One-Unknown Word Equations and Three-Unknown Constant-Free Word Equations*

Dirk Nowotka[1] and Aleksi Saarela[2]

[1] Department of Computer Science, Kiel University, 24098 Kiel, Germany, dn@zs.uni-kiel.de

[2] Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland, amsaar@utu.fi

Abstract. We prove connections between one-unknown word equations and three-unknown constant-free word equations, and use them to prove that the number of equations in an independent system of three-unknown constant-free equations is at most logarithmic with respect to the length of the shortest equation in the system. We also study two well-known conjectures. The first conjecture claims that there is a constant c such that every one-unknown equation has either infinitely many solutions or at most c. The second conjecture claims that there is a constant c such that every independent system of three-unknown constant-free equations with a nonperiodic solution is of size at most c. We prove that the first conjecture implies the second one, possibly for a different constant.

Keywords: combinatorics on words, word equations, independent systems

1 Introduction

One of the most important open problems in combinatorics on words is the following question: For a given n, what is the maximal size of an independent system of constant-free word equations on n unknowns? It is known that every system of word equations is equivalent to a finite subsystem and, consequently, every independent system is finite. This is known as *Ehrenfeucht's compactness property*. It was conjectured by Ehrenfeucht in a language theoretic setting, formulated in terms of word equations by Culik and Karhumäki [3], and proved by Albert and Lawrence [1] and independently by Guba [6]. If $n > 2$, no finite upper bound for the size of independent systems is known. The largest known independent systems have size $\Theta(n^4)$ [10]. Some related results and variations of the problem are discussed in [11].

The difference between the best known lower and upper bounds is particularly striking in the case of three unknowns: The largest known independent systems

* This work has been supported by the DFG Heisenberg grant 590179 (Dirk Nowotka), the DFG research grant 614256 and the Vilho, Yrjö and Kalle Väisälä Foundation (Aleksi Saarela).

consist of just three equations, but it is not even known whether there exists a constant c such that every independent system has size c or less. When studying independent systems, it is often additionally required that the system has a nonperiodic solution; then the largest known example consists of two equations.

There have been some recent advances regarding this topic. The first nontrivial upper bound was proved by Saarela [14]: The size of an independent system on three unknowns is at most quadratic with respect to the length of the shortest equation in the system. This bound was improved to a linear one by Holub and Žemlička [8]; this is currently the best known result.

Another well-known but less central open problem on word equations is the following question: If a one-unknown word equation with constants has only finitely many solutions, then what is the maximal number of solutions it can have? The answer is at least two, and it has been conjectured that it is exactly two. The best known upper bound, proved by Laine and Plandowski [12], is logarithmic with respect to the number of occurrences of the unknown in the equation. Similar but slightly weaker results were proved in [5] and [4].

In this article we establish a connection between three-unknown constant-free equations and one-unknown equations with constants. This is done by using an old result by Budkina and Markov [2], or a similar result by Spehner [16]. We use this connection to prove two main results.

The first main result is that the size of an independent system of three-unknown equations is logarithmic with respect to the length of the shortest equation in the system. This result is based on the logarithmic bound for the number of solutions of one-unknown equations.

The second main result is an explicit link between two existing conjectures: If there exists a constant c such that the number of solutions of a one-unknown equation is either infinite or at most c, then there exists a constant c' such that the size of an independent system of three-unknown constant-free equations with a nonperiodic solution is at most c'. Furthermore, if $c = 2$, then we can let $c' = 17$. The number 17 here is very unlikely to be optimal, and we expect that the result could be improved by a more careful analysis.

2 Preliminaries

Let Ξ be an alphabet of unknowns and Σ an alphabet of constants. A *constant-free word equation* is a pair $(u, v) \in \Xi^* \times \Xi^*$, and the solutions of this equation are the morphisms $h : \Xi^* \to \Sigma^*$ such that $h(u) = h(v)$. A *word equation with constants* is a pair $(u, v) \in (\Xi \cup \Sigma)^* \times (\Xi \cup \Sigma)^*$, and the solutions of this equation are the constant-preserving morphisms $h : (\Xi \cup \Sigma)^* \to \Sigma^*$ such that $h(u) = h(v)$. We will state many definitions that work for both types of equations.

A solution h is *periodic* if $h(pq) = h(qp)$ for all words p, q in the domain of h, and *nonperiodic* otherwise.

Usually we assume that the alphabet of constants is $\Sigma = \{a, b\}$. The case of a unary alphabet is not interesting, and if there are more than two constant letters, they can be encoded using a binary alphabet. We are specifically interested in equations with constants on one unknown x, and in constant-free equations on three unknowns x, y, z. We use the notation $[u, v, w]$ for the morphism $h : \{x, y, z\}^* \to \Sigma^*$ defined by $(h(x), h(y), h(z)) = (u, v, w)$, and the notation $[u]$ for the constant-preserving morphism $h : (\{x\} \cup \Sigma)^* \to \Sigma^*$ defined by $h(x) = u$. If U is a set of words, we use the notation $[U] = \{[u] \mid u \in U\}$.

Example 1. The equation (xab, bax) has infinitely many solutions $[(ab)^i]$, where $i \geq 0$. The equation $(xaxbab, abaxbx)$ has exactly two solutions, $[\varepsilon]$ and $[ab]$. The constant-free equation (xyz, zyx) has solutions $[(pq)^i p, (qp)^j q, (pq)^k p]$, where $p, q \in \Sigma^*$ and $i, j, k \geq 0$. It has no other nonperiodic solutions.

A set of equations is a *system of equations*. A system $\{E_1, \ldots, E_n\}$ is often written without the braces as $E_1, \ldots, E_n$. A morphism is a solution of this system if it is a solution of every E_i.

The set of all solutions of an equation E is denoted by $\mathrm{Sol}(E)$. Two equations E_1 and E_2 are *equivalent* if $\mathrm{Sol}(E_1) = \mathrm{Sol}(E_2)$. These notions can naturally be extended to systems of equations.

The set of all equations satisfied by a solution h is denoted by $\mathrm{Eq}(h)$. Two solutions h_1 and h_2 are *equivalent* if $\mathrm{Eq}(h_1) = \mathrm{Eq}(h_2)$.

A system of equations $E_1, \ldots, E_n$ is *independent* if it is not equivalent to any of its proper subsystems. Another equivalent definition would be that $E_1, \ldots, E_n$ is independent if there are solutions $h_1, \ldots, h_n$ such that $h_i \in \mathrm{Sol}(E_j)$ if and only if $i \neq j$. The sequence $(h_1, \ldots, h_n)$ is then called an *independence certificate*. (A system is a set, so the order of the equations is not formally specified, but whenever talking about certificates, it is to be understood that the order of the solutions corresponds to the order in which the equations have been written.)

If an independent system has a nonperiodic solution h, it is called *strictly independent*. If $(h_1, \ldots, h_n)$ is its independence certificate, then $(h_1, \ldots, h_n, h)$ is a *strict independence certificate*.

The above definitions can also be stated for infinite systems. However, by Ehrenfeucht's compactness property, every system of word equations is equivalent to a finite subsystem. We will consider only finite systems in this article.

Example 2. The pair of constant-free equations $(xyz, zyx), (xyyz, zyyx)$ is strictly independent. It has a strict independence certificate $([a, b, abba], [a, b, aba], [a, b, a])$. The system of constant-free equations $(x, \varepsilon), (y, \varepsilon), (z, \varepsilon)$ is independent, but not strictly independent. It has an independence certificate $([a, \varepsilon, \varepsilon], [\varepsilon, a, \varepsilon], [\varepsilon, \varepsilon, a])$.

The *length* of an equation $E = (u, v)$ is $|uv|$ and it is denoted by $|E|$. If h is a morphism, we use the notation $h(E) = (h(u), h(v))$. The equation E is *reduced* if u and v do not have a common nonempty prefix or suffix. We can always replace an equation with an equivalent reduced equation.

3 Main Questions

The following question is one of the biggest open problems on word equations:

Question 3. Let S be a strictly independent system of constant-free equations on three unknowns. How large can S be?

The largest known examples are of size two, and it has been conjectured that these examples are optimal. Even the following weaker conjecture is open:

Conjecture 4. There exists a number c such that every strictly independent system of constant-free equations on three unknowns is of size c or less.

We will refer to this conjecture as SIND-XYZ, or as SIND-XYZ(c) for a specific value of c. Currently, the best known result is the following [8]:

Theorem 5. *Every strictly independent system of constant-free equations on three unknowns is of size $O(n)$, where n is the length of the shortest equation.*

Another well-known open problem is the following:

Question 6. Let E be a one-unknown equation with only finitely many solutions. How many solutions can E have?

The best known examples have two solutions, and it has been conjectured that these examples are optimal. Even the following weaker conjecture is open:

Conjecture 7. There exists a number c such that every one-unknown equation has either infinitely many solutions or at most c.

We will refer to this conjecture as SOL-XAB, or as SOL-XAB(c) for a specific value of c. Currently, the best known result is the following [12]:

Theorem 8. *The solution set of a nontrivial one-unknown equation is either of the form $[(pq)^*p]$, where pq is primitive, or a finite set of size at most $8 \log n + O(1)$, where n is the number of occurrences of the unknown.*

As a question between Questions 6 and 3, we can state the following problem and conjecture (we are not aware of any previous research on this problem):

Question 9. Let S be a strictly independent system of one-unknown equations. How large can S be?

Conjecture 10. There exists a number c such that every strictly independent system of one-unknown equations is of size c or less.

We will refer to this conjecture as SIND-XAB, or as SIND-XAB(c) for a specific value of c.

We will prove the following implications between the three conjectures:

$$\text{SOL-XAB} \implies \text{SIND-XAB} \iff \text{SIND-XYZ},$$

or more specifically,

$$\text{SOL-XAB}(c) \implies \text{SIND-XAB}(c) \begin{cases} \Longleftarrow \text{SIND-XYZ}(c) \\ \implies \text{SIND-XYZ}(5c+7). \end{cases}$$

Using the same ideas, we will turn Theorem 8 into a result on constant-free equations on three unknowns.

4 One-Unknown Equations with Constants

In this section we prove that Conjectures SIND-XYZ and SOL-XAB imply Conjecture SIND-XAB. The next lemma is from [5].

Lemma 11. *Let E be a one-unknown equation and let pq be primitive. The set* $\mathrm{Sol}(E) \cap [(pq)^+p]$ *is either* $[(pq)^+p]$ *or has at most one element.*

Lemma 12. *Let $N \geq 3$ and let $E_1, \ldots, E_N$ be a strictly independent system of one-unknown equations. All of these equations have at least N solutions, and at most one of them has infinitely many solutions. If $N \geq 4$, then none of them has infinitely many solutions.*

Proof. If $(h_1, \ldots, h_{N+1})$ is a strict independence certificate, then E_i has solutions h_j for all $j \neq i$. Thus every equation has at least N solutions.

Let one of the equations, say E_1, have infinitely many solutions. By Theorem 8, $\mathrm{Sol}(E_1) = [(pq)^*p]$ for a primitive word pq.

Let another of the equations, say E_2, have infinitely many solutions, so $\mathrm{Sol}(E_2) = [(p'q')^*p']$ for a primitive word $p'q'$. The equations E_1 and E_2 have at least two common solutions h_3, h_4, so $(pq)^i p = (p'q')^{i'} p'$ and $(pq)^j p = (p'q')^{j'} p'$ for some $i < j$ and $i' < j'$. Then $(pq)^{j-i} = (p'q')^{j'-i'}$. By primitivity, $pq = p'q'$, and then $p = p'$ and $q = q'$, so E_1 and E_2 are equivalent, which is a contradiction. This proves that $E_2, \ldots, E_N$ have only finitely many solutions.

If $N \geq 4$, then $\mathrm{Sol}(E_1, E_2) = \mathrm{Sol}(E_2) \cap [(pq)^*p]$ is finite but contains at least three solutions h_3, h_4, h_5, which contradicts Lemma 11, so none of the equations can have infinitely many solutions in this case. □

Theorem 13. *Every strictly independent system of one-unknown equations is of size at most $8 \log n + O(1)$, where n is the length of the shortest equation. Furthermore, Conjecture SOL-XAB(c) implies Conjecture SIND-XAB(c).*

Proof. Follows from Theorem 8 and Lemma 12. □

Lemma 14. *Let $\Sigma = \{a_1, \ldots, a_k\}$ be the alphabet of constants and*

$$\alpha : (\{x\} \cup \Sigma)^* \to \{x, y, z\}^*, \ \alpha(x) = x, \ \alpha(a_i) = y^i z$$

be a morphism. Let $E_1, \ldots, E_N$ be a strictly independent system of equations on $\{x\}$. The system $\alpha(E_1), \ldots, \alpha(E_N)$ of three-unknown constant-free equations is strictly independent.

Proof. Let

$$\beta : \Sigma^* \to \{a, b\}^*, \ \beta(a_i) = a^i b$$

be a morphism. A constant-preserving morphism $h : (\{x\} \cup \Sigma)^* \to \Sigma^*$ is a solution of E_i if and only if the nonperiodic morphism

$$g_h : \{x, y, z\}^* \to \{a, b\}^*, \ g_h(x) = \beta(h(x)), \ g_h(y) = a, \ g_h(z) = b$$

is a solution of $\alpha(E_i)$ (this follows from the fact that $g_h \circ \alpha = \beta \circ h$ and the injectivity of β). So if $(h_1, \ldots, h_{N+1})$ is a strict independence certificate for $E_1, \ldots, E_N$, then $(g_{h_1}, \ldots, g_{h_{N+1}})$ is a strict independence certificate for $\alpha(E_1), \ldots, \alpha(E_N)$. □

Theorem 15. *Conjecture SIND-XYZ(c) implies Conjecture SIND-XAB(c).*

Proof. Follows from Lemma 14. ◻

5 Classification of Solutions

We are interested in strictly independent systems and their certificates. Every morphism in a certificate can be replaced by an equivalent morphism, so it would be beneficial for us if there was a simple subclass of morphisms containing a representative of every equivalence class. In the three-unknown case, this kind of a result follows from a characterization of three-generator subsemigroups of a free semigroup by Budkina and Markov [2], or alternatively from a similar result by Spehner [15, 16]. A comparison of these two results can be found in [7]. The result we present here in Theorem 16 is a simplified version that is perhaps slightly weaker, but sufficiently strong for our purposes and easier to work with.

We define classes of morphisms $\{x, y, z\}^* \to \{a, b, c\}^*$:

$$\begin{aligned}
&\mathcal{A} = \{[a, b, c]\},\\
&\mathcal{B} = \{[a^i, a^j, a^k] \mid i, j, k \geq 0\},\\
&\mathcal{C}_{xyz}(i, j) = \{[a, a^i b a^j, w] \mid w \in \{a, b\}^* \wedge (i = 0 \vee w \in b\{a, b\}^*)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \wedge (j = 0 \vee w \in \{a, b\}^* b)\},\\
&\mathcal{C}_{xyz} = \bigcup_{i,j \geq 0} \mathcal{C}_{xyz}(i, j),\\
&\mathcal{D}_{xyz}(i, j, k, l, m, p, q) = \{[a, a^i b (a^m b)^p a^j, a^k b (a^m b)^q a^l]\},\\
&\mathcal{D}_{xyz} = \bigcup \mathcal{D}_{xyz}(i, j, k, l, m, p, q),
\end{aligned}$$

where the last union is taken over all $i, j, k, l, m \geq 0$ and $p, q \geq 1$ such that $ik = jl = 0$ and $\gcd(p + 1, q + 1) = 1$. If (X, Y, Z) is a permutation of (x, y, z), then $\mathcal{C}_{XYZ}(i, j)$, $\mathcal{C}_{XYZ}$, $\mathcal{D}_{XYZ}(i, j, k, l, m, p, q)$ and $\mathcal{D}_{XYZ}$ are defined similarly, with the images of the unknowns permuted in a corresponding way. For example, in the case of $\mathcal{C}_{XYZ}(i, j)$, X maps to a, Y to $a^i b a^j$, and Z to w. Then we also define

$$\begin{aligned}
\mathcal{C} &= \mathcal{C}_{xyz} \cup \mathcal{C}_{yzx} \cup \mathcal{C}_{zxy} \cup \mathcal{C}_{zyx} \cup \mathcal{C}_{xzy} \cup \mathcal{C}_{yxz},\\
\mathcal{D} &= \mathcal{D}_{xyz} \cup \mathcal{D}_{yzx} \cup \mathcal{D}_{zxy}.
\end{aligned}$$

For $\mathcal{A}$ and $\mathcal{B}$, we do not need to consider different permutations of the unknowns because the images of the unknowns are symmetric. For $\mathcal{D}$, we need only three of the six permutations, because the images of the latter two unknowns are symmetric.

Theorem 16. *Every morphism $\{x, y, z\}^* \to \{a, b, c\}^*$ is equivalent to a morphism in $\mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \cup \mathcal{D}$.*

Proof. Follows from the characterization of Budkina and Markov [2], or alternatively from the characterization of Spehner [16]. □

By the following lemma, we can concentrate on solutions in classes $\mathcal{C}$ and $\mathcal{D}$.

Lemma 17. *A strictly independent system of $N \geq 2$ constant-free equations on $\{x, y, z\}$ has a strict independence certificate in $(\mathcal{C} \cup \mathcal{D})^{N+1}$.*

Proof. Every solution in a certificate can be replaced by an equivalent solution, so the system has a certificate in $(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \cup \mathcal{D})^{N+1}$ by Theorem 16.

The morphism in $\mathcal{A}$ is a solution of only the trivial equations (u, u), and these equations cannot be part of any independent system, so none of the solutions in the certificate can be in $\mathcal{A}$.

It was proved by Harju and Nowotka [7] that if an independent pair of equations has a nonperiodic solution, then both of the equations are balanced, that is, every unknown appears on the left-hand side as often as on the right-hand side. Every morphism in $\mathcal{B}$ is periodic and thus a solution of every balanced equation, so none of the solutions in the certificate can be in $\mathcal{B}$. □

Example 18. The nonperiodic solutions of the equation (xyz, zyx) are of the form $[(pq)^i p, q(pq)^j, (pq)^k p]$. For example, we have the following solutions:

- $[a, b, (ab)^k a] \in \mathcal{C}_{xyz}(0, 0)$ and $[b, a, (ba)^k b] \in \mathcal{C}_{yxz}(0, 0)$ (these are equivalent),
- $[a, b(ab)^j, aba] \in \mathcal{C}_{xzy}(1, 1)$,
- $[a, b(ab)^j, (ab)^k a] \in \mathcal{D}_{xyz}(0, 0, 1, 1, 1, j, k-1)$ ($j, k-1 \geq 1$, $\gcd(j+1, k) = 1$),
- $[(ba)^i b, a, (ba)^k b] \in \mathcal{D}_{yzx}(1, 1, 1, 1, 1, k, i)$ ($i, k \geq 1$, $\gcd(i+1, k+1) = 1$).

6 Class $\mathcal{C}$

In this section we study morphisms in class $\mathcal{C}$. This leads to a natural connection between three-unknown constant-free equations and one-unknown equations with constants.

Lemma 19. *Let E be a nontrivial constant-free equation on $\{x, y, z\}$. There is at most one pair (i, j) such that E has a solution in $\mathcal{C}_{xyz}(i, j)$. For this pair, $i + j \leq |E| - 1$.*

Proof. Let $E = (u, v)$ and $h \in \mathrm{Sol}(E) \cap \mathcal{C}_{xyz}(i, j)$. We can assume that one of the following is true:

1. $v = \varepsilon$.
2. $u = x^k$, $k \geq 1$, and v begins with y.
3. u begins with $x^k y$, $k \geq 1$, and v begins with y.
4. u begins with $x^k z$, $k \geq 1$, and v begins with y.
5. u begins with x and v begins with z.
6. u begins with y and v begins with z.

In all cases, we get either a contradiction or a single possible value for i as follows:

1. $u \neq \varepsilon$, so at least one of $h(x), h(y), h(z)$ is ε. The only possibility is $h(z) = \varepsilon$, and then $i = j = 0$.
2. $h(u) = a^k$ and $h(v)$ contains the letter b, which is a contradiction.
3. $h(u)$ begins with $a^{k+i}b$ and $h(v)$ begins with $a^i b$, which is a contradiction.
4. $h(y)$ must begin with a and thus $h(z)$ must begin with b, so $h(u)$ begins with $a^k b$ and $h(v)$ begins with $a^i b$. Thus $i = k$.
5. $h(z)$ cannot begin with b and thus $h(y)$ must begin with b, so $i = 0$.
6. It is not possible that $h(y)$ would begin with a and $h(z)$ with b, so $h(y)$ must begin with b and $i = 0$.

By looking at the suffixes of u and v, we will similarly see that j is uniquely determined. Moreover, $i + j \leq |E| - 1$. □

Lemma 20. *Let $S = \{E_1, \dots, E_N\}$ be a system of constant-free equations on $\{x, y, z\}$. Let S have a strict independence certificate $(h_1, \dots, h_{N+1}) \in \mathcal{C}_{xyz}^{N+1}$. There is a strictly independent system $E'_1, \dots, E'_N$ of one-unknown equations such that $|E'_n| \leq |E_n|^2$ for all n.*

Proof. The case $N < 2$ is trivial, so let $N \geq 2$. Let i, j be such that $h_{N+1} \in \mathcal{C}_{xyz}(i, j)$. By Lemma 19, $(h_1, \dots, h_N) \in \mathcal{C}_{xyz}(i, j)^N$. Let

$$\alpha : \{x, y, z\}^* \to \{a, b, z\}^*, \ \alpha(x) = a, \ \alpha(y) = a^i b a^j, \ \alpha(z) = z$$

be a morphism and let

$$h'_n : \{a, b, z\}^* \to \{a, b\}^*, \ h'_n(z) = h_n(z)$$

be a constant-preserving morphism. For every n, $h_n = h'_n \circ \alpha$ and $\alpha(E_n)$ is a one-unknown equation with constants. Then $(h'_1, \dots, h'_{N+1})$ is a strict independence certificate of the system $\alpha(E_1), \dots, \alpha(E_N)$. The length of $\alpha(E_n)$ is at most $(i + j + 1)|E_n|$, which is at most $|E_n|^2$ by Lemma 19. □

7 Class $\mathcal{D}$

In this section we study morphisms in class $\mathcal{D}$. This class looks more complicated than class $\mathcal{C}$, but actually there is a lot of structure in the morphisms in $\mathcal{D}$, which allows us to prove stronger results than for $\mathcal{C}$.

Lemma 21. *Let E be a nontrivial constant-free equation on $\{x, y, z\}$. There are i, j, k, l, m, p', q' such that $\mathrm{Sol}(E) \cap \mathcal{D}_{xyz}$ is either $\varnothing$, $\mathcal{D}_{xyz}(i, j, k, l, m, p', q')$, or the union of $\mathcal{D}_{xyz}(i, j, k, l, m, p, q)$ over all $p, q \geq 1$ such that $\gcd(p+1, q+1) = 1$.*

Proof. Let $E = (u, v)$. If $u = \varepsilon$ or $v = \varepsilon$, then $\mathrm{Sol}(E) \cap \mathcal{D}_{xyz} = \varnothing$, so let $u \neq \varepsilon \neq v$. We can assume that E is reduced and write it as

$$(x^{a_0} y_1 x^{a_1} \cdots y_r x^{a_r}, x^{b_0} z_1 x^{b_1} \cdots z_s x^{b_s}),$$

where $y_1, \dots, y_r, z_1, \dots, z_s \in \{y, z\}$. We can also assume that $r, s \geq 2$. Let $h \in \mathrm{Sol}(E) \cap \mathcal{D}_{xyz}$ and

$$h(x) = a,\ h(y_t) = a^{i_t} b (a^m b)^{p_t} a^{j_t},\ h(z_t) = a^{k_t} b (a^m b)^{q_t} a^{l_t},$$

$$(i_t, j_t, p_t) = \begin{cases} (i, j, p) & \text{if } y_t = y, \\ (k, l, q) & \text{if } y_t = z, \end{cases} \qquad (k_t, l_t, q_t) = \begin{cases} (i, j, p) & \text{if } z_t = y, \\ (k, l, q) & \text{if } z_t = z. \end{cases}$$

The left-hand side $h(u)$ begins with $a^{a_0 + i_1} b$ and the right-hand side $h(v)$ begins with $a^{b_0 + k_1} b$, so $a_0 + i_1 = b_0 + k_1$. If $y_1 = z_1$, then $i_1 = k_1$, $a_0 = b_0$, and E is not reduced, a contradiction. Thus $y_1 \neq z_1$ and $i_1 k_1 = ik = 0$. From $a_0 + i_1 = b_0 + k_1$, $i_1 k_1 = 0$, $a_0 b_0 = 0$ it then follows that $k_1 = a_0$ and $i_1 = b_0$. Similarly, by looking at the suffixes of $h(u)$ and $h(v)$ we find out that $y_r \neq z_s$, $l_s = a_r$, and $j_r = b_s$. Thus i, j, k, l are uniquely determined by the equation E.

It must be $\{p_1, q_1\} = \{p, q\}$, and $\gcd(p + 1, q + 1) = 1$, so $p_1 \neq q_1$. If $p_1 < q_1$, then $h(u)$ and $h(v)$ begin with

$$a^{a_0 + i_1} b (a^m b)^{p_1} a^{j_1 + a_1 + i_2} b \qquad \text{and} \qquad a^{b_0 + k_1} b (a^m b)^{p_1 + 1},$$

respectively, so $j_1 + a_1 + i_2 = m$. Similarly, if $p_1 > q_1$, then $l_1 + b_1 + k_2 = m$. Thus $m \in \{j_1 + a_1 + i_2, l_1 + b_1 + k_2\}$. If $j_1 + a_1 + i_2 = m \neq l_1 + b_1 + k_2$, then there are $n \neq m$, $A \geq 1$, $B \geq 0$ such that $h(u)$ and $h(v)$ begin with

$$a^{a_0 + i_1} b (a^m b)^{A(p_1 + 1) + B(q_1 + 1) - 1} a^n b \qquad \text{and} \qquad a^{b_0 + k_1} b (a^m b)^{q_1} a^{l_1 + b_1 + k_2} b,$$

respectively. It must be $A(p_1 + 1) + B(q_1 + 1) = q_1 + 1$. But then $B > 0$ would be a contradiction, and $B = 0$ would contradict $\gcd(p + 1, q + 1) = 1$. Similarly, $j_1 + a_1 + i_2 \neq m = l_1 + b_1 + k_2$ would lead to a contradiction. Thus it must be $j_1 + a_1 + i_2 = m = l_1 + b_1 + k_2$.

We can write

$$h(u) = a^{c_0} b (a^m b)^{A_1(p+1) + C_1(q+1) - 1} a^{c_1} b \cdots b (a^m b)^{A_R(p+1) + C_R(q+1) - 1} a^{c_R},$$
$$h(v) = a^{d_0} b (a^m b)^{B_1(p+1) + D_1(q+1) - 1} a^{d_1} b \cdots b (a^m b)^{B_S(p+1) + D_S(q+1) - 1} a^{d_S},$$

where $c_1, \dots, c_{R-1}, d_1, \dots, d_{S-1} \neq m$. It must be $R = S$, $c_t = d_t$, and

$$A_t(p + 1) + C_t(q + 1) = B_t(p + 1) + D_t(q + 1)$$

for all t. Moreover, all values p, q that satisfy these linear relations lead to a solution of the equation. If there are two linearly independent relations, there are no solutions. If there is one nontrivial relation $A(p + 1) = C(q + 1)$, then there is exactly one solution with $\gcd(p + 1, q + 1) = 1$. If all relations are trivial, all values of p, q satisfy them. This concludes the proof. ☐

The next lemma is a special case of Theorem 5.3 in [14]. Here, the *length type* of a solution h is the vector $(|h(x)|, |h(y)|, |h(z)|)$.

Lemma 22. *The length types of nonperiodic solutions of an independent pair of constant-free equations on three unknowns are covered by a finite union of two-dimensional subspaces of* $\mathbb{Q}^3$.

Lemma 23. *Let E_1, E_2, E_3, E_4 be a system of constant-free equations on $\{x, y, z\}$ with a strict independence certificate $(h_1, h_2, h_3, h_4, h_5)$. At most one of the h_i can be in $\mathcal{D}_{xyz}$.*

Proof. Let $h_r, h_s \in \mathcal{D}_{xyz}$, $r \neq s$. Without loss of generality, let $r, s \geq 3$. Then $h_r, h_s \in \mathrm{Sol}(E_1, E_2) \cap \mathcal{D}_{xyz}$, so the third option of Lemma 21 must be true for this set. We will show that the length types of solutions of E_1, E_2 cannot be covered by finitely many two-dimensional spaces, which contradicts Lemma 22.

The length type of $[a, a^i b(a^m b)^p a^j, a^k b(a^m b)^q a^l] \in \mathrm{Sol}(E_1, E_2) \cap \mathcal{D}_{xyz}$ is

$$(1, i + 1 + (m + 1)p + j, k + 1 + (m + 1)q + l).$$

Here i, j, k, l, m are fixed, but p, q can be arbitrary positive integers such that $\gcd(p + 1, q + 1) = 1$. For every p, there are infinitely many possible values of q, giving infinitely many length types on the line

$$L_p = \{(1, i + 1 + (m + 1)p + j, Z) \mid Z \in \mathbb{Q}\}.$$

The only way to cover these with a finite number of two-dimensional spaces is to have one of them be the unique two-dimensional space containing the whole line. This is true for any p, and different values of p give different spaces, so all length types cannot be covered by finitely many two-dimensional spaces. $\square$

8 Main Results

Putting our results together gives the following theorem, which improves the linear bound of Theorem 5 to a logarithmic one.

Theorem 24. *A strictly independent system of constant-free equations on three unknowns has at most $O(\log n)$ equations, where n is the length of the shortest equation.*

Proof. Let the system be $E_1, \ldots, E_N$, where E_1 is the shortest equation. By Lemma 17, it has a strict independence certificate $(h_1, \ldots, h_{N+1}) \in (\mathcal{C} \cup \mathcal{D})^{N+1}$. By Lemma 23, at most three of the h_i can be in $\mathcal{D}$. Let k of the solutions be in $\mathcal{C}_{xyz}$. If h_1 is one of them, we get a system of size $k - 1$, for which we can use Lemma 20, and then Theorem 13 to conclude that $k = O(\log n)$. Otherwise, we can still use the arguments in the proof of Lemma 20 to turn E_1 into a one-unknown equation E_1' with k solutions. Then, by Theorem 13, either $k = O(\log n)$ or E_1' has infinitely many solutions, but the latter leads to a contradiction like in the proof of Lemma 12. Similarly, we can prove that the number of i such that $h_i \in \mathcal{C}_{XYZ}$ is $O(\log n)$ for all permutations (X, Y, Z) of (x, y, z). $\square$

We say that two words *begin in the same way* if they begin with the same letter or are both empty. We say that equations (u_1, v_1) and (u_2, v_2) *begin in the same way* if either u_1 and u_2 begin in the same way and v_1 and v_2 begin in the same way, or u_1 and v_2 begin in the same way and v_1 and u_2 begin in the same way. Equations *ending the same way* is defined analogously.

Lemma 25. *Let $N \geq 3$ and let $E_1, \dots, E_N$ be a strictly independent system of reduced constant-free equations on $\{x, y, z\}$. All of the equations begin and end in the same way.*

Proof. Assume that all of the equations do not begin and end in the same way. Without loss of generality, we can assume that E_1 and E_2 do not begin in the same way and that they are of the form (xu, yv) and (xu', zv'), respectively. By the well-known graph lemma about word equations, every common solution of these two equations is periodic or maps one of the unknowns to the empty word. The equations E_1 and E_2 have two nonequivalent nonperiodic solutions, and these solutions must map x to the empty word. But all nonperiodic solutions mapping x to the empty word are equivalent, which is a contradiction. $\square$

By Theorem 13, Conjecture SIND-XAB could be replaced by Conjecture SOL-XAB in the next theorem. The constants are probably not optimal.

Theorem 26. *Conjecture SIND-XAB(c) implies Conjecture SIND-XYZ($5c + 7$). In particular, if SIND-XAB(2) is true, then a strictly independent system of constant-free equations on $\{x, y, z\}$ has at most 17 equations.*

Proof. Let $E_1, \dots, E_N$ be a system of reduced constant-free equations on $\{x, y, z\}$ with a strict independence certificate $(h_1, \dots, h_{N+1})$. For an equation $E_m = (u, v)$, at least one of the unknowns appears both at the beginning of u or v and at the end of u or v. By Lemma 25, this unknown does not depend on m. Without loss of generality, we can assume it is z. By Lemma 17, we can assume that $h_n \in \mathcal{C} \cup \mathcal{D}$ for all n. Because $\mathcal{C}_{xyz}(0, 0)$ and $\mathcal{C}_{yxz}(0, 0)$ are the same up to swapping a and b, we can assume that $h_n \notin \mathcal{C}_{yxz}(0, 0)$ for all n.

By Lemma 20 and the assumption about Conjecture SIND-XAB, at most $c + 1$ of the solutions h_n can be in $\mathcal{C}_{xyz}$, and the same is true for the other five permutations of the unknowns. By the assumption about z and the proof of Lemma 19, $\mathrm{Sol}(E_m) \cap \mathcal{C}_{yxz} \subseteq \mathcal{C}_{yxz}(0, 0)$ for all m, so $h_n \notin \mathcal{C}_{yxz}$ for all n. Thus at most $5c + 5$ of the solutions h_n can be in $\mathcal{C}$.

By Lemma 23, at most one of the solutions h_n can be in $\mathcal{D}_{xyz}$, and the same is true for $\mathcal{D}_{yzx}$ and $\mathcal{D}_{zxy}$. Thus at most three of the solutions h_n can be in $\mathcal{D}$.

This proves that the total number of the solutions h_n, which is $N + 1$, cannot be more than $5c + 8$. $\square$

9 Conclusion

We can mention several further research goals. Two obvious ones are improving the constants in Theorem 26, ideally so that Conjecture SIND-XAB(c) implies Conjecture SIND-XYZ(c), and proving Conjecture SOL-XAB or Conjecture SIND-XAB (ideally SOL-XAB(2)), and thus also Conjecture SIND-XYZ. Proving similar results for chains of equations instead of independent systems might be possible (see [11] for definitions).

A different topic would be to study the complexity of determining whether a three-unknown constant-free equation has a nonperiodic solution. This decision problem is known to be in NP [13]. Based on the connection to one-unknown equations, a better result could probably be obtained, because one-unknown equations can be solved efficiently, even in linear time, as proved by Jez [9].

Finally, Question 3 could be studied for more than three unknowns. This is of course a big question, and our techniques do not help here, because they are specific to the three-unknown case.

References

1. Albert, M.H., Lawrence, J.: A proof of Ehrenfeucht's conjecture. Theoret. Comput. Sci. 41(1), 121–123 (1985)
2. Budkina, L.G., Markov, A.A.: F-semigroups with three generators. Mat. Zametki 14, 267–277 (1973)
3. Culik, II, K., Karhumäki, J.: Systems of equations over a free monoid and Ehrenfeucht's conjecture. Discrete Math. 43(2–3), 139–153 (1983)
4. Dąbrowski, R., Plandowski, W.: On word equations in one variable. Algorithmica 60(4), 819–828 (2011)
5. Eyono Obono, S., Goralčík, P., Maksimenko, M.: Efficient solving of the word equations in one variable. In: Proceedings of the 19th MFCS. LNCS, vol. 841, pp. 336–341. Springer (1994)
6. Guba, V.S.: Equivalence of infinite systems of equations in free groups and semigroups to finite subsystems. Mat. Zametki 40(3), 321–324 (1986)
7. Harju, T., Nowotka, D.: On the independence of equations in three variables. Theoret. Comput. Sci. 307(1), 139–172 (2003)
8. Holub, Š., Žemlička, J.: Algebraic properties of word equations. J. Algebra 434, 283–301 (2015)
9. Jez, A.: One-variable word equations in linear time. Algorithmica 74(1), 1–48 (2016)
10. Karhumäki, J., Plandowski, W.: On the defect effect of many identities in free semigroups. In: Paun, G. (ed.) Mathematical aspects of natural and formal languages, pp. 225–232. World Scientific (1994)
11. Karhumäki, J., Saarela, A.: On maximal chains of systems of word equations. Proc. Steklov Inst. Math. 274, 116–123 (2011)
12. Laine, M., Plandowski, W.: Word equations with one unknown. Internat. J. Found. Comput. Sci. 22(2), 345–375 (2011)
13. Saarela, A.: On the complexity of Hmelevskii's theorem and satisfiability of three unknown equations. In: Proceedings of the 13th DLT. LNCS, vol. 5583, pp. 443–453. Springer (2009)
14. Saarela, A.: Systems of word equations, polynomials and linear algebra: A new approach. European J. Combin. 47, 1–14 (2015)
15. Spehner, J.C.: Quelques problémes d'extension, de conjugaison et de présentation des sous-monoïdes d'un monoïde libre. Ph.D. thesis, Univ. Paris (1976)
16. Spehner, J.C.: Les systemes entiers d'équations sur un alphabet de 3 variables. In: Semigroups. pp. 342–357 (1986)