

Word Equations with k th Powers of Variables

Aleksi Saarela

Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland

Abstract

We prove that if the equality of words $x^k u = u_1 x_1^k \cdots u_n x_n^k$ holds for three positive values of k , then it holds for all values of k . As a consequence, if $x^k = x_1^k \cdots x_n^k$ holds for three positive values of k , then the words x, x_1, \dots, x_n are powers of a common word. The most important method in our proofs is to assign numerical values to the letters, and then study the sums of the letters of words and their prefixes.

Keywords: combinatorics on words, word equation

2010 MSC: 68R15

1. Introduction

We say that words x_0, \dots, x_n *commute* if $x_i x_j = x_j x_i$ for all $i, j \in \{0, \dots, n\}$. It is well-known that words commute if and only if they are powers of a common word. It is also known that if two words satisfy a nontrivial relation, then they are powers of a common word, and more generally, if n words satisfy a nontrivial relation, then they can be written as products of $n - 1$ words. This is known as the *defect theorem*, see [7] for a survey.

One of the early results on word equations is the result of Lyndon and Schützenberger [15] that if $x^k = y^m z^n$ for some words x, y, z and numbers $k, m, n \geq 2$, then x, y, z commute. Many generalizations have been studied, see, for example, [14, 22, 2]. We are interested in the generalizations where the right-hand side can have more than two powers, but all exponents are equal (here $k \geq 1$):

$$x_0^k = x_1^k \cdots x_n^k. \quad (1)$$

We are also interested in more general equalities of the form

$$u_0 x_1^k u_1 \cdots x_m^k u_m = v_0 y_1^k v_1 \cdots y_n^k v_n. \quad (2)$$

In particular, we are interested in cases where (1) or (2) holds for many values of k at the same time.

Email address: amsaar@utu.fi (Aleksi Saarela)

Let us mention some connections and applications of (1) and (2). A subset K of a language L is called its *test set* if, for all morphisms f and g , either $f(x) \neq g(x)$ for some $x \in K$ or $f(x) = g(x)$ for all $x \in L$. This means that to check whether f and g agree on L , it is sufficient to check whether they agree on K . There are connections between test sets and the above equalities, as explained in [10]. The equalities also come up when studying pumping properties of formal languages. They are used, for example, in the study of transducers [3, 5]. The equalities (1) are related to the construction of large independent systems of word equations [12, 18]. Finally, there is a connection between Sturmian words and (1) for $k = 1, 2$ [17].

The main question about the equalities (1) is when do they imply that x_0, \dots, x_n commute. Appel and Djorup [1] proved that if $k = n$ in (1), then the words x_0, \dots, x_n must commute. Their result was generalized by Harju and Nowotka [8] for certain equalities which have many different exponents k_1, k_2, \dots instead of just one exponent k . There are many examples of words x_0, \dots, x_n such that $x_i x_j \neq x_j x_i$ for some i, j , but (1) holds for two different values of k . For instance, $(ababa)^k = (ab)^k a^k (ba)^k$ for $k \in \{1, 2\}$. No such examples were found for three different values of k , which led to the following question.

Question 1.1. *If x_0, \dots, x_n are words and $k_1, k_2, k_3 \geq 1$ are different numbers such that (1) holds for $k \in \{k_1, k_2, k_3\}$, then do x_0, \dots, x_n necessarily commute?*

In the case $\{k_1, k_2, k_3\} = \{1, 2, 3\}$, the question appeared in an article by Hakala and Kortelainen [6], a positive answer was explicitly conjectured by Plandowski [18], and a prize for a proof was offered by Holub in 2009¹. The case where one of k_1, k_2, k_3 is 1 was asked in [9], and a positive answer in the case $k_1, k_2, k_3 \geq 2$ was proved in [10] by Holub.

A positive answer to the general version of Question 1.1 was finally presented at the conference STACS 2017 [20]. This article is an extended journal version of that conference paper with more general results. Namely, we consider (2) in the case where $m = 1$ and u_0 is empty (the case where u_1 is empty is of course symmetric). We prove that in this case (2) holds either for at most two positive values of k or for all k . A positive answer to Question 1.1 follows as a consequence. The basic idea is to assign numerical values to the letters in a specific way (so that the sum of the letters of x_1 is zero), and then study the sums of the letters of words and their prefixes. The ideas have already been used to analyze other kinds of problems on word equations as well [16, 21].

We conclude the introduction with a short description of the similarities and differences between the conference version and the journal version. Section 2 is mostly copied from [20], although Lemma 2.2 is stronger than the corresponding result in [20]. The underlying structure of the main proof (which is divided into several lemmas) is roughly similar, but in this article we need some additional considerations and more complicated arguments. In particular, Section 4 and all uses of the polynomials $D_a(w)$ are entirely new.

¹<http://www.karlin.mff.cuni.cz/~holub/soubory/prizeproblem.pdf>

2. Preliminaries

Let Γ be an alphabet. We can assume that Γ is a subset of \mathbb{R} . This allows us to define $\Sigma(w)$ to be the sum of the letters of a word $w \in \Gamma^*$, that is, if $w = a_1 \cdots a_n$ and $a_1, \dots, a_n \in \Gamma$, then $\Sigma(w) = a_1 + \cdots + a_n$. The mapping Σ is a morphism from the free monoid Γ^* to the additive monoid \mathbb{R} . Words w such that $\Sigma(w) = 0$ are called *zero-sum words*.

The notation $a_1 \cdots a_n$ of course means the word consisting of the letters a_1, \dots, a_n and not a product of numbers. When we actually want to compute the product of two numbers, it should be clear from context. If w_1, \dots, w_n are words, we can also use the notation

$$\prod_{i=1}^n w_i = w_1 \cdots w_n$$

for their concatenation.

Whenever the symbol Γ appears in this article, it is always used to denote an alphabet. Occasionally we use other alphabets as well. All of them can be assumed to be subsets of \mathbb{R} . Alphabets are also assumed to be finite.

Let $a_1, \dots, a_k \in \Gamma$. The *prefix sum word* of $w = a_1 \cdots a_k$ is the word $\text{psw}(w) = b_1 \cdots b_k$, where $b_i = \Sigma(a_1 \cdots a_i)$ for all i . Of course, $\text{psw}(w)$ is usually not a word over Γ , but over some other alphabet. The word $\text{psw}(w)$ has the same length as w and the last letter is $\Sigma(w)$.

The mapping psw is injective. It is not a morphism, but we can give a simple formula for the prefix sum word of a product by using the notation $\text{psw}_r(w) = c_1 \cdots c_k$, where $r \in \mathbb{R}$ and $c_i = b_i + r$ for all i . Then, for $w_1, \dots, w_n \in \Gamma^*$,

$$\text{psw}(w_1 \cdots w_n) = \prod_{i=1}^n \text{psw}_{\Sigma(w_1 \cdots w_{i-1})}(w_i).$$

If w_1, \dots, w_{n-1} are zero-sum, then we have the simpler formula

$$\text{psw}(w_1 \cdots w_n) = \prod_{i=1}^n \text{psw}(w_i),$$

so in this case the mapping psw actually does behave like a morphism. For the n th power of a word w , we get the formula

$$\text{psw}(w^n) = \prod_{i=1}^n \text{psw}_{(i-1)\Sigma(w)}(w).$$

If w is zero-sum, then we have $\text{psw}(w^n) = \text{psw}(w)^n$.

Because letters are real numbers, there is a natural order relation for them. The largest and smallest letters in a word w can be denoted by $\max(w)$ and $\min(w)$, respectively. The length of w is denoted by $|w|$, and the number of occurrences of a letter a in w is denoted by $|w|_a$. If $\Gamma = \{a_1, \dots, a_n\}$ and $a_1 <$

$\dots < a_n$, then the *Parikh vector* of $w \in \Gamma^*$ is the vector $\Pi_w = (|w|_{a_1}, \dots, |w|_{a_n})$. The set of letters occurring in w is denoted by $\text{alph}(w)$. The size of a set S is denoted by $|S|$.

Example 2.1. Let $\Gamma = \{a, b, c\}$ and $w = bbcaacc$, where $a = 1$, $b = 2$, and $c = -2$. We have $|w| = 6$, $\max(w) = 2$, $\min(w) = -2$, and $\Pi_w = (3, 2, 2)$. Because $\Sigma(w) = 2+2-2+1+1-2-2 = 0$, w is a zero-sum word. The prefix sum word of w is $\text{psw}(w) = 2423420$, and $\max(\text{psw}(w)) = 4$ and $\min(\text{psw}(w)) = 0$.

When studying words from a combinatorial point of view, the choice of the alphabet is arbitrary (except for the size of the alphabet). Therefore, we can assign numerical values to the letters in any way we like, as long as no two letters get the same value. The next lemma shows in a formal way that, given any word u , the alphabet can be normalized so that u becomes a zero-sum word.

Lemma 2.2. *Let $u \in \Gamma^*$. There exists an alphabet Δ and an isomorphism $h : \Gamma^* \rightarrow \Delta^*$ such that for all $v \in \Gamma^*$, $h(v)$ is zero-sum if and only if Π_v is a scalar multiple of Π_u .*

Proof. If u is empty, it is sufficient to define h so that the image of every letter is positive. Otherwise, let $\Gamma = \{a_1, \dots, a_n\}$ and $a_1 < \dots < a_n$, and let a_j be the first letter of u and let $J = \{1, \dots, n\} \setminus \{j\}$. We can view \mathbb{R} as an infinite-dimensional vector space over \mathbb{Q} , and then there exists a linearly independent set $\{b_i \mid i \in J\} \subseteq \mathbb{R}$. Let

$$b_j = -\frac{1}{|u|_{a_j}} \cdot \sum_{i \in J} |u|_{a_i} b_i.$$

Because $b_j \notin \{b_i \mid i \in J\}$, we can define an alphabet $\Delta = \{b_1, \dots, b_n\}$ of size n , and an isomorphism $h : \Gamma^* \rightarrow \Delta^*$ such that $h(a_i) = b_i$ for all $i \in \{1, \dots, n\}$. Consider the linear mapping

$$\sigma : \mathbb{Q}^n \rightarrow \mathbb{R}, \quad \sigma(q_1, \dots, q_n) = q_1 b_1 + \dots + q_n b_n.$$

Then $\text{Im}(\sigma)$ is the $(n-1)$ -dimensional subspace of \mathbb{R} generated by the linearly independent set $\{b_i \mid i \in J\}$, so $\dim(\text{Ker}(\sigma)) = n - \dim(\text{Im}(\sigma)) = 1$ by the rank-nullity theorem. By the definition of b_j , $\sigma(\Pi_u) = 0$, so $\Pi_u \in \text{Ker}(\sigma)$. The one-dimensional space $\text{Ker}(\sigma)$ is generated by each one of its nonzero elements, in particular by Π_u , so $\text{Ker}(\sigma)$ consists of the scalar multiples of Π_u . For all $v \in \Gamma^*$, $\sigma(\Pi_v) = \Sigma(h(v))$, so $h(v)$ is zero-sum if and only if Π_v is a scalar multiple of Π_u . \square

The above definitions have the following graphical interpretation, which is not necessary for the proofs, but it might be helpful in understanding them (at least it was helpful in inventing the proofs): Let $w = a_1 \dots a_k$. The word $\text{psw}(w)$ (or the word w depending on the point of view) can be represented by a polygonal chain by starting at the origin, moving a_1 units up and one unit to the right, a_2 units up and one unit to the right, and so on. If $\text{psw}(w) = b_1 \dots b_k$, then

this curve is also obtained by connecting the points $(0, 0), (1, b_1), \dots, (k, b_k)$. The last point is $(|w|, \Sigma(w))$. See Figure 1 for an example. The biggest y -coordinate is $\max(\text{psw}(w))$ and the smallest y -coordinate is $\min(\text{psw}(w))$, except that we need to ignore the point $(0, 0)$ and the points between $(0, 0)$ and $(1, b_1)$. In fact, we are really only interested in the points $(1, b_1), \dots, (k, b_k)$, and the line segments connecting them are just meant to make the drawings look better. The word $\text{psw}_r(w)$ could be represented in a similar way by starting at the point $(0, r)$ instead of $(0, 0)$. The curve of $\text{psw}(uv)$ consists of the curve of $\text{psw}(u)$ followed by the curve of $\text{psw}(v)$ translated in such a way that its starting point matches the end point of the curve of $\text{psw}(u)$.

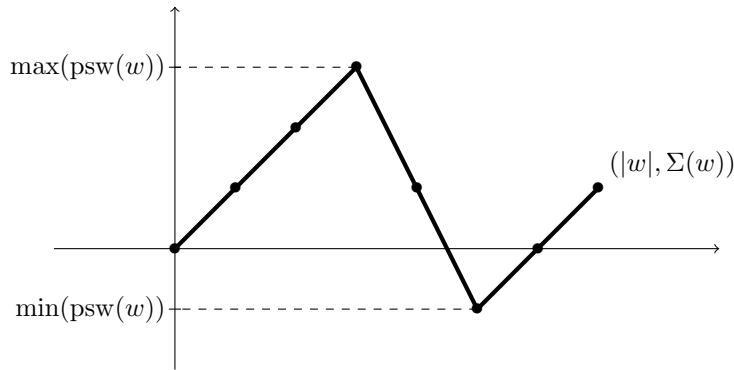


Figure 1: Graphical representation of the word $\text{psw}(w)$, where $w = aaabbaa$, $a = 1$, and $b = -2$. We have $|w| = 7$, $\Sigma(w) = 1$, $\max(\text{psw}(w)) = 3$, and $\min(\text{psw}(w)) = -1$.

The graphical interpretation is similar to the relation between Dyck words and Dyck paths, or the definition of Sturmian words as mechanical words. Representations of words as paths (or paths as words) can also be used in discrete geometry. This can, for example, lead to connections between word equations and tilings of a plane, see [4] for a survey.

3. Numbers of occurrences of letters

In this paper, we frequently need to study words of the forms $\text{psw}_{Ak+B}(x)$ and $\text{psw}_{Ak+B}(x^k)$, where x is a word, A, B are numbers, and k is a positive integer. In this section, we study how the number of occurrences of a letter changes when k changes. Let us first describe the graphical representations of these words to develop some geometric intuition, which could be used to give alternative (somewhat informal) proofs of Lemmas 3.1 and 4.3.

The starting point of the curve of $\text{psw}_{Ak+B}(x)$ is at height $Ak + B$ and the end point at height $Ak + B + \Sigma(x)$. When k grows, the curve remains the same, except that it is translated vertically if $A \neq 0$: It moves up if $A > 0$ and down if $A < 0$.

The starting point of the curve of $\text{psw}_{Ak+B}(x^k)$ is at height $Ak + B$ and the end point at height $(A + \Sigma(x))k + B$. The curve consists of k translated copies

of the curve of $\text{psw}(x)$. If $\Sigma(x) = 0$, then all k copies are at the same level, if $\Sigma(x) > 0$, then the last copy is the highest one, and if $\Sigma(x) < 0$, then the first copy is the highest one. When k grows, the highest point on the curve moves up if $A > 0$ or $A + \Sigma(x) > 0$, it moves down if $A < 0$ and $A + \Sigma(x) < 0$, and it stays the same if $\max\{A, A + \Sigma(x)\} = 0$.

Let $K \subseteq \mathbb{Z}$ and let $f : K \rightarrow \mathbb{Z}$ be a function. We use the following definitions, most of which are standard:

1. f is *constant* if $f(x) = f(y)$ for all $x, y \in K$.
2. f is *positive* if $f(x) > 0$ for all $x \in K$.
3. f is *increasing* if $f(x) \leq f(y)$ whenever $x < y$.
4. f is *strictly increasing* if $f(x) < f(y)$ whenever $x < y$.
5. f is *strictly decreasing* if $f(x) > f(y)$ whenever $x < y$.
6. f is *affine* if there exist numbers a, b such that $f(x) = ax + b$ for all $x \in K$.
7. f is *convex* if $(f(y) - f(x))/(y - x) \leq (f(z) - f(x))/(z - x)$ whenever $x < y < z$.

The following facts are direct consequences of the definitions:

1. Every constant function is affine and increasing.
2. Every affine function is convex.
3. The sum of increasing (affine, convex) functions is increasing (affine, convex, respectively).
4. The sum of increasing functions is strictly increasing if at least one of the functions is strictly increasing.
5. The sum of convex functions is not affine if at least one of the functions is not affine.

The reason why we stated the definitions of affine and convex functions (and why these concepts appear in the next lemma) is that later, in the proof of Lemma 5.4, we show that certain two functions cannot be the same, because one of them is affine and the other one is a sum of convex functions at least one of which is not affine. We do not use convexity in any other way.

Lemma 3.1. *Let $x \in \Gamma^+$, $A, B \in \mathbb{R}$, $\{\alpha, \beta\} = \{0, 1\}$, $K \subseteq \mathbb{Z}_+$, $|K| \geq 3$, $a \in \mathbb{R}$. Let $A' = A + \alpha\Sigma(x)$ and $M = \max\{\max(\text{psw}_{Ak+B}(x^{\alpha k + \beta})) \mid k \in K\}$. Consider the function*

$$\phi : K \rightarrow \mathbb{Z}, \quad \phi(k) = \lfloor \text{psw}_{Ak+B}(x^{\alpha k + \beta}) \rfloor_a.$$

- (a) *If $A = A' = 0$ and ϕ is not the zero function, then either $\alpha = 0$ and ϕ is constant, or $\Sigma(x) = 0 \neq \alpha$ and ϕ is positive, strictly increasing, and affine.*
- (b) *If $A = 0$, $A' < 0$, and $a = M$, then ϕ is constant.*
- (c) *If $A < 0$, $A' = 0$, and $a = M$, then ϕ is constant.*
- (d) *If $A < 0$, $A' < 0$, and $a = M$, then ϕ is convex but not affine.*
- (e) *If $\max\{A, A'\} > 0$ and $a = M$, then ϕ is convex but not affine.*

Proof. We have

$$\text{psw}_{Ak+B}(x^{\alpha k+\beta}) = \prod_{i=1}^{\alpha k+\beta} \text{psw}_{Ak+B+(i-1)\Sigma(x)}(x)$$

and thus

$$\phi(k) = \sum_{i=1}^{\alpha k+\beta} |\text{psw}_{Ak+B+(i-1)\Sigma(x)}(x)|_a \quad (3)$$

and

$$\begin{aligned} & \max(\text{psw}_{Ak+B}(x^{\alpha k+\beta})) \\ &= \max(\text{psw}(x)) + \max\{Ak + B + (i-1)\Sigma(x) \mid i \in \{1, \dots, \alpha k + \beta\}\} \\ &= \max(\text{psw}(x)) + \begin{cases} Ak + B & \text{if } \Sigma(x) \leq 0 \\ A'k + B + (\beta - 1)\Sigma(x) & \text{if } \Sigma(x) \geq 0 \end{cases} \end{aligned}$$

for all k .

- (a) Let $A = A' = 0$. Then one of α and $\Sigma(x)$ is zero. If $\alpha = 0$, then ϕ is clearly constant. If $\Sigma(x) = 0 \neq \alpha$, then $\phi(k) = (\alpha k + \beta) |\text{psw}_B(x)|_a$ is either the zero function (if $|\text{psw}_B(x)|_a = 0$) or positive, strictly increasing, and affine.
- (b) Let $A = 0$ and $A' < 0$. Then $\alpha = 1$, $\Sigma(x) < 0$. For a fixed k , $Ak + B + (i-1)\Sigma(x)$ is strictly decreasing with respect to i , so if $a = M$, then only the first term in the sum (3) can be positive, and therefore $\phi(k) = |\text{psw}_{Ak+B}(x)|_a$. This is constant.
- (c) Let $A < 0$ and $A' = 0$. Then $\alpha = 1$, $\Sigma(x) > 0$. For a fixed k , $Ak + B + (i-1)\Sigma(x)$ is strictly increasing with respect to i , so if $a = M$, then only the last term in the sum (3) can be positive, and therefore $\phi(k) = |\text{psw}_{A'k+B+(\beta-1)\Sigma(x)}(x)|_a$. This is constant.
- (d) Let $A < 0$ and $A' < 0$. Then $\max(\text{psw}_{Ak+B}(x^{\alpha k+\beta}))$ is strictly decreasing with respect to k , so if $a = M$, then $\phi(k) > 0$ for $k = \min(K)$ and $\phi(k) = 0$ otherwise, which makes ϕ convex and not affine.
- (e) Let $\max\{A, A'\} > 0$. Then $\max(\text{psw}_{Ak+B}(x^{\alpha k+\beta}))$ is strictly increasing with respect to k , so if $a = M$, then $\phi(k) > 0$ for $k = \max(K)$ and $\phi(k) = 0$ otherwise, which makes ϕ convex and not affine. \square

4. Distances between occurrences of letters

For a word w and a letter a , we need to consider distances between consecutive occurrences of a . We need to keep track of not just one of these distances (e.g., the largest one), but all of them, and also how many times each distance occurs. In other words, we need to consider the multiset of these distances. A multiset $\{a_1, \dots, a_n\}$ of nonnegative integers can be conveniently encoded as a polynomial $X^{a_1} + \dots + X^{a_n} \in \mathbb{Z}[X]$. This idea leads to the following definitions.

If $|w|_a = n > 0$, then there are unique words u_0, \dots, u_n such that $|u_i|_a = 0$ for all $i \in \{0, \dots, n\}$ and $w = u_0 a u_1 \dots a u_n$, and we define a polynomial in $\mathbb{Z}[X]$ that encodes the distances between consecutive occurrences of a :

$$D_a(w) = \sum_{i=1}^{n-1} X^{|u_i|}.$$

We also need to consider the lengths of the prefix u_0 and the suffix u_n , so we define two other polynomials:

$$P_a(w) = X^{|u_0|} \quad \text{and} \quad S_a(w) = X^{|u_n|}.$$

We use the notation $(P_a + D_a + S_a)(w) = P_a(w) + D_a(w) + S_a(w)$. If $|w|_a = 0$, then these polynomials are not defined.

Example 4.1. If $w = abbaabbab$, then $D_a(w) = 1 + 2X^2$, $P_a(w) = 1$, $S_a(w) = X$.

In the next lemma, we see how the polynomials behave with respect to concatenation.

Lemma 4.2. *Let a be a letter and $u_0, \dots, u_n, v_1, \dots, v_n$ words such that $|u_i|_a = 0$ for all $i \in \{0, \dots, n\}$ and $|v_i|_a > 0$ for all $i \in \{1, \dots, n\}$. Then*

$$\begin{aligned} D_a(u_0 v_1 u_1 \dots v_n u_n) &= \sum_{i=1}^n D_a(v_i) + \sum_{i=1}^{n-1} S_a(v_i) X^{|u_i|} P_a(v_{i+1}), \\ P_a(u_0 v_1 u_1 \dots v_n u_n) &= X^{|u_0|} P_a(v_1), \\ S_a(u_0 v_1 u_1 \dots v_n u_n) &= S_a(v_n) X^{|u_n|}. \end{aligned}$$

Proof. Follows directly from the definitions. □

The next lemma is a counterpart of Lemma 3.1.

Lemma 4.3. *Let $x \in \Gamma^+$, $A, B \in \mathbb{R}$, $\{\alpha, \beta\} = \{0, 1\}$, $a \in \mathbb{R}$. Let $z(k) = \text{psw}_{Ak+B}(x^{\alpha k + \beta})$ for all $k \in \mathbb{Z}_+$. Let $A' = A + \alpha \Sigma(x)$.*

(a) *If $A = A' = 0$ and $|z(k)|_a > 0$ for some $k \in \mathbb{Z}_+$, then $|z(k)|_a > 0$ for all $k \in \mathbb{Z}_+$ and there exist $f(X), g(X) \in \mathbb{Z}[X]$, $p, q \in \mathbb{Z}_{\geq 0}$ such that for all $k \in \mathbb{Z}_+$,*

$$D_a(z(k)) = kf(X) + g(X), \quad P_a(z(k)) = X^p, \quad S_a(z(k)) = X^q.$$

Moreover, if $\alpha = 0$, then $f(X) = 0$.

(b) *If $A = 0$, $A' < 0$, and $a = \max(z(k))$ for some k , then $|z(k)|_a > 0$ for all $k \in \mathbb{Z}_+$ and there exist $g(X) \in \mathbb{Z}[X]$, $p, q \in \mathbb{Z}_{\geq 0}$, $r \in \mathbb{Z}_+$ such that for all $k \in \mathbb{Z}_+$,*

$$D_a(z(k)) = g(X), \quad P_a(z(k)) = X^p, \quad S_a(z(k)) = X^{q+rk}.$$

- (c) If $A < 0$, $A' = 0$, and $a = \max(z(k))$ for some k , then $|z(k)|_a > 0$ for all $k \in \mathbb{Z}_+$ and there exist $g(X) \in \mathbb{Z}[X]$, $p, q \in \mathbb{Z}_{\geq 0}$, $r \in \mathbb{Z}_+$ such that for all $k \in \mathbb{Z}_+$,

$$D_a(z(k)) = g(X), \quad P_a(z(k)) = X^{p+rk}, \quad S_a(z(k)) = X^q.$$

Proof. We have

$$z(k) = \prod_{i=1}^{\alpha k + \beta} \text{psw}_{Ak+B+(i-1)\Sigma(x)}(x). \quad (4)$$

- (a) If $A = A' = 0$, then one of α and $\Sigma(x)$ is zero, and in both cases $z(k) = \text{psw}_B(x)^{\alpha k + \beta}$, so if $|z(k)|_a > 0$ for some $k \in \mathbb{Z}_+$, then $|z(k)|_a > 0$ for all $k \in \mathbb{Z}_+$ and

$$\begin{aligned} D_a(z(k)) &= (\alpha k + \beta)D_a(\text{psw}_B(x)) + (\alpha k + \beta - 1)P_a(\text{psw}_B(x))S_a(\text{psw}_B(x)), \\ P_a(z(k)) &= P_a(\text{psw}_B(x)), \\ S_a(z(k)) &= S_a(\text{psw}_B(x)). \end{aligned}$$

- (b) If $A = 0$, $A' < 0$, and $a = \max(z(k))$ for some k , then, like in the proof of Lemma 3.1 (b), we see that a can only occur in the first word of the product (4), which is $\text{psw}_B(x)$. Thus $|z(k)|_a = |\text{psw}_B(x)|_a > 0$ for all $k \in \mathbb{Z}_+$ and

$$\begin{aligned} D_a(z(k)) &= D_a(\text{psw}_B(x)), \\ P_a(z(k)) &= P_a(\text{psw}_B(x)), \\ S_a(z(k)) &= S_a(\text{psw}_B(x))X^{(\alpha k + \beta - 1)|x|}. \end{aligned}$$

- (c) If $A < 0$, $A' = 0$, and $a = \max(z(k))$ for some k , then, like in the proof of Lemma 3.1 (c), we see that a can only occur in the last word of the product (4), which is $\text{psw}_{B+(\beta-1)\Sigma(x)}(x)$. Thus $|z(k)|_a = |\text{psw}_{B+(\beta-1)\Sigma(x)}(x)|_a > 0$ for all $k \in \mathbb{Z}_+$ and

$$\begin{aligned} D_a(z(k)) &= D_a(\text{psw}_{B+(\beta-1)\Sigma(x)}(x)), \\ P_a(z(k)) &= P_a(\text{psw}_{B+(\beta-1)\Sigma(x)}(x))X^{(\alpha k + \beta - 1)|x|}, \\ S_a(z(k)) &= S_a(\text{psw}_{B+(\beta-1)\Sigma(x)}(x)). \quad \square \end{aligned}$$

5. Main result

In this section, let $\alpha_1, \dots, \alpha_n$ and β_1, \dots, β_n be fixed numbers such that $\{\alpha_i, \beta_i\} = \{0, 1\}$ for all $i \in \{1, \dots, n\}$, and let $\alpha_n = 1$, $\beta_n = 0$. We are studying equalities of the form

$$x^k u = \prod_{i=1}^n x_i^{\alpha_i k + \beta_i}. \quad (5)$$

Let us explain why we assume $\alpha_n = 1$. If it were $\alpha_n = 0$, then one of u and x_n would be a suffix of the other (or the equality would not hold for any k),

and we could cancel out this suffix from both sides. After the cancellation, one of u and x_n would be empty. If x_n were empty, we could just leave $x_n^{\alpha_n k + \beta_n}$ out from the product. If u were empty and x_n nonempty, then the length of the left-hand side and the length of the right-hand side of (5) would be equal for at most one value of k (a similar length argument is used in the proof of Lemma 5.1).

The equalities (2) with $m = 1$ and u_0 empty can be written in the form (5) (the number n and the words x_1, \dots, x_n here are not the same as the n and x_1, \dots, x_n in (2)), and conversely, (5) can be written in the form (2). The main difference between (2) with $m = 1$ and u_0 empty and (5) is that in (2), there is always a word v_i between the k th powers y_i^k and y_{i+1}^k , but the words are allowed to be empty, while as in (5), the words x_1, \dots, x_n have to be nonempty, but there can be two consecutive k th powers x_i^k and x_{i+1}^k . Note that allowing α_i, β_i to be arbitrary nonnegative integers would not make (5) any more general because we can always write $x_i^{\alpha_i k + \beta_i} = y^k z$, where $y = x_i^{\alpha_i}$ and $z = x_i^{\beta_i}$.

For words x, u, x_1, \dots, x_n , we define a set

$$K(x, u, x_1, \dots, x_n) = \left\{ k \in \mathbb{Z}_+ \mid x^k u = \prod_{i=1}^n x_i^{\alpha_i k + \beta_i} \right\}.$$

Our goal is to show that this set is always either the whole \mathbb{Z}_+ or of size at most two. By Lemma 2.2, we can assume that x is zero-sum. There are two cases depending on whether there exists an index i such that $\alpha_i \Sigma(x_i) \neq 0$.

If such an i does not exist, then, by using the mapping psw , we can replace the words x, u, x_1, \dots, x_n by new words so that the lengths of the words and the set $K(x, u, x_1, \dots, x_n)$ are preserved. With certain assumptions, this process increases the size of the alphabet. But if we chose the original words x, u, x_1, \dots, x_n in such a way that the alphabet was maximal, then of course the alphabet cannot grow, and this leads to $K(x, u, x_1, \dots, x_n) = \mathbb{Z}_+$. This is the idea of Lemma 5.3.

If there exists an index i such that $\alpha_i \Sigma(x_i) \neq 0$, then we can take the prefix sum words of both sides of (5) and look at the numbers of occurrences of a certain letter (with the help of Lemma 3.1), and the distances between consecutive occurrences (with the help of Lemma 4.3). If we assume $|K(x, u, x_1, \dots, x_n)| \geq 3$, then this leads to a contradiction. This is done in the proof of Lemma 5.4.

First, we are going to prove a couple of simple facts about the equalities (5).

Lemma 5.1. *Let $u \in \Gamma^*$, $x, x_1, \dots, x_n \in \Gamma^+$, $|K(x, u, x_1, \dots, x_n)| \geq 2$. Then*

$$|x| = \sum_{i=1}^n \alpha_i |x_i|, \quad |u| = \sum_{i=1}^n \beta_i |x_i|, \quad \Sigma(x) = \sum_{i=1}^n \alpha_i \Sigma(x_i), \quad \Sigma(u) = \sum_{i=1}^n \beta_i \Sigma(x_i).$$

Moreover, the left-hand side and the right-hand side of (5) have the same length and the same sum for all $k \in \mathbb{Z}_+$.

Proof. Taking the length of (5) gives $k|x| + |u| = \sum_{i=1}^n (\alpha_i k + \beta_i) |x_i|$ for all $k \in K(x, u, x_1, \dots, x_n)$. This can hold for two different values of k only if

$|x| = \sum_{i=1}^n \alpha_i |x_i|$ and $|u| = \sum_{i=1}^n \beta_i |x_i|$, and then it holds for all values of k . The claims about sums can be proved in a similar way. \square

Lemma 5.2. *Let $u \in \Gamma^*$, $x, x_1, \dots, x_n \in \Gamma^+$, $|K(x, u, x_1, \dots, x_n)| \geq 2$. Then $\text{alph}(x) = \text{alph}(xu)$.*

Proof. Let $a \in \text{alph}(xu)$ be arbitrary. Then $S_a(x^k u)$ is the same for all k , namely it is equal to $S_a(xu)$. If j is the largest index such that $|x_j|_a > 0$, then $S_a(\prod_{i=1}^n x_i^{\alpha_i k + \beta_i}) = S_a(x_j) X^{L(k)}$, where $L(k) = |\prod_{i=j+1}^n x_i^{\alpha_i k + \beta_i}|$, and this must also be the same for all $k \in K(x, u, x_1, \dots, x_n)$. If it were $j < n$, then $L(k) = |\prod_{i=j+1}^{n-1} x_i^{\alpha_i k + \beta_i}| + |x_n^k|$ would be strictly increasing with respect to k , which is a contradiction. Therefore $j = n$, that is, $|x_n|_a > 0$, and then $|x_n^{\alpha_n k + \beta_n}|_a = |x_n^k|_a = k|x_n|_a$ is strictly increasing with respect to k , and thus the number of occurrences of a on the right-hand side of (5) is strictly increasing. The number of occurrences of a on the left-hand side of (5) can be strictly increasing only if $|x|_a > 0$, so $a \in \text{alph}(x)$. Because $a \in \text{alph}(xu)$ was arbitrary, we have shown that $\text{alph}(x) = \text{alph}(xu)$. \square

Lemma 5.3. *Let $u \in \Gamma^*$, $x, x_1, \dots, x_n \in \Gamma^+$, $K = K(x, u, x_1, \dots, x_n)$, $|K| \geq 2$. We assume that the zero-sum words in Γ^* are exactly the words whose Parikh vectors are scalar multiples of Π_x , and that $\alpha_i \Sigma(x_i) = 0$ for all $i \in \{1, \dots, n\}$. If $K \neq \mathbb{Z}_+$, then there exist words y, v, y_1, \dots, y_n such that*

$$\begin{aligned} (|y|, |v|, |y_1|, \dots, |y_n|) &= (|x|, |u|, |x_1|, \dots, |x_n|), \\ K(y, v, y_1, \dots, y_n) &= K, \\ |\text{alph}(yv)| &> |\text{alph}(xu)|. \end{aligned}$$

Proof. Let $y = \text{psw}(x)$, $v = \text{psw}(u)$, and $y_i = \text{psw}_{B_i}(x_i)$ for all i , where $B_i = \sum_{j=1}^{i-1} \beta_j \Sigma(x_j)$. Then $(|y|, |v|, |y_1|, \dots, |y_n|) = (|x|, |u|, |x_1|, \dots, |x_n|)$. For all k , we have $\text{psw}(x^k u) = y^k v$ because $\Sigma(x) = 0$, and

$$\text{psw}\left(\prod_{i=1}^n x_i^{\alpha_i k + \beta_i}\right) = \prod_{i=1}^n y_i^{\alpha_i k + \beta_i} \quad (6)$$

because $\alpha_i \Sigma(x_i) = 0$ for all i . Two words are equal if and only if their prefix sum words are equal, so $K(y, v, y_1, \dots, y_n) = K$. To complete the proof, we assume that $|\text{alph}(yv)| \leq |\text{alph}(xu)|$ and show that $K = \mathbb{Z}_+$.

By Lemma 5.2, $\text{alph}(x) = \text{alph}(xu)$. Let $m = |\text{alph}(x)|$. Let $xu = a_1 \cdots a_N$ and $yv = b_1 \cdots b_N$, where $a_1, \dots, a_N, b_1, \dots, b_N$ are letters. Because $yv = \text{psw}(xu)$, $b_i = a_1 + \cdots + a_i$ for all i . If $b_i = b_j$ for some $i < j$, then $a_{i+1} \cdots a_j$ is zero-sum, and by the assumption about Parikh vectors of zero-sum words, every letter of $\text{alph}(x)$ occurs in $a_{i+1} \cdots a_j$, so $j - i \geq m$. Thus any m consecutive letters in yv are pairwise distinct. This means that any m consecutive letters in yv contain each letter of $\text{alph}(yv)$ exactly once, because $|\text{alph}(yv)| \leq |\text{alph}(xu)| = m$. It follows that $b_{i+m} = b_i$ for all i and yv is a prefix of a word in $(b_1 \cdots b_m)^*$. Because x is zero-sum, the last letter of y must be 0, and because

$a_1 \cdots a_i$ is not zero-sum for any $i < m$, it must be $b_m = 0$ and $y \in (b_1 \cdots b_m)^*$. It follows that $y^k v$ is a prefix of a word in $(b_1 \cdots b_m)^*$ for all k , and therefore so is (6) for $k \in K$. Consequently, if the last letter of y_i is b_j , then the first letter of y_{i+1} must be b_{j+1} , and if also $\alpha_i = 1$, then the first letter of y_i must be b_{j+1} . It then follows that (6) is a prefix of a word in $(b_1 \cdots b_m)^*$ for all k . We have shown that for all k , $y^k v$ and (6) are prefixes of a common word, and they have the same length by Lemma 5.1, so they are equal. It follows that $K = \mathbb{Z}_+$. \square

Lemma 5.4. *Let $u \in \Gamma^*$, $x, x_1, \dots, x_n \in \Gamma^+$, $K = K(x, u, x_1, \dots, x_n)$. We assume that $\Sigma(x_n) \leq 0 = \Sigma(x)$, and that $\alpha_m \Sigma(x_m) \neq 0$ for some $m \in \{1, \dots, n\}$. Then $|K| \leq 2$.*

Proof. We assume that $|K| \geq 3$ and derive a contradiction. First, we set up some notation. Let $A_i = \sum_{j=1}^{i-1} \alpha_j \Sigma(x_j)$ and $B_i = \sum_{j=1}^{i-1} \beta_j \Sigma(x_j)$ for all i and $z_i(k) = \text{psw}_{A_i k + B_i}(x_i^{\alpha_i k + \beta_i})$ for all i, k . By taking the prefix sum word of (5), we get

$$\text{psw}(x)^k \text{psw}(u) = \prod_{i=1}^n z_i(k) \quad (7)$$

for all $k \in K$. Let

$$\begin{aligned} I_1 &= \{i \in \{1, \dots, n\} \mid A_i = A_{i+1} = 0\}, \\ I_2 &= \{i \in \{1, \dots, n\} \mid \min\{A_i, A_{i+1}\} < 0 = \max\{A_i, A_{i+1}\}\}, \\ I_3 &= \{i \in \{1, \dots, n\} \mid \max\{A_i, A_{i+1}\} \neq 0\}. \end{aligned}$$

Then $\{1, \dots, n\}$ is a disjoint union of I_1, I_2, I_3 . We can recall the geometric intuition described at the beginning of Section 3 and notice that as k grows, both endpoints of the curve of $z_i(k)$ stay on the same level if $i \in I_1$, one of the endpoints moves down and the other stays on the same level if $i \in I_2$, and either at least one of the endpoints moves up or both move down if $i \in I_3$.

For any letter a , we can count the number of occurrences of a in (7) and we get

$$k |\text{psw}(x)|_a + |\text{psw}(u)|_a = \sum_{i \in I_1} |z_i(k)|_a + \sum_{i \in I_2} |z_i(k)|_a + \sum_{i \in I_3} |z_i(k)|_a \quad (8)$$

for all $k \in K$. Let a be the largest letter for which there exists $k \in K$ such that

$$\sum_{i \in I_2} |z_i(k)|_a + \sum_{i \in I_3} |z_i(k)|_a > 0.$$

Such a letter a exists because $I_2 \cup I_3 \neq \emptyset$ by the assumption that $\alpha_m \Sigma(x_m) \neq 0$ for some m . If we let

$$J_j = I_j \cap \{i \in \{1, \dots, n\} \mid \exists k \in K : |z_i(k)|_a > 0\}$$

for all $j \in \{1, 2, 3\}$, then (8) can be written as

$$k |\text{psw}(x)|_a + |\text{psw}(u)|_a = \sum_{i \in J_1} |z_i(k)|_a + \sum_{i \in J_2} |z_i(k)|_a + \sum_{i \in J_3} |z_i(k)|_a \quad (9)$$

for all $k \in K$. For $i \in J_2 \cup J_3$, $a = \max\{\max(z_i(k)) \mid k \in K\}$.

Next, we show that $J_3 = \emptyset$. As a function of k , the first sum on the right-hand side of (9) is affine by Lemma 3.1 (a) and the second sum is affine by Lemma 3.1 (b), (c). If it were $J_3 \neq \emptyset$, then the third sum would be convex but not affine by Lemma 3.1 (d), (e), and thus the whole right-hand side would be not affine, which would be a contradiction, because the left-hand side is clearly affine. Therefore it must be $J_3 = \emptyset$. Now we also know that whether a occurs in $z_i(k)$ does not depend on k . More specifically, by Lemma 3.1 (a), (b), (c), if $i \in J_1 \cup J_2$, then $|z_i(k)|_a > 0$ for all $k \in K$, and if $i \notin J_1 \cup J_2$, then $|z_i(k)|_a = 0$ for all $k \in K$.

Next, we show that $|\text{psw}(x)|_a > 0$. We have

$$S_a(\text{psw}(x)^k \text{psw}(u)) = S_a(\text{psw}(x) \text{psw}(u)),$$

which does not depend on k . If $j = \max(J_1 \cup J_2)$, then

$$S_a\left(\prod_{i=1}^n z_i(k)\right) = S_a(z_j(k)) X^{|z_{j+1}(k) \cdots z_n(k)|},$$

and this should also be the same for all $k \in K$. As a function of k , the degree of $S_a(z_j(k))$ is increasing by Lemma 4.3. If it were $j < n$, then the exponent $|z_{j+1}(k) \cdots z_n(k)| = |z_{j+1}(k) \cdots z_{n-1}(k)| + |x_n^k|$ would be strictly increasing, which would be a contradiction. Therefore $j = n$, that is, $n \in J_1 \cup J_2$. From the assumption $\Sigma(x_n) \leq 0$ it follows that $A_n \geq A_{n+1}$, and $A_{n+1} = \Sigma(x) = 0$ by Lemma 5.1, so it must be $n \in J_1$. Every term on the right-hand side of (9) is increasing, and $|z_n(k)|_a$ is strictly increasing by Lemma 3.1 (a), so the whole right-hand side is strictly increasing. The left-hand side must also be strictly increasing, and thus $|\text{psw}(x)|_a > 0$.

Finally, we study how the lengths of the gaps between occurrences of a in (7) change when k grows. The idea is that on the left-hand side, these lengths essentially stay the same (the number of a 's and thus the number of gaps increases, but the lengths of the new gaps are mostly the same as the lengths of some of the old ones), but on the right-hand side, some of the gaps become longer. To formalize this idea and to derive a contradiction, we can use the mapping $(P_a + D_a + S_a)$. It follows from Lemma 4.3 (a) that there exist $f_1(X), g_1(X) \in \mathbb{Z}[X]$ such that for all $k \in K$,

$$(P_a + D_a + S_a)(\text{psw}(x^k u)) = k f_1(X) + g_1(X).$$

Let $J_1 \cup J_2 = \{i_1, \dots, i_N\}$, where $i_1 < \dots < i_N$. Let $i_0 = 0$, $i_{N+1} = n + 1$, and

$L_j(k) = |z_{i_j+1}(k) \cdots z_{i_{j+1}-1}(k)|$ for $j \in \{0, \dots, N\}$. By Lemma 4.2,

$$\begin{aligned} D_a \left(\prod_{i=1}^n z_i(k) \right) &= \sum_{i \in J_1 \cup J_2} D_a(z_i(k)) + \sum_{j=1}^{N-1} S_a(z_{i_j}(k)) X^{L_j(k)} P_a(z_{i_{j+1}}(k)), \\ P_a \left(\prod_{i=1}^n z_i(k) \right) &= X^{L_0(k)} P_a(z_{i_1}(k)), \\ S_a \left(\prod_{i=1}^n z_i(k) \right) &= S_a(z_{i_N}(k)) X^{L_N(k)}. \end{aligned}$$

It follows from Lemma 4.3 that there exist $f_2(X), g_2(X) \in \mathbb{Z}[X]$ such that for all k ,

$$\sum_{i \in J_1 \cup J_2} D_a(z_i(k)) = kf_2(X) + g_2(X),$$

and that all the terms

$$S_a(z_{i_j}(k)) X^{L_j(k)} P_a(z_{i_{j+1}}(k)), X^{L_0(k)} P_a(z_{i_1}(k)), S_a(z_{i_N}(k)) X^{L_N(k)} \quad (10)$$

are of the form $X^{\gamma k + \delta}$. Thus there exist $M \geq 0, \gamma_1, \dots, \gamma_M, \delta_1, \dots, \delta_M \in \mathbb{Z}_{\geq 0}$ such that for all $k \in K$,

$$(P_a + D_a + S_a) \left(\prod_{i=1}^n z_i(k) \right) = kf_2(X) + g_2(X) + \sum_{i=1}^M X^{\gamma_i k + \delta_i}.$$

If $I = \{i \in \{1, \dots, M\} \mid \gamma_i = 0\}$ and $J = \{1, \dots, M\} \setminus I$, then we can write

$$k(f_1(X) - f_2(X)) + g_1(X) - g_2(X) - \sum_{i \in I} X^{\delta_i} = \sum_{i \in J} X^{\gamma_i k + \delta_i}.$$

The degree of the left-hand side is the same for all except possibly one k . The degree of the right-hand side, on the other hand, is strictly increasing with respect to k if $J \neq \emptyset$, so if we can show that $J \neq \emptyset$, then this is a contradiction and the proof is complete. We assumed that there exists m such that $\alpha_m \Sigma(x_m) \neq 0$. If this m is in $J_1 \cup J_2$, then it is in J_2 and the degree of either $P_a(z_{i_m}(k))$ or $S_a(z_{i_m}(k))$ is strictly increasing with respect to k by Lemma 4.3 (b), (c). Then one of the terms (10) is of the form $X^{\gamma k + \delta}$ with $\gamma > 0$ and therefore $J \neq \emptyset$. On the other hand, if $m \notin J_1 \cup J_2$, then $i_j < m < i_{j+1}$ for some j , and $|z_m(k)|$ and thus also $L_j(k)$ is strictly increasing. Then one of the terms (10) is of the form $X^{\gamma k + \delta}$ with $\gamma > 0$ and therefore $J \neq \emptyset$. This completes the proof. \square

Now we can state and prove our main result.

Theorem 5.5. *Let $u \in \Gamma^*$, $x, x_1, \dots, x_n \in \Gamma^+$, $K = K(x, u, x_1, \dots, x_n)$. Then either $|K| \leq 2$ or $K = \mathbb{Z}_+$.*

Proof. We can assume that $\text{alph}(xu)$ is maximal in the sense that if

$$(|y|, |v|, |y_1|, \dots, |y_n|) = (|x|, |u|, |x_1|, \dots, |x_n|) \quad \text{and} \quad K(y, v, y_1, \dots, y_n) = K$$

for some words y, v, y_1, \dots, y_n , then $|\text{alph}(yv)| \leq |\text{alph}(xu)|$. By Lemma 2.2, we can assume that x is zero-sum and that the zero-sum words in Γ^* are exactly the words whose Parikh vectors are scalar multiples of Π_x . The use of Lemma 2.2 just renames the letters, so the assumption about maximality of $\text{alph}(xu)$ still holds. We can assume that $\Sigma(x_n) \leq 0$ by replacing every letter by its negation if necessary. This operation preserves zero-sum words, so all of the assumptions still hold. If there exists m such that $\alpha_m \Sigma(x_m) \neq 0$, then $|K| \leq 2$ by Lemma 5.4, and if there does not exist m such that $\alpha_m \Sigma(x_m) \neq 0$, then $K = \mathbb{Z}_+$ by Lemma 5.3. \square

The result can also be formulated as follows.

Corollary 5.6. *Let $u, u_1, \dots, u_n \in \Gamma^*$ and $x, x_1, \dots, x_n \in \Gamma^+$. Then*

$$x^k u = \prod_{i=1}^n u_i x_i^k \tag{11}$$

holds either for all or for at most two values of $k \in \mathbb{Z}_+$. Similarly,

$$u x^k = \prod_{i=1}^n x_i^k u_i \tag{12}$$

holds either for all or for at most two values of $k \in \mathbb{Z}_+$.

Proof. Equality (11) is of the form (5), so the first claim follows from Theorem 5.5. If we denote the reversal of a word w by w^R , then (12) is equivalent to

$$(x^R)^k u^R = \prod_{i=0}^{n-1} u_{n-i}^R (x_{n-i}^R)^k,$$

which is of the same form as (11), so also the second claim follows. \square

Answer to Question 1.1 follows as a corollary.

Corollary 5.7. *Let $x, x_1, \dots, x_n \in \Gamma^*$. If $x^k = x_1^k \dots x_n^k$ for three positive integers k , then the words x, x_1, \dots, x_n commute.*

Proof. It follows from Theorem 5.5 that $x^k = x_1^k \dots x_n^k$ for all k . From $x^n = x_1^n \dots x_n^n$ it follows that the words x, x_1, \dots, x_n commute by the result in [1]. \square

6. Conclusion

As one possible direction for further research, we can ask the following question: For how many values of k does (2) need to hold to guarantee that it holds for all $k \geq 0$? This was first studied by Kortelainen [13]. Currently it is known that $m + n$ different values of k are sufficient [19]. Holub and Kortelainen [11] proved that if $m = 1$ and there exists $i \geq 2$ such that (2) holds for all $k \in \{i, i + 1, i + 2\}$, then (2) holds for all $k \geq 0$. In this article, we have come quite close to showing that if $m = 1$ and (2) holds for three values of k , then (2) holds for all $k \geq 0$. If we tried to replace the left-hand side of (5) by $u_0 x^k u_1$, there would be some problems, at least in the part of the proof of Lemma 5.4 where it is proved that $|\text{psw}(x)|_a > 0$. If a different way to handle this part of the proof was found, then completely solving the case $m = 1$ might be possible. The case $m > 1$, on the other hand, seems significantly more difficult.

Another direction for future research would be to try to apply the methods used in this paper to some other entirely different problems on word equations, as was done in [16]. We hope and believe that, in addition to the immediate impact of solving an open problem, this article will also lead to further advances in the future.

References

- [1] Appel, K. I., & Djourup, F. M. (1968). On the equation $z_1^n z_2^n \cdots z_k^n = y^n$ in a free semigroup. *Trans. Amer. Math. Soc.*, *134*, 461–470.
- [2] Barbin-Le Rest, E., & Le Rest, M. (1985). Sur la combinatoire des codes à deux mots. *Theoret. Comput. Sci.*, *41*, 61–80. doi:10.1016/0304-3975(85)90060-X.
- [3] Baschenis, F., Gauwin, O., Muscholl, A., & Puppis, G. (2017). Untwisting two-way transducers in elementary time. In *Proceedings of the 32nd LICS* (pp. 1–12). IEEE. doi:10.1109/LICS.2017.8005138.
- [4] Brlek, S. (2011). Interactions between digital geometry and combinatorics on words. In *Proceedings of the 8th WORDS* (pp. 1–12). volume 63 of *EPTCS*. doi:10.4204/EPTCS.63.1.
- [5] Gallot, P., Muscholl, A., Puppis, G., & Salvati, S. (2017). On the decomposition of finite-valued streaming string transducers. In *Proceedings of the 34th STACS* (pp. 34:1–34:14). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik volume 66 of *LIPICs*. doi:10.4230/LIPICs.STACS.2017.34.
- [6] Hakala, I., & Kortelainen, J. (1997). On the system of word equations $x_1^i x_2^i \cdots x_m^i = y_1^i y_2^i \cdots y_n^i$ ($i = 1, 2, \dots$) in a free monoid. *Acta Inform.*, *34*, 217–230. doi:10.1007/s002360050081.
- [7] Harju, T., & Karhumäki, J. (2004). Many aspects of defect theorems. *Theoret. Comput. Sci.*, *324*, 35–54. doi:10.1016/j.tcs.2004.03.051.

- [8] Harju, T., & Nowotka, D. (2005). On the equation $x^k = z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$ in a free semigroup. *Theoret. Comput. Sci.*, 330, 117–121. doi:10.1016/j.tcs.2004.09.012.
- [9] Holub, Š. (1999). A solution of the equation $(x_1^2 \cdots x_n^2)^3 = (x_1^3 \cdots x_n^3)^2$. In *Contributions to general algebra, 11 (Olomouc/Velké Karlovice, 1998)* (pp. 105–111). Heyn.
- [10] Holub, Š. (2001). Local and global cyclicity in free semigroups. *Theoret. Comput. Sci.*, 262, 25–36. doi:10.1016/S0304-3975(00)00156-0.
- [11] Holub, Š., & Kortelainen, J. (2007). On systems of word equations with simple loop sets. *Theoret. Comput. Sci.*, 380, 363–372. doi:10.1016/j.tcs.2007.03.026.
- [12] Karhumäki, J., & Plandowski, W. (1994). On the defect effect of many identities in free semigroups. In G. Paun (Ed.), *Mathematical aspects of natural and formal languages* (pp. 225–232). World Scientific.
- [13] Kortelainen, J. (1998). On the system of word equations $x_0 u_1^i x_1 u_2^i x_2 \cdots u_m^i x_m = y_0 v_1^i y_1 v_2^i y_2 \cdots v_n^i y_n$ ($i = 0, 1, 2, \dots$) in a free monoid. *J. Autom. Lang. Comb.*, 3, 43–57.
- [14] Lentin, A. (1965). Sur l'équation $a^M = b^N c^P d^Q$ dans un monoïde libre. *C. R. Acad. Sci. Paris*, 260, 3242–3244.
- [15] Lyndon, R. C., & Schützenberger, M.-P. (1962). The equation $a^M = b^N c^P$ in a free group. *Michigan Math. J.*, 9, 289–298. doi:10.1307/mmj/1028998766.
- [16] Nowotka, D., & Saarela, A. (2018). An optimal bound on the solution sets of one-variable word equations and its consequences. In *Proceedings of the 45th ICALP* (pp. 136:1–136:13). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik volume 107 of *LIPICs*. doi:10.4230/LIPICs.ICALP.2018.136.
- [17] Peltomäki, J., & Whiteland, M. (2017). A square root map on Sturmian words. *Electron. J. Combin.*, 24, P1.54.
- [18] Plandowski, W. (2003). Test sets for large families of languages. In *Proceedings of the 7th DLT* (pp. 75–94). Springer volume 2710 of *LNCS*. doi:10.1007/3-540-45007-6_6.
- [19] Saarela, A. (2015). Systems of word equations, polynomials and linear algebra: A new approach. *European J. Combin.*, 47, 1–14. doi:10.1016/j.ejc.2015.01.005.
- [20] Saarela, A. (2017). Word equations where a power equals a product of powers. In *Proceedings of the 34th STACS* (pp. 55:1–55:9). Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik volume 66 of *LIPICs*. doi:10.4230/LIPICs.STACS.2017.55.

- [21] Saarela, A. (2018). Studying word equations by a method of weighted frequencies. *Fund. Inform.*, 162, 223–235. doi:10.3233/FI-2018-1722.
- [22] Spehner, J.-C. (1976). *Quelques problèmes d'extension, de conjugaison et de présentation des sous-monoïdes d'un monoïde libre*. Ph.D. thesis Univ. Paris.