Journal of Algebra 434 (2015) 283–301

Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

Algebraic properties of word equations

Štěpán Holub [*,1], Jan Žemlička

Department of Algebra, Charles University, Sokolovská 83, 175 86 Praha, Czech Republic

ARTICLE INFO

Article history:
Received 26 September 2014
Available online 23 April 2015
Communicated by J.T. Stafford

MSC:
68R15

Keywords:
Independent systems of word equations
Multivariate polynomials

ABSTRACT

The question about maximal size of independent system of word equations is one of the most striking problems in combinatorics on words. Recently, Aleksi Saarela has introduced a new approach to the problem that is based on linear-algebraic properties of polynomials encoding the equations and their solutions. In this paper we develop further this approach and take into account other algebraic properties of polynomials, namely their factorization. This, in particular, allows to improve the bound for the number of independent equations with maximal rank from quadratic to linear.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The question about maximal size of independent system of word equations is one of the most striking problems in combinatorics on words. The conjecture that such a system cannot be infinite (known as Ehrenfeucht's conjecture) had been open for more than a decade, until solved in [1] by embedding the free monoid into a metabelian group, and independently in [4] by using matrix representation (and generalizing the result to free groups). Both these solutions indicated that, despite strongly combinatorial nature

* Corresponding author.
E-mail addresses: holub@karlin.mff.cuni.cz (Š. Holub), zemlicka@karlin.mff.cuni.cz (J. Žemlička).

[1] Supported by the Czech Science Foundation grant number 13-01832S.

http://dx.doi.org/10.1016/j.jalgebra.2015.03.021
0021-8693/© 2015 Elsevier Inc. All rights reserved.

of words, algebraic methods may be necessary to approach problems concerning word equations.

Since the proof of Ehrenfeucht's conjecture, the question about the possible size of independent systems moved to the center of investigation. The above algebraic methods, based ultimately on Hilbert's basis theorem, did not help to approach the problem. The solution is simple for two unknowns because any nontrivial word equation over two unknowns possesses only periodic solutions. However, already for three unknowns, no upper bound is known up to date, leaving open even the possibility that arbitrarily large independent systems of equations over three unknowns exist. This should be contrasted to the fact that the largest known independent system over three unknowns consists of three equations. Some partial results have been obtained in [6] and [3]. For general number n of unknowns, independent systems of size $\Theta(n^4)$ have been explicitly constructed in [7]. For more details see also [5].

The research got a new impetus recently, when Aleksi Saarela, in [10], introduced an idea how to encode word equations to the language of polynomial algebra and exploit its linear algebraic properties. This approach allowed to obtain an upper bound for the size of an independent system of equations over three unknowns, namely quadratic in the length of the shortest equation. In general, the method allows to limit the size of independent systems that have solutions with maximal possible rank. In this paper we develop further this approach, and applying classical algebraic tools, in particular irreducible factorization of multivariate polynomials, we are able to improve Saarela's bounds from quadratic to linear.

Basic facts about word equations and corresponding linear algebra are introduced in Section 2. The properties of principal solutions (Theorem 1) and their rank are presented here. The main goal of Section 3 is to translate Saarela's results from less usual notation of general monoid rings to classical terminology of multivariate rational polynomials which allows to employ divisibility of obtained polynomials. An application of these tools gives bounds proved in the final section of this paper. As a consequence of two observations, elementary algebraic (Lemma 11) and finer one of geometrical nature (Theorem 14), we get two pairs of upper bounds. The first are bounds on the number of pairwise linearly nonequivalent solutions of rank $n-1$ of two strongly independent equations in n unknowns (Theorem 19) and the second are the bounds on the maximal size of strongly independent system of word equations (Theorem 21).

2. Solutions of word equations

In this section we review some facts about systems of word equations and their solutions.

Throughout the paper, n is a number of unknowns, i.e. an integer usually greater then 2, and $\mathbb{N}_0$ denotes the set of all nonnegative integers. If $\Xi = \{x_1, x_2, \dots, x_n\}$ is a set of unknowns, then a pair $(u, v) \in \Xi^* \times \Xi^*$ is an *equation*. A morphism $h : \Xi^* \to \Sigma^*$ is a *solution* of a system of equations $T = \{(u_i, v_i) \mid i \in I\}$ if $h(u_i) = h(v_i)$ for each $i \in I$.

If $h(c)$ is the empty word for at least one c in the domain alphabet, then we say that h is *erasing*. The set of letters occurring in a word w is denoted by $\mathrm{alp}(w)$ and $\mathrm{alp}(h)$ denotes $\bigcup_{x\in\Xi}\mathrm{alp}(h(x))$.

The system is *trivial* if $u_i = v_i$ for all $i \in I$. The vector

$$L(h) = (|h(x_1)|, |h(x_2)|, \dots, |h(x_n)|) \in \mathbb{N}_0^n$$

is called the *length type* of h.

For a solution h of the system T, we shall implicitly assume that the domain alphabet Ξ of h is equal to $\mathrm{alp}(T) = \bigcup_{i\in I}\mathrm{alp}(u_i v_i)$. The number of occurrences of a letter c in a word w is denoted by $|w|_c$.

We say that the solution h' of T *divides* the solution h if $h = \vartheta \circ h'$ where ϑ is not erasing and defined on $\mathrm{alp}(h')$. The solution is called *principal* if it is minimal in the divisibility ordering just defined. In other words, if h is principal and $h = \vartheta \circ h'$ for a solution h', then ϑ is a renaming of letters.

For the sake of completeness we prove the following theorem. Our proof partly follows the standard reference, Proposition 9.5.2 in [9], which, however, is not formulated precisely.

Theorem 1. *Let $h : \Xi^* \to \Sigma^*$ be a solution of a system T. Then there is a unique (up to renaming of letters) principal solution g of T and a unique morphism $\vartheta : \mathrm{alp}(g)^+ \to \Sigma^+$ such that $h = \vartheta \circ g$.*

Moreover,

- $|\mathrm{alp}(g)| < |\mathrm{alp}(T)|$ *if T is not trivial; and*
- *g and $L(\vartheta)$ depend only on $L(h)$ and T.*

Proof. If $h = \vartheta \circ g$ where ϑ is not a renaming of letters, then

$$\sum_{x\in\mathrm{alp}(T)} |h(x)| > \sum_{x\in\mathrm{alp}(T)} |g(x)| \quad \text{or} \quad |\mathrm{alp}(h)| > |\mathrm{alp}(g)|\,.$$

Consequently, by induction, there is at least one principal solution g such that $h = \vartheta \circ g$ for some ϑ.

In order to show that ϑ is given uniquely and that g and $L(\vartheta)$ are given by $L(h)$, we again proceed by induction. Clearly, $h(x)$ is empty if and only if $g(x)$ is empty. We can therefore suppose that h is not erasing.

If T is trivial, then the only principal solution is identity, $\vartheta = h$ and $L(\vartheta) = L(h)$.

Let $u \neq v$ for some $(u, v) \in T$, and let rx be a prefix of u and ry a prefix of v, where $x \neq y$ are letters, and $r \in \Xi^*$.

Let us first assume that $|h(x)| < |h(y)|$. Then $h(x)$ is a prefix of $h(y)$. Moreover, also $|g(x)| < |g(y)|$ and $g(x)$ is a prefix of $g(y)$. Define $\varphi : \Xi^+ \to \Xi^+$ by

$$\varphi(z) = \begin{cases} xy, & \text{if } z = y, \\ z, & \text{if } z \neq y, \end{cases}$$

and h', g' by

$$h'(z) = \begin{cases} h(x)^{-1}h(y) & \text{if } z = y, \\ h(z) & \text{if } z \neq y, \end{cases} \qquad g'(z) = \begin{cases} g(x)^{-1}g(y) & \text{if } z = y, \\ g(z) & \text{if } z \neq y. \end{cases}$$

Then g' and $h' = \vartheta \circ g'$ are solutions of the system $T' = \{(\varphi(u_i), \varphi(v_i)) \mid i \in I\}$. If $g' = \vartheta' \circ g''$, where g'' is a solution of T', then $g'' \circ \varphi$ is a solution of T and $g = \vartheta' \circ g'' \circ \varphi$, which implies that ϑ' is a renaming of letters. Therefore g' is a principal solution of (u', v'). By induction assumption, ϑ and g' are unique, and g' and $L(\vartheta)$ are uniquely given by $L(h')$ which is in turn determined by $L(h)$. Since $g = g' \circ \varphi$, we obtain that also g is unique and determined by $L(h)$. Since φ is invertible (in the free group), we have $u' \neq v'$ and the induction yields

$$|\mathrm{alp}(g)| = |\mathrm{alp}(g')| < |\mathrm{alp}(T')| = |\mathrm{alp}(T)|.$$

The proof is symmetric for $|h(x)| < |h(y)|$. If $|h(x)| = |h(y)|$, then the proof is analogous if we define

$$\varphi(y) = x, \quad \text{and} \quad \varphi(z) = z \text{ otherwise},$$

and h' and g' are restrictions of h and g respectively on the alphabet of T' which is $\mathrm{alp}(T) \setminus \{y\}$. In this case, the system T' can be trivial but we have $\mathrm{alp}(T') < \mathrm{alp}(T)$. $\square$

Let $h : \{x_1, x_2, \ldots, x_n\}^* \to \{a_1, a_2, \ldots, a_m\}^*$ be a morphism. In order to employ linear and polynomial algebra we define non-negative rational vectors $\mathfrak{g}(h)_i$, $i = 1, 2, \ldots, m$, by

$$\mathfrak{g}(h)_i = (|h(x_1)|_{a_i}, |h(x_2)|_{a_i}, \ldots, |h(x_n)|_{a_i}).$$

The set of vectors $\{\mathfrak{g}(h)_1, \mathfrak{g}(h)_2, \ldots, \mathfrak{g}(h)_m\}$ is denoted by G_h, Γ_h is the rational vector space generated by G_h, and the dimension of Γ_h will be called the *rank* of h. Note that the space Γ_h will turn out to be an important linear-algebraic characteristic of a solution h of rank $n-1$, that is when Γ_h is a hyperplane.

Since rational vectors will serve as one of the main tools in this paper, let us introduce corresponding notation; by $\cdot$ we denote the standard dot product on $\mathbb{Q}^n$, for every vector $\alpha \in \mathbb{Q}^n$, the i-th coordinate of α is denoted by $(\alpha)_i$ and $\alpha_\oplus$, $\alpha_\ominus$ represent the uniquely determined nonnegative vectors for which $\alpha = \alpha_\oplus - \alpha_\ominus$ and $\alpha_\oplus \cdot \alpha_\ominus = 0$. If $M \subseteq \mathbb{Q}^n$, then $M\mathbb{Q}$ is the subspace of the vector space $\mathbb{Q}^n$ generated by M.

If $\alpha \in \mathbb{N}_0^m$, then the endomorphism ϑ_α of $\{a_1, a_2, \ldots, a_m\}^*$ will be defined by the condition $a_i \mapsto a_i^{(\alpha)_i}$. Rank of principal solutions has the following property.

Lemma 2. *Let g be a principal solution of a system T. Then the rank of g is equal to the cardinality of* $\mathrm{alp}(g)$. *If $h = \vartheta \circ g$, then the rank of h is at most the rank of g.*

Proof. We want to show that the set G_g is linearly independent. Suppose it is not. Then there are two distinct vectors $\alpha_1, \alpha_2 \in \mathbb{N}^{|\mathrm{alp}(g)|}$ such that $L(\vartheta_{\alpha_1} \circ g) = L(\vartheta_{\alpha_2} \circ g)$. Since both $\vartheta_{\alpha_1} \circ g$ and $\vartheta_{\alpha_2} \circ g$ are solutions of T, we obtain a contradiction with Theorem 1.

It is easy to see that $G_h\mathbb{Q}$ is a subspace of $G_g\mathbb{Q}$, which implies the second statement. □

Remark 1. There are several kinds of a rank defined in the literature, see for example [2]. The "combinatorial rank" is the smallest cardinality of a set A such that $h(x) \in A^*$ for each $x \in \Xi$. The combinatorial rank is used in [10].

From the algebraic point of view, most natural is probably the "free rank": the size of the basis of the smallest free monoid containing each $h(x)$, $x \in \Xi$.

We remark without proof that all the ranks, including the "linear rank" we use in this paper, coincide for principal solutions. It is therefore convenient and recommended, whenever possible, to consider, instead of a general solution $h = \vartheta \circ g$, the principal solution g that divides it, see also Lemma 2. The morphism ϑ typically destroys some properties of the rank. For example, if it maps two different letters to the same one. As another example, consider a principal solution of rank three and $\vartheta : a \mapsto abc,\ b \mapsto bca,\ c \mapsto cab$. Then ϑ preserves both combinatorial and free rank but lowers the linear rank to one. Also, the linear rank and the combinatorial rank are not preserved by the usual binary encoding $a_i \mapsto ab^i$.

Note that Theorem 1 implies that a principal solution of a nontrivial system in n unknowns has rank at most $n-1$, which is known in the combinatorics on words as the *defect effect*.

Put $L_g = \{L(\vartheta_\alpha \circ g) \mid \alpha \in \mathbb{N}_0^k\}$. In order to point out the property of sets L_g which will play an important role in linking word equations and linear algebra, let us define the notion of *rank* of a subset M of the rational vector space $\mathbb{Q}^n$ as such an integer r that $\dim M\mathbb{Q} = r$ and M is not covered by any finite union of $(r-1)$-dimensional subspaces of $M\mathbb{Q}$. Moreover for $M \subseteq \mathbb{Q}^n$ let $M\mathbb{N}$ denote the set $\left\{\sum_i a_i\alpha_i \mid a_i \in \mathbb{N},\ \alpha_i \in M\right\}$ and formulate a lemma describing ranks of some subsets of lattice points of r-dimensional Euclidean space

Lemma 3. *If $G \subset \mathbb{Q}^n$ and $N \subset \mathbb{Z}^n$ are such that $G\mathbb{N} \subseteq N \subseteq G\mathbb{Q}$, then N is of rank* $\dim G\mathbb{Q}$.

Proof. Let r denote $\dim G\mathbb{Q}$. Since $\dim N\mathbb{Q} = r$ by hypotheses, it remains to show that N is not covered by a finite number of spaces with the dimension less than r. It is enough to prove that $G\mathbb{N}$ is not covered so.

Consider the set $M = \{(k, k^2, \ldots, k^r) \mid k \in \mathbb{N}\} \subset \mathbb{N}^r$ and a linearly independent subset $\{\gamma_1, \ldots, \gamma_r\}$ of the set G. Each r distinct elements of M are linearly independent

in $\mathbb{Q}^r$ since they form a Vandermonde matrix. Therefore also each r distinct elements of the set

$$M_G = \left\{ \sum_{i=1}^{r} k^i \gamma_i \mid k \in \mathbb{N} \right\}$$

are linearly independent. The claim follows since $M_G \subset G\mathbb{N}$ and since any subspace of $\mathbb{Q}^n$ with the dimension less than r contains at most $r-1$ elements of M_G. $\square$

The following consequence of the previous lemma relates the rank of a subset of $\mathbb{Z}^n$ to the rank of a morphism.

Lemma 4. *If $h : \Xi^* \to \{a_1, a_2, \dots, a_m\}^*$ is a morphism of rank r, then the set L_h is of rank r.*

Proof. Since $G_h\mathbb{N} \subset L_h \subset G_h\mathbb{Q}$, the claim follows from Lemma 3. $\square$

The last two lemmas of this section investigate several properties of rank, which will serve as a useful tool in the next section devoted to polynomial description of word equations.

Lemma 5. *If $\mathcal{L} = \bigcup_{i \le k} \mathcal{L}_i \subseteq \mathbb{Q}^n$ is a set of rank r such that $\mathcal{L}_i\mathbb{N} \subseteq \mathcal{L}_i$ for each i, then there exists i such that $\mathcal{L}_i$ is of rank r.*

Proof. Since $\mathcal{L} \subseteq \bigcup_{i \le k} \mathcal{L}_i\mathbb{Q}$, there exists i such that $\dim \mathcal{L}_i\mathbb{Q} = r$. Thus $\mathcal{L}_i$ contains a linearly independent set $\gamma_1, \dots, \gamma_r$. By the hypothesis $\mathcal{L}_i\mathbb{N} \subseteq \mathcal{L}_i$, hence $\mathcal{L}_i$ is of rank r by Lemma 3. $\square$

If $\lambda \in \mathbb{Z}^n \setminus \{0\}$, let $\mathcal{N}(\lambda)$ denote the set $\{\alpha \in \mathbb{Z}^n \mid \lambda \cdot \alpha = 0\}$ and put $\mathcal{N}(\lambda)^+ = \mathcal{N}(\lambda) \cap \mathbb{N}^n$.

Lemma 6. *Let $\lambda \in \mathbb{Z}^n \setminus \{0\}$.*

(1) *$\mathcal{N}(\lambda)$ is of rank $n-1$,*
(2) *$\mathcal{N}(\lambda)^+$ is of rank $n-1$ whenever $\lambda_\oplus \neq 0 \neq \lambda_\ominus$.*

Proof. (1) Since $\mathcal{N}(\lambda)$ contains the set $G\mathbb{N}$ for every base $G \subseteq \mathbb{Z}^n$ of the $(n-1)$-dimensional vector space $\{\mathbf{u} \in \mathbb{Q}^n \mid \lambda \cdot \mathbf{u} = 0\}$, the assertion follows from Lemma 3.

(2) As $\lambda_\oplus \neq 0 \neq \lambda_\ominus$, there exists a positive $\mathbf{v}$ such that $\lambda \cdot \mathbf{v} = 0$. Hence for every base $\gamma_1, \dots, \gamma_{n-1} \in \mathbb{Z}^n$ of the vector space $\{\mathbf{u} \in \mathbb{Q}^n \mid \lambda \cdot \mathbf{u} = 0\}$ there is $c \in \mathbb{N}$ such that $G = \{\gamma_1 + c\mathbf{v}, \dots, \gamma_{n-1} + c\mathbf{v}\} \subset \mathbb{N}^n$ is a linearly independent set. Since $G\mathbb{N} \subseteq \mathcal{N}(\lambda)^+$, it remains to apply Lemma 3. $\square$

3. Word equations and polynomials

We review the crucial idea of A. Saarela from [10] that allows to employ linear algebra. As it is based on expressing word equations by polynomials, we recall needed notions from polynomial algebra. $\mathbb{Q}(x)$ denotes the field of fractions of the polynomial ring $\mathbb{Z}[x]$, and $\mathbb{Q}(\mathbf{X})$ the field of fractions of the polynomial ring $\mathbb{Z}[\mathbf{X}] = \mathbb{Z}[X_1, X_2, \dots, X_n]$. Let $\gamma \in \mathbb{N}_0^n$. We denote by $\mathbf{X}^\alpha$ the monomial $\prod_{i=1}^n X_i^{(\alpha)_i} \in \mathbb{Z}[\mathbf{X}]$ and by $\Omega_\gamma : \mathbb{Z}[\mathbf{X}] \to \mathbb{Z}[x]$ the evaluation homomorphism defined by $\Omega_\gamma : X_i \mapsto x^{(\gamma)_i}$, that is,

$$\Omega_\gamma\left(p(X_1, \dots, X_n)\right) = p(x^{(\gamma)_1}, \dots, x^{(\gamma)_n}) .$$

In order to simplify the notation, we shall write $p(\gamma)$ instead of $\Omega_\gamma(p)$.

If we choose (a subset of) $\mathbb{Z}$ as the alphabet, then there is a natural representation of a word $w = a_0 a_1 \cdots a_k$ by the polynomial $P(w) = \sum_{i=0}^k a_i x^i \in \mathbb{Z}[x]$. In this representation, there is an ambiguity caused by trailing zeros. This can be avoided by using only nonzero digits, or by specifying the length of the word.

Representation of an equation is less obvious. Let

$$E = (x_{i_1} x_{i_2} \dots x_{i_r} \,,\, x_{j_1} x_{j_2} \dots x_{j_s}) \tag{$*$}$$

be an equation in n unknowns $\Xi = \{x_1, x_2, \dots, x_n\}$. Then we define

$$S_{E,x_j} = \sum_{a:i_a=j} \prod_{t=1}^{a-1} X_{i_t} - \sum_{a:j_a=j} \prod_{t=1}^{a-1} X_{j_t} \in \mathbb{Z}[\mathbf{X}]$$

(where the empty product is equal to 1). Comparing this definition with the one given in [10], note that, in order to exploit properties of multivariate polynomials, we work in the usual polynomial ring $\mathbb{Z}[\mathbf{X}]$ instead of the (isomorphic) monoid ring $\mathbb{Z}[X; \mathcal{M}]$.

If E and E' are two equations, we will be interested in determinants

$$t_{jk}^{E,E'} := S_{E,x_j} S_{E',x_k} - S_{E',x_j} S_{E,x_k} .$$

Given a length type $\beta \in \mathbb{N}_0^n$, we obtain the polynomial $S_{E,x_j}(\beta) = \Omega_\beta(S_{E,x_j}) \in \mathbb{Z}[x]$. We define

$$\mathcal{S}_E := (S_{E,x_1}, S_{E,x_2}, \dots, S_{E,x_n}) \in \mathbb{Z}[\mathbf{X}]^n ,$$
$$\mathcal{S}_E(\beta) := (S_{E,x_1}(\beta), S_{E,x_2}(\beta), \dots, S_{E,x_n}(\beta)) \in \mathbb{Z}[x]^n .$$

If $h : \Xi^* \to \mathbb{Z}^*$ is a word homomorphism (that is, $\mathbb{Z}^*$ is understood as a free monoid over the alphabet $\mathbb{Z}$), then denote

$$\mathcal{P}(h) = (P(h(x_1)), P(h(x_2)), \dots, P(h(x_n))) \in \mathbb{Z}[x]^n .$$

The point of these definitions is that h, with the length type $L = L(h)$, is a solution of E if and only if

$$\mathcal{S}_E(L) \cdot \mathcal{P}(h) = 0,$$

that is, if $\mathcal{P}(h)$ is a solution of the homogeneous linear equation $\mathcal{S}_E(L)$. The claim is verified by a straightforward check of definitions.

Trivial cases are described in the following lemma.

Lemma 7.

(1) $\mathcal{S}_E = 0$ *if and only if* E *is trivial.*
(2) *If* $\mathcal{S}_E \neq 0$ *and* $\mathcal{S}_E(\beta) = 0$*, then* $(\beta)_i = 0$ *for at least two coordinates* i*.*

Proof. We shall use the notation of $(*)$. The proof consists in verifying the following claims directly from definitions.

If E is trivial, then $\mathcal{S}_E = 0$. Let now E be nontrivial and suppose w.l.o.g. that $r \leq s$. Let $k \geq 1$ be the smallest integer such that $i_k \neq j_k$ or, if the left side of E is a prefix of the right side, $k = r + 1$. Then $S_{E,x_{j_k}} \neq 0$.

If $k = r + 1$, then also $S_{E,x_{j_k}}(\beta) \neq 0$ for any β. If $i_k \neq j_k$ and $(\beta)_{i_k} \neq 0$, then $S_{E,x_{j_k}}(\beta) \neq 0$. Similarly, $S_{E,x_{i_k}}(\beta) \neq 0$ if $(\beta)_{j_k} \neq 0$. $\square$

The following lemma is based on the observation that a solution of rank $n-1$ in fact represents $n-1$ linearly independent solutions.

Lemma 8. *Let* $h : \Xi^* \to \{a_1, a_2, \ldots, a_m\}^*$ *be a solution of rank* $n-1$ *of equations* E *and* E'*, and let* $\alpha \in \mathbb{N}^m$*. Then* $\mathcal{S}_E(L(\vartheta_\alpha \circ h))$ *and* $\mathcal{S}_{E'}(L(\vartheta_\alpha \circ h))$ *are linearly dependent over* $\mathbb{Q}(x)$*.*

Proof. Let $h_i = \tau_i \circ \vartheta_\alpha \circ h$, $i = 1, 2, \ldots, m$, where $\tau_i : \mathrm{alp}(h)^* \to \{0, 1\}^*$ is defined by $\tau_i(a_j) = \delta_{ij}$. Clearly, each h_i is a solution of both E and E'. Observe that the evaluation $x \mapsto 1$ applied to $\mathcal{P}(h_i)$ yields $(\alpha)_i \mathfrak{g}(h)_i$. Since such an evaluation yields $n-1$ linearly independent vectors, we deduce that also among $\mathcal{P}(h_1), \mathcal{P}(h_2), \ldots, \mathcal{P}(h_m)$ there are (at least) $n-1$ linearly independent vectors from $\mathbb{Q}(x)^n$. Since

$$\mathcal{S}_E(L(\vartheta_\alpha \circ h)) \cdot \mathcal{P}(h_i) = \mathcal{S}_{E'}(L(\vartheta_\alpha \circ h)) \cdot \mathcal{P}(h_i) = 0$$

for each $i = 1, 2, \ldots, n-1$, the proof is completed. $\square$

As we need to know more about polynomials $t_{jk}^{E,E'}$ we formulate several straightforward observations on multivariate polynomials.

Lemma 9. *Let $\alpha, \beta \in \mathbb{N}_0^n$, $c \in \mathbb{N}$ and $\gamma \in \mathbb{N}_0^n$. Then:*

(1) $\mathbf{X}^\alpha(\gamma) = x^{\alpha\cdot\gamma}$,
(2) $[\mathbf{X}^\alpha - \mathbf{X}^\beta](\gamma) = x^{\beta\cdot\gamma}\left(x^{(\alpha-\beta)\cdot\gamma} - 1\right)$ *in* $\mathbb{Q}(x)$,
(3) $[\mathbf{X}^\alpha - \mathbf{X}^\beta](\gamma) = 0$ *if and only if* $(\alpha-\beta)\cdot\gamma = 0$,
(4) $\mathbf{X}^{c\alpha} - \mathbf{X}^{c\beta} = (\mathbf{X}^\alpha - \mathbf{X}^\beta)\sum_{i=0}^{c-1}\mathbf{X}^{i\alpha+(c-1-i)\beta}$.

Proof. (1) It follows immediately from definitions.

(2) The equality is a result of a direct computation:

$$[\mathbf{X}^\alpha - \mathbf{X}^\beta](\gamma) = x^{\alpha\cdot\gamma} - x^{\beta\cdot\gamma} = x^{\beta\cdot\gamma}\left(x^{\alpha\cdot\gamma-\beta\cdot\gamma} - 1\right) = x^{\beta\cdot\gamma}(x^{(\alpha-\beta)\cdot\gamma} - 1).$$

(3) As $\mathbb{Q}(x)$ is a field, $x^{\beta\cdot\gamma}(x^{(\alpha-\beta)\cdot\gamma} - 1) = 0$ if and only if $x^{(\alpha-\beta)\cdot\gamma} = 1$, which is equivalent to $(\alpha-\beta)\cdot\gamma = 0$. Now, it remains to apply (2).

(4) An easy computation. □

We say that $\lambda \in \mathbb{Z}^n$ has *coprime coefficients* whenever $\gcd_{i\le n}((\lambda)_i) = 1$.

Lemma 10. *Let $\lambda \in \mathbb{Z}^n \setminus \{0\}$ have coprime coefficients and $\{\gamma_1, \dots, \gamma_{n-1}\} \subseteq \mathcal{N}(\lambda)$ be linearly independent in $\mathbb{Q}^n$. If $\alpha, \beta \in \mathbb{N}_0^n$ such that $[\mathbf{X}^\alpha - \mathbf{X}^\beta](\gamma_i) = 0$ for each $i = 1, \dots, n-1$, then $(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}) \mid (\mathbf{X}^\alpha - \mathbf{X}^\beta)$.*

Proof. By Lemma 9(3), the equality $(\alpha-\beta)\cdot\gamma_i = 0$ holds for each $i < n$. Hence there exists a nonzero rational number c such that $\alpha - \beta = c\lambda$. Note that $c \in \mathbb{Z}$ because λ has coprime coefficients and $\alpha - \beta \in \mathbb{Z}^n$. By symmetry, we may suppose without loss of generality that $c \in \mathbb{N}$. Put

$$\mu = (\min((\alpha)_1, (\beta)_1), \dots, \min((\alpha)_n, (\beta)_n)).$$

It is easy to see that $c\lambda_\oplus = \alpha - \mu$ and $c\lambda_\ominus = \beta - \mu$, thus

$$\mathbf{X}^\alpha - \mathbf{X}^\beta = (\mathbf{X}^{\alpha-\mu} - \mathbf{X}^{\beta-\mu})\mathbf{X}^\mu = (\mathbf{X}^{c\lambda_\oplus} - \mathbf{X}^{c\lambda_\ominus})\mathbf{X}^\mu.$$

Finally, note that $(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}) \mid (\mathbf{X}^{c\lambda_\oplus} - \mathbf{X}^{c\lambda_\ominus})$ by Lemma 9(4). □

The following lemma is a core observation allowing to employ properties of the unique factorization domain $\mathbb{Z}[\mathbf{X}]$.

Lemma 11. *Let $\lambda \in \mathbb{Z}^n \setminus \{0\}$ have coprime coefficients and let $N \subseteq \mathcal{N}(\lambda)$ be of rank $n-1$. Then*

- *$p \in \mathbb{Z}[X]$ satisfies $p(\gamma) = 0$ for all $\gamma \in N$ if and only if $\left(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}\right) \mid p$,*

and

- $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$ *is irreducible.*

Proof. In the proof we shall often implicitly use the well known fact that the evaluation mapping Ω_α is a homomorphism.

By Lemma 9(3), we have $[\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}](\gamma) = 0$ for each $\gamma \in \mathcal{N}(\lambda)$, hence $\left(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}\right) \mid p$ implies $p(\gamma) = 0$ for each $\gamma \in N$.

To prove the other implication, assume to the contrary that there is a polynomial p such that $p(\gamma) = 0$ for each $\gamma \in N$ and p is not divisible by $\left(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}\right)$. Fix such a p with minimal possible number of monomials. More precisely, since every nonzero coefficient of an arbitrary polynomial in $\mathbb{Z}[\mathbf{X}]$ is a sum of copies of either 1 or -1, there exist $s, r \in \mathbb{N}$ and two sequences $(\alpha_i \mid i \le s)$, $(\beta_i \mid i \le r)$ of elements of $\mathbb{N}_0^n$ such that

$$p = \sum_{i \le s} \mathbf{X}^{\alpha_i} - \sum_{i \le r} \mathbf{X}^{\beta_i},$$

and we suppose that $s + r$ is minimal among all polynomials contradicting the assertion. In $p(\gamma)$, we focus on $\mathbf{X}^{\alpha_1}(\gamma)$. If $p(\gamma) = 0$, then there is some $j \le r$ such that $\mathbf{X}^{\alpha_1}(\gamma) = \mathbf{X}^{\beta_j}(\gamma)$. Since $p(\gamma) = 0$ for each $\gamma \in N$, we deduce that $N = \bigcup_{j \le r} N_j$ where

$$N_j = \{\gamma \in N \mid [\mathbf{X}^{\alpha_1} - \mathbf{X}^{\beta_j}](\gamma) = 0\}.$$

By Lemma 5, there exists a linearly independent set $\{\gamma_1, \dots, \gamma_{n-1}\}$ in N_j for some j. Therefore $N_j = N$. Lemma 10 yields that $\left(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}\right) \mid \left(\mathbf{X}^{\alpha_1} - \mathbf{X}^{\beta_j}\right)$, which implies that $p' = p - \left(\mathbf{X}^{\alpha_1} - \mathbf{X}^{\beta_j}\right)$ is not divisible by $\left(\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}\right)$. On the other hand, $p'(\gamma) = p(\gamma) - [\mathbf{X}^{\alpha_1} - \mathbf{X}^{\beta_j}](\gamma) = 0$ for every $\gamma \in N = N_j$, a contradiction to the minimality of $r + s$.

It remains to prove irreducibility of $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$. Suppose that $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus} = gh$ and let

$$N_g = \left\{\gamma \mid g(\gamma) = 0\right\}, \qquad N_h = \left\{\gamma \mid h(\gamma) = 0\right\}.$$

Clearly, for each $\gamma \in \mathcal{N}(\lambda)$ either $g(\gamma) = 0$ or $h(\gamma) = 0$. Hence $\mathcal{N}(\lambda) = N_g \cup N_h$ and at least one of the two sets is of rank $n - 1$ by Lemma 5. By the first part of the proof, we have that $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$ divides either g or h, which we wanted to show. $\square$

An immediate consequence of the last result and Lemma 9(4) is that the polynomial $\mathbf{X}^{\lambda_0} - \mathbf{X}^{\lambda_1}$ is irreducible if and only if $\lambda_0 \cdot \lambda_1 = 0$ and $\lambda_0 - \lambda_1$ has coprime coefficients.

Combining the previous results on multivariate polynomials with Lemma 4, we obtain the following observation.

Lemma 12. *Let $h : \Xi^* \to \{a_1, a_2, \dots, a_m\}^*$ be a solution of rank $n - 1$ of equations E and E'. Then, for each $j, k = 1, 2, \dots, n$, the determinant $t_{jk}^{E,E'}$ is divisible by $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$*

where λ has coprime coefficients and $\mathcal{N}(\lambda) = \Gamma_h$. In particular, $\mathcal{S}_E(\beta)$ and $\mathcal{S}_{E'}(\beta)$ are linearly dependent over $\mathbb{Q}(x)$ for each $\beta \in \Gamma_h \cap \mathbb{N}_0^n$.

Proof. Let $t_{jk} = t_{jk}^{E,E'}$. By Lemma 8,

$$t_{jk}(L(\vartheta_\alpha \circ h)) = 0$$

for each $j, k = 1, 2, \ldots, n$ and each $\alpha \in \mathbb{N}^m$. By Lemma 4, the set L_h is a subset of Γ_h of rank $n-1$, and Lemma 11 implies that $t_{jk} \in \mathbb{Z}[\mathbf{X}]$ is divisible by $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$. Therefore $t_{jk}(\beta) = 0$ for each $\beta \in \Gamma_h \cap \mathbb{N}_0^n$. This concludes the proof. □

Let p be a polynomial in $\mathbb{Z}[\mathbf{X}]$ such that

$$p = \sum_{i \in I} r_i$$

where $r_i = \pm \mathbf{X}^{\alpha_i}$ with $\alpha_i \in \mathbb{N}_0^n$ and $r_i \neq -r_j$ for $i \neq j$. We say that the monomial r_j, $j \in I$, is *minimal* in p, if it has no divisor r_i, $i \in I$ with $i \neq j$, that is, if α_j is a minimal element of $\{\alpha_i \mid i \in I\}$ with respect to the usual product order on $\mathbb{N}_0^n$.

We shall need the following little combinatorial fact.

Lemma 13. *Let A be a set and let S be a subset of A^k such that for each $j \leq k$ there are vectors $s_1, s_2 \in S$ satisfying $(s_1)_j \neq (s_2)_j$ and $(s_1)_{j'} = (s_2)_{j'}$ for each $j' \neq j$. Then S has cardinality at least $k+1$.*

Proof. Proceed by induction. For $k = 1$, the claim is obvious. Let now $k > 1$ and consider the set $S' \subset A^{k-1}$ resulting from S by projection on first $k-1$ coordinates. By assumption, the projection is not injective, hence $|S'| < |S|$. It is easy to see that S' satisfies the hypothesis, which implies $|S'| \geq k$, and the proof is complete. □

Theorem 14. *Let*

$$p = \left(\mathbf{X}^{\lambda_0^{(1)}} - \mathbf{X}^{\lambda_1^{(1)}}\right)\left(\mathbf{X}^{\lambda_0^{(2)}} - \mathbf{X}^{\lambda_1^{(2)}}\right)\cdots\left(\mathbf{X}^{\lambda_0^{(k)}} - \mathbf{X}^{\lambda_1^{(k)}}\right)\sum_{j \in J} e_j \mathbf{X}^{\alpha_j}, \qquad (**)$$

where $\left(\mathbf{X}^{\lambda_0^{(i)}} - \mathbf{X}^{\lambda_1^{(i)}}\right)$ are distinct irreducible polynomials, all $\lambda_b^{(i)}$ are non-zero elements of $\mathbb{N}_0^n$, $e_j = \pm 1$ and $\alpha_j \neq \alpha_{j'}$ for any $j, j' \in J$ such that $e_j \neq e'_j$.

Then p contains at least $k+1$ minimal monomials.

Proof. Let $t \in \mathbb{Q}_+^n$. If $t \cdot \lambda_0^{(i)} \neq t \cdot \lambda_1^{(i)}$ for all $i = 1, 2, \ldots, k$, then we say that t is a *separating type*. The *profile* of a separating type t is the k-tuple $Z(t) := (z_1, \ldots, z_k) \in \{1, -1\}^k$, where

$$z_i := \operatorname{sgn}\left(t \cdot \lambda_0^{(i)} - t \cdot \lambda_1^{(i)}\right).$$

For a separating type t we define $b_{i,t} \in \{0,1\}$, $i = 1, 2, \ldots, k$, $j_t \in J$, and

$$\rho_t := \sum \lambda^{(i)}_{b_{i,t}} + \alpha_{j_t}$$

so that $t \cdot \lambda^{(i)}_{b_{i,t}} < t \cdot \lambda^{(i)}_{1-b_{i,t}}$ for each i, and $t \cdot \alpha_{j_t} \le t \cdot \alpha_j$ for all $j \in J$. Clearly, $t \cdot \rho_t \le t \cdot \beta$ for any monomial $\pm \mathbf{X}^\beta$ resulting from the expansion of (∗∗). Moreover, if $t \cdot \rho_t = t \cdot \beta$, then

$$\beta = \sum \lambda^{(i)}_{b_{i,t}} + \alpha_k$$

for a suitable $k \in J$ and $t \cdot \alpha_{j_t} = t \cdot \alpha_k$. Therefore, either $\alpha_{j_t} = \alpha_k$, or α_{j_t} and α_k are incomparable. We conclude that for a separating type t, the monomial $\mathbf{X}^{\rho_t}$ is minimal. Also $\rho_{t_1} \ne \rho_{t_2}$ if t_1 and t_2 are separating types with different profiles, since then $t_1 \cdot \rho_{t_1} < t_1 \cdot \rho_{t_2}$.

For every $i \le k$ put $\lambda^{(i)} = \lambda^{(i)}_0 - \lambda^{(i)}_1$ and note that $\lambda^{(i)}$ has coprime coefficients, $\lambda^{(i)}_0 = \lambda^{(i)}_\oplus$, and $\lambda^{(i)}_1 = \lambda^{(i)}_\ominus$. It remains to show that there exist at least $k+1$ separating types with distinct profiles.

First, we show for each $j \le k$ that there exists a vector $t \in \mathbb{N}^n$ such that $t \cdot \lambda^{(j)}_0 = t \cdot \lambda^{(j)}_1$, while $t \cdot \lambda^{(i)}_0 \ne t \cdot \lambda^{(i)}_1$ for every $i \ne j$. Assume to the contrary that for every $t \in \mathcal{N}\left(\lambda^{(j)}\right)^+$ there exists $i \ne j$ such that $t \cdot \lambda^{(i)}_0 = t \cdot \lambda^{(i)}_1$. Thus $\mathcal{N}(\lambda^{(j)})^+ \subseteq \bigcup_{i \ne j} V_i$ where $V_i = \left(\mathcal{N}(\lambda^{(i)}) \cap \mathcal{N}(\lambda^{(j)})\right)\mathbb{Q}$, $i \ne j$. Since both λ_i and λ_j, $i \ne j$, have coprime coefficients and $\lambda_i \ne \lambda_j$, each V_i, $i \ne j$, is an $(n-2)$-dimensional subspace. By Lemma 6(2), $\mathcal{N}\left(\lambda^{(j)}\right)^+$ is of rank $n-1$, a contradiction.

Therefore there is a neighborhood of t in $\mathbb{Q}^n_+$ such that $t' \cdot \lambda^{(i)}_0 \ne t' \cdot \lambda^{(i)}_1$, $i \ne j$, for each t' from the neighborhood. This implies that $(Z(t_+))_j = 1$, $(Z(t_-))_j = -1$, and $(Z(t_+))_i = (Z(t_-))_i$, $i \ne j$, for suitable t_+ and t_- from the neighborhood. The proof is completed by Lemma 13. □

4. Independent systems of word equations

In this section, we apply our findings to the question of independence of word equations.

Two systems of equations are *equivalent* if they have the same set of solutions (of all ranks). A system is *independent* if it is not equivalent to any of its proper subsystems. The *compactness property*, proved in [1] and [4], states that each system of word equations over n unknowns contains an equivalent finite subsystem. However, very little is known about the possible size of independent systems. There is a lower bound $\Omega(n^4)$ by a direct construction in [8], but a nearly complete lack of upper bounds. In fact, results from [10] discussed here are the only upper bounds known. Note that they depend on the length of equations and apply only to solutions of rank $n-1$. We therefore say that a system of equations T over n variables is *strongly independent* if any proper subsystem of T has

a solution of rank $n-1$ that is not a solution of T. We can now formulate our goal as to find an upper bound for the size of a strongly independent system.

Let us recall that $\Gamma_h = G_h\mathbb{Q}$ for each morphism h. We say that morphisms h and h' satisfying $\Gamma_h = \Gamma_{h'}$ are *linearly equivalent*. Note that h and $\vartheta_\alpha \circ h$ are linearly equivalent for each $\alpha \in \mathbb{N}^m$. Our goal can be achieved by bounding the number of pairwise linearly nonequivalent solutions of rank $n-1$ of two equations.

Let us first look at erasing solutions. Let $\delta_k(E)$ denote the equation in $n-1$ unknowns resulting from E by erasing the variable x_k.

Lemma 15. *Let h be an erasing solution of rank $n-1$ of E. Then $h(x_k)$ is the empty word for exactly one k, and $\delta_k(E)$ is trivial. Moreover, $\Gamma_h = \mathcal{N}(e_k)$ where e_k is the canonical basis vector defined by $(e_k)_i = \delta_{ki}$.*

Proof. If $h(x_k)$ is the empty word, then $\mathfrak{g}(h)_i \cdot e_k = 0$ for each i. Since h is erasing and of rank $n-1$, we deduce that h erases exactly one k and $\Gamma_h = \mathcal{N}(e_k)$. The restriction of h on $\Xi \setminus \{x_k\}$ is a solution of rank $n-1$ of an equation $\delta_k(E)$ over $n-1$ variables, therefore $\delta_k(E)$ is trivial by the defect effect (see Remark 1). $\square$

We have the following consequence.

Lemma 16. *Let h and h' be linearly nonequivalent erasing solutions of rank $n-1$ of equations E_1 and E_2. Then E_1 and E_2 are equivalent.*

Proof. By Lemma 15, h and h' erase different variables x_a and x_b respectively, and both $E_i = (u_i, v_i)$, $i = 1, 2$, are of the form

$$u_i = r_0^{(i)}\, x_{j_1^{(i)}}\, r_1^{(i)}\, x_{j_2^{(i)}}\, r_2^{(i)} \cdots x_{j_{m_i}^{(i)}}\, r_{m_i}^{(i)}\,,$$
$$v_i = z_0^{(i)}\, x_{j_1^{(i)}}\, z_1^{(i)}\, x_{j_2^{(i)}}\, z_2^{(i)} \cdots x_{j_{m_i}^{(i)}}\, z_{m_i}^{(i)}$$

where $j_k^{(i)} \notin \{a, b\}$, $i = 1, 2$, $k \le m_i$, and all $r_k^{(i)}$, $z_k^{(i)}$ are words in $\{x_a, x_b\}^*$ such that $|r_k^{(i)}|_{x_a} = |z_k^{(i)}|_{x_a}$ and $|r_k^{(i)}|_{x_b} = |z_k^{(i)}|_{x_b}$ for each $r_k^{(i)}$, $z_k^{(i)}$ with $k \le m_i$. Therefore both E_1 and E_2 are equivalent to $x_a x_b = x_b x_a$. $\square$

Note the following fact.

Lemma 17. *Let E and E' be strongly independent. Then there are $k, l = 1, 2, \ldots, n$ such that $t_{kl}^{E,E'} \neq 0$.*

Proof. Independence implies that E and E' are both nontrivial. If $t_{kl}^{E,E'} = 0$ for all $k, l = 1, 2, \ldots, n$, then $\mathcal{S}_E(\beta)$ and $\mathcal{S}_{E'}(\beta)$ are linearly dependent for each β. By Lemma 7 and Lemma 15, $\mathcal{S}_E(L(h))$ and $\mathcal{S}_{E'}(L(h))$ are both nonzero for any morphism h of rank

$n-1$. Therefore $\mathcal{S}_E(L(h)) \cdot \mathcal{P}(h) = 0$ if and only if $\mathcal{S}_{E'}(L(h)) \cdot \mathcal{P}(h) = 0$, and E and E' are not strongly independent. □

The next observation is a variant of a similar claim in the proof of [10, Theorem 5.3]:

Lemma 18. *For every pair of equations E and E' and every $1 \le k < \ell \le n$, the polynomial $t_{k\ell}^{E,E'}$ contains at most $2(|E|_{x_k} + |E|_{x_\ell})$ minimal monomials.*

Proof. Let

$$E = (x_{i_1} x_{i_2} \dots x_{i_r}\,,\, x_{j_1} x_{j_2} \dots x_{j_s}), \qquad E' = (x_{i'_1} x_{i'_2} \dots x_{i'_{r'}}\,,\, x_{j'_1} x_{j'_2} \dots x_{j'_{s'}}),$$

and let

$$Q_k^+ = \sum_{a:i'_a=k} \prod_{t=1}^{a-1} X_{i'_t}\,, \qquad Q_k^- = -\sum_{a:j'_a=k} \prod_{t=1}^{a-1} X_{j'_t}\,,$$

$$Q_\ell^+ = \sum_{a:i'_a=\ell} \prod_{t=1}^{a-1} X_{i'_t}\,, \qquad Q_\ell^- = -\sum_{a:j'_a=\ell} \prod_{t=1}^{a-1} X_{j'_t}\,,$$

so that $S_{E',x_k} = Q_k^+ + Q_k^-$ and $S_{E',x_\ell} = Q_\ell^+ + Q_\ell^-$. Then

$$t_{k\ell}^{E,E'} = S_{E,x_k} Q_\ell^+ + S_{E,x_k} Q_\ell^- + S_{E,x_\ell} Q_k^+ + S_{E,x_\ell} Q_k^-\,.$$

Since monomials in Q_ℓ^+ are totally ordered by divisibility, there is at most one minimal monomial in $\mu\, Q_\ell^+$, for each monomial μ in S_{E,x_k}. Analogous arguments hold for all four summands in the above expression of $t_{k\ell}^{E,E'}$. Since S_{E,x_k} contains $|E|_{x_k}$ monomials and S_{E,x_ℓ} contains $|E|_{x_\ell}$ monomials, the proof is completed. □

We can now prove the desired bounds.

Theorem 19. *Let E, E' be strongly independent equations in n unknowns and m be the number of their pairwise linearly nonequivalent solutions of rank $n-1$. Then*

(1) $m \le |E| + |E'|$,
(2) *there exist indices $1 \le k < \ell \le n$ such that $m \le 2(|E|_{x_k} + |E|_{x_\ell})$.*

Proof. Using Lemma 17, let $h_1, h_2, \dots, h_m$ be pairwise linearly nonequivalent solutions of rank $n-1$ of E and E'. By Lemma 16, we can suppose that $h_1, h_2, \dots, h_{m-1}$ are nonerasing. For each $i = 1, 2, \dots, m-1$, let $\lambda^{(i)}$ be a vector with coprime coefficients such that $\Gamma_{h_i} = \mathcal{N}\left(\lambda^{(i)}\right)$. Let $1 \le k < \ell \le n$ be such that $t = t_{k\ell}^{E,E'} \ne 0$.

Since $\mathbf{X}^{\lambda_{\oplus}^{(i)}} - \mathbf{X}^{\lambda_{\ominus}^{(i)}}$, $i = 1, 2, \ldots, m$, are pairwise non-associated irreducible polynomials, and $(\mathbf{X}^{\lambda_{\oplus}^{(i)}} - \mathbf{X}^{\lambda_{\ominus}^{(i)}}) \mid t$ by Lemma 12, the product $\prod_{i=1}^{m}(\mathbf{X}^{\lambda_{\oplus}^{(i)}} - \mathbf{X}^{\lambda_{\ominus}^{(i)}})$ is a divisor of the polynomial t. If $\lambda_{\oplus}^{(i)} = 0$ or $\lambda_{\ominus}^{(i)} = 0$, then $\Gamma_{h_i} = \mathcal{N}\left(\lambda^{(i)}\right)$ implies that the morphism h_i is erasing.

(1) As degree of every factor is positive and $\deg(t) \leq |E| + |E'|$, the number of pairwise linearly nonequivalent solutions of rank $n-1$ is bounded by $|E| + |E'|$.

(2) Furthermore, all factors $\mathbf{X}^{\lambda_{\oplus}^{(i)}} - \mathbf{X}^{\lambda_{\ominus}^{(i)}}$, $i = 1, 2 \ldots, m-1$, satisfy hypotheses of Theorem 14, which implies that t contains at least m minimal monomials. The proof is completed by Lemma 18. □

As anticipated in [10, p. 16], the improvement given by Theorem 19 with respect to [10, Theorem 5.3.] has consequences for the size of strongly independent systems of equations. The following lemma shows the main idea (cf. Theorem 3.5 of [10]).

Lemma 20. *Let h be a solution of rank $n-1$ of nontrivial equations E and E'. Let h' be a solution of E of rank $n-1$ that is not a solution of E'. Then h and h' are not linearly equivalent.*

Proof. Suppose that h and h' are linearly equivalent. Then $L(h') \in \Gamma_h$, and $\mathcal{S}_E(L(h'))$ and $\mathcal{S}_{E'}(L(h'))$ are linearly dependent by Lemma 12. Since h' is of rank $n-1$, at most one letter can be erased, which implies that both $\mathcal{S}_E(L(h'))$ and $\mathcal{S}_{E'}(L(h'))$ are nonzero by Lemma 7. Then $\mathcal{S}_E(L(h')) \cdot \mathcal{P}(h') = 0$ implies $\mathcal{S}_{E'}(L(h')) \cdot \mathcal{P}(h') = 0$, a contradiction. □

Remark 2. Let us stress the message of the previous lemma. Independence of equations is defined by distinct sets of solutions. Lemma 20, however, shows that the strong independence has more linear algebraic flavor: independent equations are distinguished not only by different solutions h and h' but also by different spaces Γ_h and $\Gamma_{h'}$. More precisely, if two equations have a common solution h, then they are equivalent for length types from the whole Γ_h.

The following example (suggested by Aleksi Saarela) shows that this is not a vacuous property, since an equation can have (even infinitely many) different but linearly equivalent principal solutions. Consider the well known conjugacy equation (xz, zy). It has infinitely many distinct principal solutions

$$g_i : x \mapsto ab, \quad y \mapsto ba, \quad z \mapsto (ab)^i a,$$

but $\Gamma_{g_i} = \mathcal{N}((1, -1, 0))$ for all i.

We are ready to prove the following bounds.

Theorem 21. *Let $T = \{E_1, E_2, \ldots, E_m\}$ be a strongly independent system. Then $m \leq |E_1| + |E_2| + 2$ and there are $1 \leq k < \ell \leq n$ such that $m \leq 2(|E_1|_{x_k} + |E_1|_{x_\ell}) + 2$.*

If T has a solution of rank $n-1$, then $m \leq |E_1|+|E_2|+1$ and there are $1 \leq k < \ell \leq n$ such that $m \leq 2(|E_1|_{x_k} + |E_1|_{x_\ell}) + 1$.

Proof. For each $i = 1, \dots, m$, the system $T \setminus \{E_i\}$ has a solution φ_i of rank $n-1$ that is not a solution of E_i. Let $m \geq 3$ (otherwise the claim is trivial), and let $1 \leq i, j, k \leq m$ be three distinct numbers. Then φ_i is a solution of E_j and E_k, while φ_j is a solution of E_k but not of E_j. Lemma 20 implies that φ_i and φ_j are not linearly equivalent. Therefore E_1 and E_2 have $m-2$ common solutions φ_i, $i = 3, 4, \dots, m$, of rank $n-1$, pairwise linearly nonequivalent. The first two bounds now follow from Theorem 19.

Let T have a solution φ_0. Lemma 20 again implies that φ_0 is not linearly equivalent to any of φ_i, $i = 1, 2, \dots, m$, and we have the second claim. □

Recall that an equation (u, v) is *balanced*, if $|u|_x = |v|_x$ for each $x \in \Xi$. If E is not balanced and h is a solution of rank $n-1$, then Γ_h is uniquely determined by the length constraints induced by the equation. This implies that strongly independent systems contain balanced equations only. This was first proved in [6] for equations in three unknowns. In [10], the result was reproved and generalized to the form presented here.

Our approach and notation allows to characterize balanced equations by the following simple formula.

Lemma 22. *E be a balanced equation in n unknowns if and only if*

$$(S_{E,x_1}, S_{E,x_2}, \dots, S_{E,x_n}) \cdot (X_1 - 1, X_2 - 1, \dots, X_n - 1) = 0.$$

Proof. Let

$$E = (x_{i_1} x_{i_2} \dots x_{i_r}\ ,\ x_{j_1} x_{j_2} \dots x_{j_s})\,.$$

Then

$$S_{E,x_\ell}(X_\ell - 1) = \left(\sum_{a:i_a=\ell} \prod_{t=1}^{a-1} X_{i_t} - \sum_{a:j_a=\ell} \prod_{t=1}^{a-1} X_{j_t} \right) (X_\ell - 1) =$$
$$= \left(\sum_{a:i_a=\ell} \prod_{t=1}^{a} X_{i_t} - \sum_{a:j_a=\ell} \prod_{t=1}^{a} X_{j_t} \right) - \left(\sum_{a:i_a=\ell} \prod_{t=1}^{a-1} X_{i_t} - \sum_{a:j_a=\ell} \prod_{t=1}^{a-1} X_{j_t} \right).$$

This implies that, for each $a < r$ the monomial $\mu = \prod_{t=1}^{a} X_{i_t}$ vanishes in

$$(S_{E,x_1}, S_{E,x_2}, \dots, S_{E,x_n}) \cdot (X_1 - 1, X_2 - 1, \dots, X_n - 1)$$

since μ is contained in $S_{E,x_k}(X_k - 1)$ and $-\mu$ is contained in $S_{E,x_{k+1}}(X_{k+1} - 1)$. Similarly, the monomial $\prod_{t=1}^{a} X_{j_t}$ vanishes for each $a < s$. Therefore

$$(S_{E,x_1}, S_{E,x_2}, \ldots, S_{E,x_n}) \cdot (X_1 - 1, X_2 - 1, \ldots, X_n - 1) = \prod_{t=1}^{s} X_{i_t} - \prod_{t=1}^{r} X_{j_t},$$

which is zero if and only if the equation is balanced. □

Of special interest is the case of three variables, and that for two reasons. First, because this is the simplest case for which there is no bound known independent of the length of equations. In other words, it is an open question whether independent systems of equations over three variables can be arbitrarily large. Second, because for $n = 3$, solutions of rank $n-1$ are precisely all nontrivial (that is, nonperiodic) solutions. Lemma 22 implies the following property of balanced equations over three variables.

Theorem 23. *Let E_1 and E_2 be balanced equations in variables $\{x_1, x_2, x_3\}$. Then*

$$(t_{23}, t_{31}, t_{12}) = t\,(X_1 - 1, X_2 - 1, X_3 - 1)$$

for some polynomial $t \in \mathbb{Z}[X_1, X_2, X_3]$, where $t_{ij} = t_{ij}^{E_1,E_2}$.

Proof. Work in the vector space $\mathbb{Q}(X_1, X_2, X_3)^3$. Let $\mathbf{v}_i = (S_{E_i,x_1}, S_{E_i,x_2}, S_{E_i,x_3})$, $i = 1, 2$. The claim holds for $t = 0$ if $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent. Otherwise, the cross product $(t_{23}, t_{31}, t_{12}) = \mathbf{v}_1 \times \mathbf{v}_2$ is equal to $t\,(X_1 - 1, X_2 - 1, X_3 - 1)$, $t \in \mathbb{Q}(X_1, X_2, X_3)$, by Lemma 22. Since $t_{23}, t_{31}, t_{12} \in \mathbb{Z}[X_1, X_2, X_3]$, it is easy to see that also $t \in \mathbb{Z}[X_1, X_2, X_3]$. □

For independent systems over three unknowns we have the following result (compare with [10, Corollary 6.4]).

Theorem 24. *Let $E_1, \ldots, E_m$, $m \geq 2$, be an independent system of equations in three unknowns having a nonperiodic solution. Then*

(1) *all E_i are balanced,*
(2) *$m \leq |E_1| + |E_2| + 1$, and*
(3) *$m \leq 2(|E_1|_x + |E_1|_y) + 1$ for any pair x, y of unknowns.*

Proof. Suppose that $E_i = (u_i, v_i)$ is not balanced and let φ_j be a solution of $T \setminus \{E_j\}$, with $j \neq i$, that is not a solution of $E_j = (u_j, v_j)$. The constraint $|h'(u_i)| = |h'(v_i)|$ imposed on any solution h' of E_i implies that all nonperiodic solutions of E_i are linearly equivalent. By Lemma 20, each nonperiodic solution of E_i is a solution of whole T. Therefore φ_j is periodic and $|\varphi_j(u_j)| \neq |\varphi_j(v_j)|$. We deduce that the length type of any common solution of E_i and E_j has to satisfy two independent linear constraints, a contradiction with the existence of a common nonperiodic solution. This proves that all E_i are balanced.

Since balanced equations agree on periodic solutions (each periodic morphism is a solution) and T is independent, we obtain that T is strongly independent. Theorem 21 implies (2), and also (3) for some pair of unknowns. From proofs of Theorem 21 and Theorem 19 we can see that the choice of x_k and x_ℓ for which (3) holds depends solely on the condition $t_{k\ell} \neq 0$. Theorem 23 implies that we can choose any pair of unknowns. □

We conclude by an example, which shows that our results allow to obtain in a simple way concrete information about particular cases. Consider the following system of two independent equations in three unknowns $x = x_1$, $y = x_2$, and $z = x_3$:

$$E_1 = (xyxz, zxyx),$$
$$E_2 = (xyxxz, zxxyx).$$

We denote $X = X_1$, $Y = X_2$ and $Z = X_3$ and calculate

$$(S_{E_1,x}, S_{E_1,y}, S_{E_1,z}) = \left(1 + XY - Z - XYZ,\ X - XZ,\ X^2Y - 1\right),$$
$$(S_{E_2,x}, S_{E_2,y}, S_{E_2,z}) = \left(1 + XY + X^2Y - Z - ZX - X^2YZ,\ X - X^2Z,\ X^3Y - 1\right)$$

and

$$(t_{23}, t_{31}, t_{12}) = X(X^2Y - Z)(X - 1, Y - 1, Z - 1).$$

The polynomial $t = X(X^2Y - Z)$ characterizes possible nonperiodic common solutions of E_1 and E_2. Note that the bound of Theorem 19 comes from the bound on the number of hyperplanes covering length types of solutions of rank $n-1$. In other words, this is the number of factors of the form $\mathbf{X}^{\lambda_\oplus} - \mathbf{X}^{\lambda_\ominus}$ dividing the corresponding determinant. In the present example, there is only one such factor, namely $(X^2Y - Z)$. This means that each possible common solution h of E_1 and E_2 of rank two must satisfy $2|h(x)| + |h(y)| = |h(z)|$. Whether such a solution exists can be checked by standard means.

References

[1] M.H. Albert, J. Lawrence, A proof of Ehrenfeucht's conjecture, Theoret. Comput. Sci. 41 (1985) 121–123.

[2] Christian Choffrut, Juhani Karhumäki, Combinatorics of words, in: Handbook of Formal Languages, vol. 1, Springer, Berlin, 1997, pp. 329–438.

[3] Elena Czeizler, Juhani Karhumäki, On non-periodic solutions of independent systems of word equations over three unknowns, Internat. J. Found. Comput. Sci. 18 (4) (2007) 873–897.

[4] V.S. Guba, Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems, Mat. Zametki 40 (3) (1986) 321–324, 428.

[5] Tero Harju, Juhani Karhumäki, Wojciech Plandowski, Independent systems of equations, in: Algebraic Combinatorics on Words, in: Encyclopedia Math. Appl., Cambridge University Press, 2002.

[6] Tero Harju, Dirk Nowotka, On the independence of equations in three variables, Theoret. Comput. Sci. 307 (1) (2003) 139–172.

[7] Juhani Karhumäki, Wojciech Plandowski, On the size of independent systems of equations in semigroups, in: Mathematical Foundations of Computer Science 1994, Košice, 1994, in: Lecture Notes in Comput. Sci., vol. 841, Springer, Berlin, 1994, pp. 443–452.

[8] Juhani Karhumäki, Wojciech Plandowski, On the defect effect of many identities in free semigroups, in: Mathematical Aspects of Natural and Formal Languages, World Scientific, 1994, pp. 225–233.
[9] M. Lothaire, Combinatorics on Words, Cambridge Math. Lib., Cambridge University Press, Cambridge, 1997.
[10] Aleksi Saarela, Systems of word equations, polynomials and linear algebra: a new approach, European J. Combin. 47 (2015) 1–14.