. Given an $f \in \mathbb{F}\langle X \rangle_d$, we write $f \approx 0$ in $N_{n,d}$ if and only if either f = 0 in $N_{n,d}$ or $f = \sum_i \alpha_i w_i$ in $N_{n,d}$ for $\alpha_i \in \mathbb{F}$ and $w_i \in \langle X \rangle$ satisfying $\deg(w_i) > \deg(f)$.

Remark 3.3. If $f \approx 0$ in $N_{n,d}$, then $ufv \approx 0$ in $N_{n,d}$ for all $u, v \in \langle X \rangle_d^{\#}$. On the other hand, note that in case $f \approx 0$ and $h \approx 0$ in $N_{n,d}$ we can have that $f + h \not\approx 0$ in $N_{n,d}$. Similarly, in case $f \approx 0$ in $N_{n,d}$ and f = h in $N_{n,d}$ we can have that $h \not\approx 0$ in $N_{n,d}$.

We will use the following remark to obtain relations for $N_{n,d}$.

Remark 3.4. Define the inversion * on $\mathbb{F}\langle X \rangle$ as follows: $x_i^* = x_i$ for all i, $(ab)^* = b^*a^*$ for all $a, b \in \langle X \rangle$ and $* : \mathbb{F}\langle X \rangle \to \mathbb{F}\langle X \rangle$ is a linear map. Then for any relation $f \in \mathbb{F}\langle X \rangle_d$ for $N_{n,d}$ we have that f^* is also a relation for $N_{n,d}$.

In the rest of this section we assume that $\mathbb{F} = \mathbb{F}_2$ unless otherwise stated. By Theorem 2.5 and Example 2.8, we have

(2)
$$L_{111}(x, y, z) = 0$$
 in $N_{3,d}$

(3)
$$L_{21}(x,y) + L_{12}(x,y) = 0$$
 in $N_{3,d}$

Moreover, the ideal of relations for $N_{3,d}$ is generated by x^3 and left hand sides of equalities (2) and (3) for $x, y, z \in \langle X \rangle_d$. Note that $L_{111}(x, y, z) = xyz + xzy + yxz + yzx + zxy + zyx$ and $L_{21}(x, y) = x^2y + xyx + yx^2$. By the straightforward calculation we can see that

$$xL_{21}(y,z) = L_{21}(y,xz) + L_{111}(x,y,yz) - yL_{111}(x,y,z)$$

holds in $\mathbb{F}\langle X \rangle$. Then equality (2) implies that

(4)
$$xL_{21}(y,z) = L_{21}(y,xz)$$
 in $N_{3,d}$

By Remark 3.4, we have

(5)
$$L_{21}(x,y)z = L_{21}(x,yz)$$
 in $N_{3,d}$.

Lemma 3.5. We have $L_{21}(a,b) \approx 0$ in $N_{3,d}$ for all $a, b \in \langle X \rangle_d$ with deg(a) > 1.

Proof. In this proof we work in $N_{3,d}$. It is convenient to rewrite equalities (4) and (5) as follows

(6)
$$xyzy = xy^2z + xzy^2 + L_{21}(y, xz),$$

(7)
$$xyxz = x^2yz + yx^2z + L_{21}(x, yz),$$

respectively. Making a substitution $z \to yz$ in (7), we obtain

$$xyxyz = x^2y^2z + yx^2yz + L_{21}(x, y^2z).$$

Application of (7) to yx^2yz gives us

(8)
$$xyxyz = y^2x^2z + L_{21}(x, y^2z) + L_{21}(y, x^2z).$$

By Remark 3.4, we have

(9)
$$zxyxy = zy^2x^2 + L_{21}(y, zx^2) + L_{21}(x, zy^2).$$

Making substitutions $y \to yz$ and $z \to y$ in (7), we obtain

$$xyzxy = x^2yzy + yzx^2y + L_{21}(x, yzy).$$

Application of (6) to x^2yzy and the equality $yzx^2y = y^2zx^2 + zx^2y^2 + L_{21}(y, zx^2)$ in $\mathbb{F}\langle X \rangle$ gives us

$$xyzxy = x^{2}y^{2}z + x^{2}zy^{2} + zx^{2}y^{2} + y^{2}zx^{2} + L_{21}(y, x^{2}z) + L_{21}(y, zx^{2}) + L_{21}(x, yzy) + L_{21}(x, yzy) + L_{21}(y, yzy) + L_$$

Consider $L_{21}(xy, z) = xyxyz + zxyxy + xyzxy$. Applying (8), (9), (10) to the first, second and third summands of $L_{21}(xy, z)$, respectively, we obtain

$$L_{21}(xy,z) = L_{111}(x^2, y^2, z) + L_{21}(x, y^2z) + L_{21}(y, x^2z) + L_{21}(x, zy^2) + L_{21}(y, zx^2) + L_{21}(y, x^2z) + L_{21}(y, zx^2) + L_{21}(x, yzy).$$

By equalities (2) and (3), $L_{21}(xy, z) = f(x, y, z)$ for

$$f(x, y, z) = L_{12}(x, y^2 z) + L_{12}(y, x^2 z) + L_{12}(x, zy^2) + L_{12}(y, zx^2) + L_{12}(y, zx^2) + L_{12}(y, zx^2) + L_{12}(x, yzy).$$

Consider a, b from the formulation of the lemma. Let us recall that $\deg(a) > 1$. We set a = xy for $x, y \in \langle X \rangle_d$ such that

• if $\deg(a)$ is even, then $\deg(x) = \deg(y) = r \ge 1$;

• if deg(a) is odd, then deg(x) = r + 1 and deg(y) = $r \ge 1$.

We have $L_{21}(a,b) = f(x,y,b)$. Note that f(x,y,b) is a sum of words of degrees $k_1 = \deg(x) + 4\deg(y) + 2\deg(b)$ and $k_2 = 4\deg(x) + \deg(y) + 2\deg(b)$. For $k_0 = \deg(L_{21}(a,b)) = 2\deg(x) + 2\deg(y) + \deg(b)$ we claim that

$$k_0 < k_1$$
 and $k_0 < k_2$.

For short, we write s for deg(b). In the case of even deg(a) we have $k_0 = 4r + s$, $k_1 = k_2 = 5r + 2s$ and in the case of odd deg(a) we have $k_0 = 4r + s + 2$, $k_1 = 5r + 2s + 1, k_2 = 5r + 2s + 4$. Hence, the claim is proven and $L_{21}(a, b) \approx 0$. \Box

Lemma 3.6. We have $uL_{21}(a,b)v \approx 0$ in $N_{3,d}$ for all $a,b \in \langle X \rangle_d$, $u,v \in \langle X \rangle_d^{\#}$ with $\deg(ua^2bv) \ge 4$.

Proof. Assume that u = v = 1. In case deg(a) > 1 the required statement follows from Lemma 3.5. Let $\deg(a) = 1$ and $\deg(b) \ge 2$. Then equality (3) gives us the required stetement.

Assume that $\deg(u) > 0$. Since the degree of the right hand side of formula (4) is equal to the degree of its left hand side, formula (4) completes the proof. \square

Now we can proof Theorem 3.1.

Proof. Denote by $V_r^{\mathbb{F}}$ the \mathbb{F} -span of all words of $\langle X \rangle_d$ of degree r. We write $U_r^{\mathbb{F}}$ for the image of $V_r^{\mathbb{F}}$ in $N_{3,d}^{\mathbb{F}}$. Note that if a summand of relation (3) has degree three, then the rest of summands of this relation have degree three. Therefore, $N_{3,d}^{\mathbb{F}} = U_3^{\mathbb{F}} \oplus W^{\mathbb{F}}$ for $W^{\mathbb{F}} = \sum_{r>3} U_r^{\mathbb{F}}$. Let $B = \{b_1, \ldots, b_s\}$ be an L-basis for $N_{3,d}^{\mathbb{L}}$ consisting of words. Denote by B_r

the set of all $b \in B$ of degree r. We claim that

(11) every $w \in \langle X \rangle_d$ with $\deg(w) \ge 4$ belongs to \mathbb{F} -span of $\sqcup_{r \ge 4} B_r$ in $N_{3,d}^{\mathbb{F}}$.

We set $\deg(w) = r \ge 4$. Since $N_{3,d}^{\mathbb{L}}$ is homogeneous with respect to the degree, we have $w = \sum_i \beta_i b_i$ in $N_{3,d}^{\mathbb{L}}$ for some $\beta_i \in \mathbb{F}$ and $b_i \in B_r$ (see also Remarks 2.1) and 2.2). Thus relation (2) together with Lemma 3.6 imply that $w - \sum_i \beta_i b_i \approx 0$ in $N_{3,d}^{\mathbb{F}}$. Thus, $w = \sum_i \beta_i b_i + \sum_j \gamma_j w_j$ for some $\gamma_j \in \mathbb{F}$ and $w_j \in \langle X \rangle_d$ with $\deg(w_j) > r$. Then we consider w_j and so on. Since $N_{3,d}^{\mathbb{F}}$ is nilpotent, this process will stop at some step. The claim is proven. Similarly to part (b) of Remark 2.2, we can see that the set $\sqcup_{r\geq 4}B_r$ is linearly independent in $N_{3,d}^{\mathbb{F}}$. Thus, $\sqcup_{r\geq 4}B_r$ is a basis for $W^{\mathbb{F}}$.

Denote by B_{111} the set of all $x_i x_j x_k$ with pairwise different $i, j, k \in \{1, \ldots, d\}$ such that (i, j, k) is not ordered and denote by B_{21} the set of words $x_i^2 x_j$ with $1 \leq i \neq j \leq d$. It is not difficult to see that the set $B_{111} \sqcup B_{21}$ is a basis for $U_3^{\mathbb{L}}$; and $B_{111} \sqcup B_{21}$ together with words $x_i x_j x_i$, where $1 \le i < j \le d$, is a basis for $U_3^{\mathbb{F}}$. The theorem is proven.

The next remark follows from the proof of Theorem 3.1.

Remark 3.7. Let $\mathbb{F} = \mathbb{F}_2$ and \mathbb{L} be an infinite field of characteristic two. If we add elements $x_i x_j x_i$, $1 \le i < j \le d$, to the basis for $N_{3,d}^{\mathbb{L}}$ from Theorem 2 of [9], then we obtain a basis for $N_{3,d}^{\mathbb{F}}$.

4. The case of $\#\mathbb{F} = n-1$

Let us remark that the case of $\#\mathbb{F} \ge n$ is considered in Corollary 2.9. In this section we investigate the case of field \mathbb{F} with n-1 elements.

Theorem 4.1. Let $\#\mathbb{F} = n-1$. Then for every infinite field \mathbb{L} with char $\mathbb{L} = \operatorname{char} \mathbb{F}$ we have $C_{n,d}^{\mathbb{L}} \leq C_{n,d}^{\mathbb{F}} \leq C_{n,d}^{\mathbb{L}} + 1.$

We split the proof of the theorem into several lemmas. The following lemma is a modification of Lemma 1 of Chapter 6 of [16].

Lemma 4.2. For every field \mathbb{F} we have

$$n xy^{n} = L_{n-1,1}(y, xy) - (n-1)L_{n-1,1}(y, yx) + L_{n-2,1,1}(y, x, y^{2})$$

in $\mathbb{F}\langle X \rangle$.

Proof. For short, we write x for x_1 and y for x_{n+1} . In $\mathbb{Z}\langle X \rangle$ holds

(12)
$$L_{n-1,1}(x,xy) = \sum_{i=0}^{n-1} x^i xy x^{n-i-1} \\ = x \sum_{i=0}^{n-1} x^i y x^{n-i-1} = x L_{n-1,1}(x,y).$$

Note that two elements of $\mathbb{Z}\langle X \rangle$ are equal to each other if and only if its homogeneous components with respect to the multidegree are equal to each other. Thus, making the substitution $x \to x_1 + \cdots + x_n$ in the above equality and taking the homogeneous component of multidegree 1^{n+1} (i.e. of degree one in each of letters x_1,\ldots,x_n,y , we obtain

$$\sum_{i=1}^{n} L_{1^{n}}(x_{1}, \dots, x_{i-1}, x_{i+1}, \dots, x_{n}, x_{i}y) = \sum_{i=1}^{n} x_{i}L_{1^{n}}(x_{1}, \dots, x_{i-1}, x_{i+1}, \dots, x_{n}, y).$$

We make substitutions $x_2 \to y, \ldots, x_n \to y$ and recall that $x_1 = x$. Thus in $\mathbb{Z}\langle X \rangle$ we have

(13)
$$L_{1^{n}}(\underbrace{y,\ldots,y}_{n-1},xy) + (n-1)L_{1^{n}}(x,\underbrace{y,\ldots,y}_{n-2},y^{2}) \\ = xL_{1^{n}}(y,\ldots,y,y) + (n-1)yL_{1^{n}}(x,\underbrace{y,\ldots,y}_{n-2},y).$$

By (12), the second summand of the right hand side of (13) is equal to (n-2)!(n-1) $(1)^2 L_{n-1,1}(y, yx)$. Hence

(14)
$$(n-1)! L_{n-1,1}(y,xy) + (n-1)! L_{n-2,1,1}(y,x,y^2) = n! xy^n + (n-2)! (n-1)^2 L_{n-1,1}(y,yx)$$

in $\mathbb{Z}\langle X \rangle$. Obviously, we can divide (14) by (n-1)!. As the result, we obtain that the required equality hold over \mathbb{Z} and, therefore, it holds over \mathbb{F} .

In the rest of this section we assume that $\#\mathbb{F} = n-1$ and $n \ge 3$. By Theorem 2.5 (see also Example 2.8), in $N_{n,d}$ we have

(15)
$$L_{1,n-1}(x,y) + L_{n-1,1}(x,y) = 0$$
 and

(16) $L_{\theta}(x_1,\ldots,x_r) = 0$, where $\#\underline{\theta} = r$, $|\underline{\theta}| = n$ and $\underline{\theta} \notin \{(1,n-1), (n-1,1)\}$.

Lemma 4.3. In case $\#\mathbb{F} = n - 1$ the following equalities hold in $N_{n,d}$:

- (a) $xL_{n-1,1}(y,z) = L_{n-1,1}(y,xz);$
- (b) $L_{n-1,1}(x,y)z = L_{n-1,1}(x,yz);$ (c) $L_{n-1,1}(x,yz) = -L_{n-1,1}(z,yx).$

Proof. We can assume that $x = x_1$, $y = x_2$, and $z = x_3$. By Lemma 4.2,

$$xy^{n} = L_{n-1,1}(y, xy) + L_{n-2,1,1}(y, x, y^{2})$$

in $\mathbb{F}\langle X \rangle$. Note that two elements of $\mathbb{F}\langle X \rangle$ are equal to each other if and only if its homogeneous components with respect to the multidegree are equal to each other. Thus, making the substitution $y \to y + z$ in the above equality and taking the homogeneous component of multidegree (1, n - 1, 1) (i.e. of degree one in x, z and degree n - 1 in y), we obtain

$$xL_{n-1,1}(y,z) = L_{n-1,1}(y,xz) + L_{n-2,1,1}(y,z,xy) + L_{n-2,1,1}(y,x,yz) + L_{n-2,1,1}(y,x,zy) + L_{n-3,1,1,1}(y,z,x,y^2)$$

in $\mathbb{F}\langle X \rangle$. Applying (16), we complete the proof of part (a) of the lemma. Part (b) follows from Remark 3.4.

By (15), $xL_{n-1,1}(y,z) + xL_{n-1,1}(z,y) = 0$ in $N_{n,d}$. Thus part (a) of the lemma implies part (c).

Lemma 4.4. In case $\#\mathbb{F} = n-1$ we have $uL_{n-1,1}(a,b)v \approx 0$ in $N_{n,d}$ for all $a, b \in \langle X \rangle_d$, $u, v \in \langle X \rangle_d^{\#}$ with $\deg(ubv) \geq 2$.

Proof. We work in $N_{n,d}$. Note that the equalities from parts (a) and (b) of Lemma 4.3 are homogeneous with respect to the degree. Applying these equalities, without loss of generality we can assume that $\deg(b) \geq 2$. We set $f = L_{n-1,1}(a,b)$ and b = cx for a letter x and a word $c \in \langle X \rangle_d$. Part (c) of Lemma 4.3 and (15) imply that $f = -L_{n-1,1}(x, ca) = L_{n-1,1}(ca, x)$. The inequality $\deg(L_{n-1,1}(ca, x)) > \deg(f)$ completes the proof.

Now we can proof Theorem 4.1.

Proof. We set $p = \operatorname{char} \mathbb{F} = \operatorname{char} \mathbb{L}$. Consider a $k > C = C_{n,d}^{\mathbb{L}}$. Let u be a word of $\langle X \rangle_d$ with $\deg(u) = k$. We set u = vx for a letter x and a word $v \in \langle X \rangle_d$ with $\deg(v) \geq C$. Since the ideal of relations for $N_{n,d}^{\mathbb{L}}$ is generated by elements with coefficients 0 and 1 (see Remark 2.1), we have

$$v = \sum_i \alpha_i L_{\underline{\theta}_i}(\underline{a}_i)$$

in $\mathbb{F}\langle X \rangle$ for $\alpha_i \in \mathbb{F}_p$, $|\underline{\theta}_i| = n$, and $\underline{a}_i = (a_{i1}, \dots, a_{ir})$, where $r = \#\underline{\theta}_i$ and $a_{i1}, \dots, a_{ir} \in \langle X \rangle_d$. Lemma 4.4 together with relation (16) imply $u = vx \approx 0$ in $N_{n,d}^{\mathbb{F}}$. Thus if w = 0 in $N_{n,d}^{\mathbb{F}}$ for all $w \in \langle X \rangle_d$ with $\deg(w) = k + 1$, then w = 0 in $N_{n,d}^{\mathbb{F}}$ for all $w \in \langle X \rangle_d$ with $\deg(w) = k$. Hence $C_{n,d}^{\mathbb{F}} \leq C_{n,d}^{\mathbb{L}} + 1$. The inequality $C_{n,d}^{\mathbb{L}} \leq C_{n,d}^{\mathbb{F}}$ follows from part 2 of Corollary 2.9.

Corollary 2.9, Theorem 4.1 together with Theorem 5.1 from [13] imply the following corollary.

Corollary 4.5. For an arbitrary field \mathbb{F} of characteristic p and $d \geq 2$ we have

- $C_{4,d} = 10$, if p = 0;
- $3d < C_{4,d}$, if p = 2;
- $3d + 1 \le C_{4,d} \le 3d + 5$, if p = 3;
- $10 \le C_{4,d} \le 13$, if p > 3.

Acknowledgements

This paper was supported by FAPESP No. 2011/51047-1.

References

- A. Kanel-Belov, L.H. Rowen, Computational aspects of polynomial identities, Research Notes in Mathematics 9, A.K. Peters, Wellesley, MA, 2005, 378 pp.
- [2] A.J. Belov, M.I. Kharitonov, Subexponential esimations in Shirshov height theorem, Mat. Sb. 203 (2012), no. 4, 81–102 (Russian).
- [3] M. Domokos, S.G. Kuzmin, A.N. Zubkov, Rings of matrix invariants in positive characteristic, J. Pure Appl. Algebra 176 (2002), 61–80.
- [4] S. Donkin, Invariants of several matrices, Invent. Math. 110 (1992), 389–401.
- [5] V. Drensky, E. Formanek, *Polynomial identity rings*, Advanced Courses in Mathematics, CRM Barcelona, Birkhäuser, Basel-Boston, 2004.
- B. Eick, Computing nilpotent quotients of associative algebras and algebras satisfying a polynomial identity, Int. J. Algebra Comput. 21 (2011), 1339–1355.
- [7] E.N. Kuzmin, On the Nagata-Higman theorem, Mathematical Structures Computational Mathematics — Mathematical Modeling, Proceedings Dedicated to the 60th Birthday of Academician L. Iliev, Sofia, 1975, 101–107 (Russian).
- [8] A.A. Lopatin, The algebra of invariants of 3 × 3 matrices over a field of arbitrary characteristic, Commun. Algebra 32 (2004), No. 7, 2863–2883.
- [9] A.A. Lopatin, Relatively free algebras with the identity x³ = 0, Commun. Algebra 33 (2005), No. 10, 3583–3605.
- [10] A.A. Lopatin, Indecomposable invariants of quivers for dimension (2,...,2) and maximal paths, Commun. Algebra, 38 (2010), no. 10, 3539–3555.
- [11] A.A. Lopatin, Indecomposable invariants of quivers for dimension (2,...,2) and maximal paths, II, Sib. Elektron. Mat. Izv., 7 (2010), 350–371.
- [12] A.A. Lopatin, On minimal generating systems for matrix O(3)-invariants, Linear Mult. Algebra, 59 (2011), No. 1, 87–99.
- [13] A.A. Lopatin, On the nilpotency degree of the algebra with identity $x^n = 0$, J. Algebra, **371** (2012), 350–366.
- [14] Yu.P. Razmyslov, Trace identities of full matrix algebras over a field of characteristic 0, Izv. Akad. Nauk SSSR, Ser. Mat. 38 (1974), No. 4, 723–756 (Russian). English translation: Math. USSR, Izv. 8 (1974), No. 4, 727–760.
- [15] I.P. Shestakov, N. Zhukavets, On associative algebras satisfying the identity $x^5 = 0$, Algebra Discrete Math. (2004), No. 1, 112–120.
- [16] K.A. Zhevlakov, A.M. Slin'ko, I.P. Shestakov, A.I. Shirshov, *Rings that are nearly associative*, Nauka, Moscow, 1978, 431 pp. (in Russian). English translation: Pure and Applied Math. 104, Academic Press, 1982, 371 pp.
- [17] A.I. Shirshov, On rings with polynomial identities, Mat. Sb. 43 (1957), 277–283. English translation: Twelve Papers in Algebra, Amer. Math. Soc. Transl. 119, 1983, 139pp.
- [18] M.R. Vaughan-Lee, An algorithm for computing graded algebras, J. Symbolic Comput. 16 (1993), 345–354.
- [19] E. Zelmanov, Some open problems in the theory of infinite dimensional algebras, J. Korean Math. Soc. 44 (2007), No. 5, 1185–1195.

ARTEM A. LOPATIN OMSK BRANCH OF FEDERAL STATE BUDGETARY SCIENTIFIC ESTABLISHMENT SOBOLEV INSTITUTE OF MATHEMATICS, SIBERIAN BRANCH OF THE RUSSIAN ACADEMY OF SCIENCES (OB IM SBRAS). PEVTSOVA STREET, 13, 644043, OMSK, RUSSIA. *E-mail address:* ARTEM_LOPATIN@YAHOO.COM

IVAN P. SHESTAKOV INSTITUTE OF MATHEMATICS AND STATISTICS, UNIVERSITY OF SÃO PAULO, CAIXA POSTAL 66281, SÃO PAULO-SP, BRAZIL, 05311-970. *E-mail address:* Shestak@ime.usp.br

12