

PROPERTIES OF PSEUDO-PRIMITIVE WORDS AND THEIR APPLICATIONS

LILA KARI, BENOÎT MASSON, SHINNOSUKE SEKI

*Department of Computer Science, The University of Western Ontario,
London, Ontario, N6A 5B7, Canada
{lila, benoit, sseki}@csd.uwo.ca*

Received (Day Month Year)
Accepted (Day Month Year)
Communicated by (xxxxxxxxxx)

A pseudo-primitive word with respect to an antimorphic involution θ is a word which cannot be written as a catenation of occurrences of a strictly shorter word t and $\theta(t)$. Properties of pseudo-primitive words are investigated in this paper. These properties link pseudo-primitive words with essential notions in combinatorics on words such as primitive words, (pseudo)-palindromes, and (pseudo)-commutativity. Their applications include an improved solution to the extended Lyndon-Schützenberger equation $u_1 u_2 \cdots u_\ell = v_1 \cdots v_n w_1 \cdots w_m$, where $u_1, \dots, u_\ell \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$ for some words u, v, w , integers $\ell, n, m \geq 2$, and an antimorphic involution θ . We prove that for $\ell \geq 4$, $n, m \geq 3$, this equation implies that u, v, w can be expressed in terms of a common word t and its image $\theta(t)$. Moreover, several cases of this equation where $\ell = 3$ are examined.

Keywords: antimorphic involution, (pseudo-)primitive word, (extended) Lyndon-Schützenberger equation, (pseudo-)periodic word, (pseudo-)palindrome, (weak) defect effect

2000 Mathematics Subject Classification: 68Q70, 68R15

1. Introduction

For elements u, v, w in a free group, the equation of the form $u^\ell = v^n w^m$ ($\ell, n, m \geq 2$) is known as the *Lyndon-Schützenberger equation* (LS equation for short). Lyndon and Schützenberger [13] investigated the question of finding all possible solutions for this equation in a free group, and proved that if the equation holds, then u, v , and w are all powers of a common element. This equation can be also considered on the semigroup of all finite words over a fixed alphabet Σ , and an analogous result holds.

Theorem 1 (see, e.g., [7, 13, 14]) *For words $u, v, w \in \Sigma^+$ and integers $\ell, n, m \geq 2$, the equation $u^\ell = v^n w^m$ implies that u, v, w are powers of a common word.*

The Lyndon-Schützenberger equation has been generalized in several ways; e.g., the equation of the form $x^k = z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$ was investigated by Harju and Nowotka [8] and its special cases in [1, 11]. Czeizler et al. [3] have recently proposed another extension, which was originally motivated by the information encoded as DNA strands for DNA computing. In this framework, a DNA strand is modeled by a word w and encodes the same information as its Watson-Crick complement. In formal language theory, the Watson-Crick complementarity of DNA strands is modeled by an antimorphic involution θ [9, 15], i.e., a function θ on an alphabet Σ^* that is (a) antimorphic, $\theta(xy) = \theta(y)\theta(x)$, $\forall x, y \in \Sigma^*$, and (b) involution, $\theta^2 = id$, the identity. Thus, we can model the property whereby a DNA single strand binds to and is completely equivalent to its Watson-Crick complement, by considering a word u and its image $\theta(u)$ equivalent, for a given antimorphic involution θ .

For words u, v, w , integers $\ell, n, m \geq 2$, and an antimorphic involution θ , an extended Lyndon-Schützenberger equation (ExLS equation) is of the form

$$u_1 u_2 \cdots u_\ell = v_1 \cdots v_n w_1 \cdots w_m, \quad (1)$$

with $u_1, \dots, u_\ell \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$. The question arises as to whether an equation of this form implies the existence of a word t such that $u, v, w \in \{t, \theta(t)\}^+$. A given triple (ℓ, n, m) of integers is said to *impose pseudo-periodicity, with respect to θ , on u, v, w* , or simply, to *impose θ -periodicity on u, v, w* if (1) implies $u, v, w \in \{t, \theta(t)\}^+$ for some word t . Furthermore, we say that the triple (ℓ, n, m) *imposes θ -periodicity* if it imposes θ -periodicity on all u, v, w . The known results on ExLS equations [3] are summarized in Table 1.

Table 1. Summary of the known results regarding the extended Lyndon-Schützenberger equation.

ℓ	n	m	θ -periodicity
≥ 5	≥ 3	≥ 3	YES
3 or 4	≥ 3	≥ 3	?
2	≥ 2	≥ 2	?
≥ 3	2	≥ 2	NO

This paper is a step towards solving the unsettled cases of Table 1, by using the following strategy. Concise proofs exist in the literature for Theorem 1, that make use of fundamental properties such as:

- (i) The periodicity theorem of Fine and Wilf (FW theorem),
- (ii) The fact that a primitive word cannot be a proper infix of its square, and
- (iii) The fact that the class of primitive words is closed under cyclic permutation.

(For details of each, see [2].) In contrast, the proof given in [3] for the result about ExLS equations, stating that $(\geq 5, \geq 3, \geq 3)$ imposes θ -periodicity, involves tech-

niques designed for this specific purpose only. Should Properties (i), (ii), (iii) be generalized so as to take into account the informational equivalence between a word u and $\theta(u)$, they could possibly form a basis for a concise proof of the solutions to the ExLS equation. The approach we use in this paper is thus to seek analog properties for this extended case, and use the results we obtain to approach the unsettled cases in Table 1.

Czeizler, Kari, and Seki generalized Property (i) in [4]. There, first the notion of primitive words was extended to that of pseudo-primitive words with respect to a given antimorphic involution θ (or simply, θ -primitive words). A word u is said to be θ -primitive if there does not exist another word t such that $u \in t\{t, \theta(t)\}^+$. For example, if θ is the mirror image over $\{a, b\}^*$ (the identity function on $\{a, b\}$ extended to an antimorphism on $\{a, b\}^*$), $aabb$ is θ -primitive, while $abba$ is not because it can be written as $ab\theta(ab)$. Based on the θ -primitivity of words, Property (i) was generalized as follows: “For words u, v , if a word in $u\{u, \theta(u)\}^*$ and a word in $v\{v, \theta(v)\}^*$ share a long enough prefix (for details, see Theorems 5 and 6), then $u, v \in t\{t, \theta(t)\}^*$ for some θ -primitive word t .”

In contrast, little is known about Properties (ii) and (iii) except that they cannot be generalized as suggested in the previous example: non-trivial overlaps between two words in $\{t, \theta(t)\}^+$ are possible, and cyclic permutations do not in general preserve the θ -primitivity of words. As a preliminary step towards an extension of Property (ii), Czeizler et al. examined the non-trivial overlap of the form $v_1 \cdots v_m x = y v_{m+1} \cdots v_{2m}$, where $m \geq 1$, v_i is either v or $\theta(v)$ for some θ -primitive word v ($1 \leq i \leq 2m$), and both x and y are properly shorter than v [3]. Some of the results obtained there will be employed in this paper.

One purpose of this paper is to explore further the extendability of Properties (ii) and (iii). The main result here is Theorem 12, which states that for a θ -primitive word x , neither $x\theta(x)$ nor $\theta(x)x$ can be a proper infix of a word $x_1 x_2 x_3$, where $x_1, x_2, x_3 \in \{x, \theta(x)\}$. Based on this result, we consider two problems: For a θ -primitive word x , (1) does $v, yvz \in \{x, \theta(x)\}^+$ imply that y and z are in $\{x, \theta(x)\}^*$?; and (2) if the catenation of words u, v is in $\{x, \theta(x)\}^+$, under what conditions does $u, v \in \{x, \theta(x)\}^*$ hold? In particular, our investigation into the second problem will reveal close relationships between primitive words, θ -primitive words, and θ -palindromes (fixed points of θ). These relationships further present several cyclic permutations under which the θ -primitivity of words is preserved.

The results thus obtained enable us to prove that the triple $(4, \geq 3, \geq 3)$ imposes θ -periodicity (Theorem 48) in a much simpler manner than the proof in [3] for $(\geq 5, \geq 3, \geq 3)$. Even for $(3, n, m)$ ExLS equations, these results give some insight and narrow down the open cases of ExLS equations.

The paper is organized as follows: in the next section, we provide required notions and notation. Section 3 begins with the proof of some basic properties of θ -primitive words, and then proves some consequences of overlaps between θ -primitive words of a similar flavour with Properties (ii) and (iii) (e.g., Theorem 12, Corollary 24).

These tools are used in Section 4, where we prove that the $(4, \geq 3, \geq 3)$ ExLS equation has only θ -periodic solutions (Theorem 48), and study particular cases of $(3, n, m)$ ExLS equations.

2. Preliminaries

An alphabet is a finite and non-empty set of symbols. In the sequel, we shall use a fixed non-singleton alphabet Σ . The set of all words over Σ is denoted by Σ^* , which includes the empty word λ , and let $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$. The length of a word $w \in \Sigma^*$ is denoted by $|w|$. A word v is an *infix* (resp. *prefix*, *suffix*) of a word w if $w = xvy$ (resp. $w = vy$, $w = xv$) for some $x, y \in \Sigma^*$; in any case, if $w \neq v$, then the infix (prefix, suffix) is said to be *proper*. For a word w , denote by $\text{Pref}(w)$ the set of prefixes of w and by $\text{Suff}(w)$ the set of its suffixes.

A language L is a subset of Σ^* . For a non-negative integer $n \geq 0$, we write L^n for the language consisting of all words of the form $w_1 \cdots w_n$ such that each w_i is in L . We also write $L^{\geq n}$ for $L^n \cup L^{n+1} \cup L^{n+2} \cup \cdots$. Analogously, we can define $L^{\leq n} = L^0 \cup L^1 \cup \cdots \cup L^n$. For $L^{\geq 0}$ and $L^{\geq 1}$, we employ the traditional notation L^* and L^+ .

A mapping $\theta : \Sigma^* \rightarrow \Sigma^*$ is called an *antimorphic involution* of Σ^* if $\theta(xy) = \theta(y)\theta(x)$ for any $x, y \in \Sigma^*$ (antimorphism), and θ^2 is equal to the identity (involution). Throughout this paper, θ denotes an antimorphic involution. The *mirror image*, which maps a word to its reverse, is a typical example of antimorphic involution. A word $w \in \Sigma^*$ is called a θ -*palindrome* if $w = \theta(w)$. A word which is a θ -palindrome with respect to a given but unspecified antimorphic involution θ is also called a *pseudo-palindrome* [5].

A non-empty word $w \in \Sigma^+$ is said to be *primitive* if $w = v^n$ implies $n = 1$ for any word $v \in \Sigma^+$. It is known that any non-empty word $w \in \Sigma^+$ can be written as a power of a unique primitive word, which is called the *primitive root* of w , and denoted by $\rho(w)$. Two words which *commute* share a primitive root, that is, $uv = vu$ implies $\rho(u) = \rho(v)$ (see [2]). In literature, it is said that $uv = vu$ causes a *defect effect* (for details of defect effects and defect theorems, see [2, 14]). The LS equation also causes defect effect, since $u^\ell = v^n w^m$ with $\ell, n, m \geq 2$ implies $\rho(u) = \rho(v) = \rho(w)$ (Theorem 1). The following results describe other relations causing a defect effect.

Lemma 2 ([4]) *Let $u \in \Sigma^+$ such that $u = pq$ for some θ -palindromes $p, q \in \Sigma^+$. If $q \in \text{Pref}(u)$ and $|q| \geq |p|$, then $\rho(p) = \rho(q) = \rho(u)$.*

Theorem 3 ([2]) *Let $u, v \in \Sigma^+$. If there exist $\alpha(u, v) \in u\{u, v\}^*$ and $\beta(u, v) \in v\{u, v\}^*$ which share a prefix of length at least $|u| + |v|$, then $\rho(u) = \rho(v)$.*

The notion of primitive word was generalized into that of pseudo-primitive word by Czeizler, Kari, and Seki [4]. For an antimorphic involution θ , a non-empty word $w \in \Sigma^+$ is said to be *pseudo-primitive with respect to θ* , or simply θ -*primitive*, if $w \in \{v, \theta(v)\}^n$ implies $n = 1$ for any word $v \in \Sigma^+$. In [4] it was proved that for

any non-empty word $w \in \Sigma^+$, there exists a unique θ -primitive word t satisfying $w \in t\{t, \theta(t)\}^*$. Such a word t is called the θ -primitive root of w . The next lemma describes a property of the θ -primitive root of a θ -palindrome of even length.

Lemma 4. *Let $x \in \Sigma^+$ be a θ -primitive word and p be a θ -palindrome of even length. If $p = x_1x_2 \cdots x_m$ for some $m \geq 1$ and $x_1, \dots, x_m \in \{x, \theta(x)\}$, then m has to be even.*

Proof. Suppose that the equality held for some odd m . Then x must be of even length because $|p|$ is even. Hence $x_{(m-1)/2}$ becomes a θ -palindrome. Thus $x = y\theta(y)$ for some $y \in \Sigma^+$. However, this contradicts the θ -primitivity of x . \square

The *theorem of Fine and Wilf* (FW theorem) is one of the fundamental results on periodicity [6]. It states that for two words $u, v \in \Sigma^+$, if a power of u and a power of v share a prefix of length at least $|u| + |v| - \gcd(|u|, |v|)$, then $\rho(u) = \rho(v)$, where $\gcd(\cdot, \cdot)$ denotes the greatest common divisor of two arguments (for its proof, see, e.g., [2]). This theorem has been generalized in [4], by taking into account the equivalence between a word and its image under θ , in the following two forms.

Theorem 5 ([4]) *Let $u, v \in \Sigma^+$. If a word in $\{u, \theta(u)\}^*$ and a word in $\{v, \theta(v)\}^*$ share a prefix of length at least $\text{lcm}(|u|, |v|)$, then $u, v \in \{t, \theta(t)\}^+$ for some θ -primitive word $t \in \Sigma^+$, where $\text{lcm}(\cdot, \cdot)$ denotes the least common multiple of two arguments.*

Theorem 6 ([4]) *Let $u, v \in \Sigma^+$ with $|u| \geq |v|$. If a word in $\{u, \theta(u)\}^*$ and a word in $\{v, \theta(v)\}^*$ share a prefix of length at least $2|u| + |v| - \gcd(|u|, |v|)$, then $u, v \in \{t, \theta(t)\}^+$ for some θ -primitive word $t \in \Sigma^+$.*

In a way, we can say that these theorems describe relations causing a *weak defect effect* because they all imply that $u, v \in \{t, \theta(t)\}^+$ for some θ -primitive word $t \in \Sigma^+$, which is strictly weaker than the usual defect effect $\rho(u) = \rho(v)$ [4]. Various relations causing such a weak defect effect were presented in [4].

Besides, the commutativity $xy = yx$ was extended to the θ -commutativity $xy = \theta(y)x$ in [10]. This is a special case of $xy = zx$, whose solutions are given as $x = r(tr)^i$, $y = (tr)^j$, and $z = (rt)^j$ for some $i \geq 0$, $j \geq 1$, and $r, t \in \Sigma^*$ such that rt is primitive (see, e.g., [2]). The next proposition immediately follows from this; note that the θ -commutativity equation guarantees that both r, t are θ -palindromes.

Proposition 7 ([10]) *For $x, y \in \Sigma^+$, the solutions of $xy = \theta(y)x$ are given by $x = r(tr)^i$ and $y = (tr)^j$ for some $i \geq 0$, $j \geq 1$, and θ -palindromes r, t such that rt is primitive.*

Although this equation does not cause even a weak defect effect, one encounters it often when considering word equations which involve θ . Note that for words $u, v \in \Sigma^*$, it was proved in [4] that the system $uv = \theta(uv)$ and $vu = \theta(vu)$ causes a weak defect effect: $u, v \in \{t, \theta(t)\}^*$ for some $t \in \Sigma^+$. Thus for words x, y, z satisfying

$xy = zx$, if both y and z are θ -palindromes, then the representation of solutions of $xy = zx$ implies $tr = \theta(tr)$ and $rt = \theta(rt)$. Hence the next result holds.

Proposition 8 ([3]) *For a word $x \in \Sigma^+$ and two θ -palindromes $y, z \in \Sigma^+$, the equation $xy = zx$ implies that $x, y, z \in \{t, \theta(t)\}^*$ for some $t \in \Sigma^+$.*

3. Properties of Pseudo-Primitive Words

The primitivity of words is one of the most essential notions in combinatorics on words. The past few decades saw a considerable number of studies on this topic (see e.g., [2, 12, 16]). In contrast, research on the pseudo-primitivity of words has just been initiated in [3, 4]. For instance, although the class of pseudo-primitive words was proved to be properly included in that of primitive words [4], nothing else is known about the relation between these two classes. The purpose of this section is to prove various properties of pseudo-primitive words.

Throughout this section, θ is assumed to be a given antimorphic involution. We begin this section with a simple extension of a known result on the primitive root (Lemma 9) to the θ -primitive root (Lemma 10).

Lemma 9 (e.g., [16]) *For words $u, v \in \Sigma^+$ and a primitive word $w \in \Sigma^+$, the following properties hold:*

1. $u^n \in w^+$ implies $u \in w^+$;
2. $uv, u \in w^+$ or $uv, v \in w^+$ implies $u, v \in w^+$.

Lemma 10. *For words $u, v \in \Sigma^+$ and a θ -primitive word $x \in \Sigma^+$, the following properties hold:*

1. for some $n \geq 1$, $u^n \in \{x, \theta(x)\}^+$ implies $u \in \{x, \theta(x)\}^+$;
2. $uv, u \in \{x, \theta(x)\}^+$, or $uv, v \in \{x, \theta(x)\}^+$ implies $u, v \in \{x, \theta(x)\}^+$;
3. $\theta(u)v, u \in \{x, \theta(x)\}^+$, or $u\theta(v), v \in \{x, \theta(x)\}^+$ implies $u, v \in \{x, \theta(x)\}^+$.

Proof. The first property follows from Theorem 5, while the others are immediately proved by comparing the length of words. \square

As mentioned in the introduction, if a word w is primitive, then the equation $w^2 = ywz$ implies either $y = \lambda$ or $z = \lambda$. Since a θ -primitive word is primitive, this applies to θ -primitive words, too; a θ -primitive word x cannot be a proper infix of x^2 . However, due to the informational equivalence between x and $\theta(x)$, we should consider equations like $x^2 = y\theta(x)z$ as well, and in fact this equation can hold with non-empty y and z . Nevertheless, we can state an analogous theorem based on the next lemma.

Lemma 11 ([4]) *Let $x \in \Sigma^+$ be a θ -primitive word, and $x_1, x_2, x_3, x_4 \in \{x, \theta(x)\}$. If $x_1x_2y = zx_3x_4$ for some non-empty words $y, z \in \Sigma^+$ with $|y|, |z| < |x|$, then $x_2 \neq x_3$.*

Theorem 12. For a θ -primitive word $x \in \Sigma^+$, neither $x\theta(x)$ nor $\theta(x)x$ can be a proper infix of a word in $\{x, \theta(x)\}^3$.

Proof. Let $x_1, x_2, x_3 \in \{x, \theta(x)\}$ and suppose that $x\theta(x)$ is a proper infix of $x_1x_2x_3$. That is to say, there exist words $y, z, y', z' \in \Sigma^+$, $0 < |y|, |z|, |y'|, |z'| < |x|$ such that $zx\theta(x) = x_1x_2y$ and $x\theta(x)y' = z'x_2x_3$. By Lemma 11, the first equation implies that $x_2 \neq x$ and the second that $x_2 \neq \theta(x)$, this is in contradiction with $x_2 \in \{x, \theta(x)\}$. We prove similarly that $\theta(x)x$ cannot be a proper infix of $x_1x_2x_3$. \square

This theorem will lead us to two propositions (Propositions 16 and 20), as well as to several other results. The main usage of these propositions in this paper is the following “splitting strategy,” which shall prove useful in solving ExLS equations in Section 4. Given “complicated” words in $\{x, \theta(x)\}^+$ for a θ -primitive word x , these propositions make it possible to split such words into “simple” component words which are still in $\{x, \theta(x)\}^+$. Then, Lemmas 9 and 10 are often applicable to subdivide these simple components into smaller units in $\{x, \theta(x)\}^+$.

Recall that a primitive word cannot be a proper infix of its square. It is hence evident that for a primitive word w , if a word u in w^+ contains w as its infix like $u = ywz$ for some $y, z \in \Sigma^*$, then $y, z \in w^*$. For such w , more generally, $v, ywz \in w^+$ implies $y, z \in w^*$. This raises a naturally extended question of whether for a θ -primitive word x , if $v, ywz \in \{x, \theta(x)\}^+$, then $y, z \in \{x, \theta(x)\}^*$ holds or not. Although this is not always the case, we provide some positive cases based on the following lemma, which is a natural consequence of Theorem 12.

Lemma 13. Let x be a θ -primitive word, and $v \in \Sigma^+$. For $y, z \in \Sigma^*$, either $yx\theta(x)z \in \{x, \theta(x)\}^*$ or $y\theta(x)xz \in \{x, \theta(x)\}^*$ implies $y, z \in \{x, \theta(x)\}^*$.

Proof. We prove that $yx\theta(x)z \in \{x, \theta(x)\}^*$ implies $y, z \in \{x, \theta(x)\}^*$. Let $yx\theta(x)z = x_1 \cdots x_n$ for some $n \geq 2$ and $x_1, \dots, x_n \in \{x, \theta(x)\}$. In light of Theorem 12, there must exist such i that $y = x_1 \cdots x_{i-1}$, $x\theta(x) = x_i x_{i+1}$, and $z = x_{i+2} \cdots x_n$. \square

Lemma 14. Let x be a θ -primitive word, and $v \in \Sigma^+$. If $v, ywz \in \{x, \theta(x)\}^*$ for some $y, z \in \Sigma^*$ and either $x\theta(x)$ or $\theta(x)x$ is an infix of v , then $y, z \in \{x, \theta(x)\}^*$.

Proof. Here we consider only the case when $x\theta(x)$ is an infix of v . Due to Lemma 13, we can let $v = x'x\theta(x)x''$ for some $x', x'' \in \{x, \theta(x)\}^*$. Thus, $ywz = yx'x\theta(x)x''z \in \{x, \theta(x)\}^{\geq 2}$. From this, the same lemma derives $yx', x''z \in \{x, \theta(x)\}^*$. Based on Lemma 10, we obtain $y, z \in \{x, \theta(x)\}^*$. \square

Lemma 14 is a generalization of Lemma 13, and makes it possible to prove the following two propositions.

Proposition 15. Let x be a θ -primitive word, and $v \in \Sigma^+$. If $v, ywz \in \{x, \theta(x)\}^{\geq 2}$ for some $y, z \in \Sigma^*$ and v is primitive, then $y, z \in \{x, \theta(x)\}^*$.

Proof. Let $v = x_1 \cdots x_m$ for some $m \geq 2$ and $x_1, \dots, x_m \in \{x, \theta(x)\}$. Since v is primitive, there exists $1 \leq i \leq m$ such that $x_i x_{i+1} \in \{x\theta(x), \theta(x)x\}$. Now we can employ Lemma 14 to get this result. \square

Proposition 16. *Let x be a θ -primitive word, and $v \in \Sigma^+$. If $v, yvz \in \{x, \theta(x)\}^+$ for some $y, z \in \Sigma^*$ and v is a non-empty θ -palindrome, then $y, z \in \{x, \theta(x)\}^*$.*

Proof. Let $v = x_1 \cdots x_n$ for some $n \geq 1$ and $x_1, \dots, x_n \in \{x, \theta(x)\}$. If n is odd, then $v = \theta(v)$ implies $x_{(n+1)/2} = \theta(x_{(n+1)/2})$ and this means $x = \theta(x)$. Thus we have $v, yvz \in x^+$, and hence $y, z \in x^*$. If n is even, then $x_{n/2} x_{n/2+1} \in \{x\theta(x), \theta(x)x\}$ so that $y, z \in \{x, \theta(x)\}^*$ due to Lemma 14. \square

From now on, we address the following question: “for a θ -primitive word x and two words $u, v \in \Sigma^*$ such that $uv \in \{x, \theta(x)\}^+$, under what conditions on u, v , we can say $u, v \in \{x, \theta(x)\}^*$?”. Here we provide several such conditions. Among them is Proposition 20, which serves for the splitting strategy. As its corollary, we will obtain relationships between primitive words and θ -primitive words (Corollaries 21 and 22).

Proposition 17. *Let x be a θ -primitive word, $u \in \text{Suff}(\{x, \theta(x)\}^+)$, and $v \in \text{Pref}(\{x, \theta(x)\}^+)$. If $uv = x_1 \cdots x_m$ for some integer $m \geq 2$ and $x_1, \dots, x_m \in \{x, \theta(x)\}$, then either $u, v \in \{x, \theta(x)\}^+$ or $x_1 = \cdots = x_m$.*

Proof. Let us prove that when $u, v \notin \{x, \theta(x)\}^+$, $x_1 = \cdots = x_m$ must hold. Let $u = z'_s x'_{i-1} \cdots x'_1$ for some $i \geq 1$, $x'_i, \dots, x'_1 \in \{x, \theta(x)\}$, and some non-empty words $z'_p, z'_s \in \Sigma^+$ such that $z'_p z'_s = x'_i$. We can also let $v = x''_1 \cdots x''_{j-1} z''_p$ for some $j \geq 1$, $x''_1, \dots, x''_j \in \{x, \theta(x)\}$, and $z''_p, z''_s \in \Sigma^+$ such that $z''_p z''_s = x_j$. Now we have $x'_i \cdots x'_1 x''_1 \cdots x''_j = z'_p u v z''_s = z'_p x_1 \cdots x_m z''_s$. Since $0 < |z'_p| < |x|$, Theorem 12 implies $x_1 = \cdots = x_m$. \square

Corollary 18. *Let x be a θ -primitive word, and $u \in \text{Suff}(\{x, \theta(x)\}^+)$, $v \in \text{Pref}(\{x, \theta(x)\}^+)$. If uv is in $\{x, \theta(x)\}^{\geq 2}$ and primitive, then $u, v \in \{x, \theta(x)\}^+$.*

Proposition 17 gives the following two propositions which play an important role in investigating the ExLS equation.

Proposition 19. *Let x be a θ -primitive word, and $u, v \in \Sigma^+$. If $uv, vu \in \{x, \theta(x)\}^n$ for some $n \geq 2$, then one of the following statements holds:*

1. $u, v \in \{x, \theta(x)\}^+$;
2. $uv = x^n$ and $vu = \theta(x)^n$;
3. $uv = \theta(x)^n$ and $vu = x^n$.

Proof. We have $v \in \text{Pref}(\{x, \theta(x)\}^+)$ and $u \in \text{Suff}(\{x, \theta(x)\}^+)$ because $vu \in \{x, \theta(x)\}^n$. Proposition 17 implies that either the first property holds or $uv \in$

$\{x^n, \theta(x)^n\}$. Here we consider only the case when $uv = x^n$. Then $u = x^i x_p$ and $v = x_s x^{n-i-1}$ for some $1 \leq i \leq n$ and $x_p, x_s \in \Sigma^+$ with $x = x_p x_s$. Thus, we have $x_p v u x_s = x^{n+1}$, from which can deduce $vu = \theta(x)^n$ with the aid of Theorem 12 and the fact that x cannot be a proper infix of its square. \square

Proposition 20. *Let $x \in \Sigma^+$ be a θ -primitive word, and $p, q \in \Sigma^+$ be θ -palindromes. If pq is primitive, and $pq = x_1 \cdots x_n$ for some $n \geq 2$ and $x_1, \dots, x_n \in \{x, \theta(x)\}$, then there are integers $k, m \geq 1$ such that $n = 2m$, $p = x_1 \cdots x_{2k}$, and $q = x_{2k+1} \cdots x_{2m}$.*

Proof. It is clear from $pq = x_1 \cdots x_n$ that $p \in \text{Pref}(\{x, \theta(x)\}^+)$ and $q \in \text{Suff}(\{x, \theta(x)\}^+)$. Since both p and q are θ -palindromes, these mean that $p \in \text{Suff}(\{x, \theta(x)\}^+)$ and $q \in \text{Pref}(\{x, \theta(x)\}^+)$. Hence we can apply Proposition 17 to obtain $p = x_1 \cdots x_i$ and $q = x_{i+1} \cdots x_n$ for some i (since pq is primitive, the case $x_1 = \cdots = x_n$ is impossible).

The integer i has to be even ($i = 2k$ for some $k \geq 1$). Suppose not, then p being a θ -palindrome implies that $x_{(i+1)/2}$ is a θ -palindrome, and hence so is x . As a result, $pq = x^n$ but this contradicts the assumption that pq is primitive. Similarly, $n - i$ proves to be even, too, and we obtain $n = 2m$. \square

The next two corollaries follow from Proposition 20. The first one provides us with a sufficient condition for a primitive word that is a catenation of two non-empty θ -palindromes to be θ -primitive.

Corollary 21. *For non-empty θ -palindromes p, q , if pq is primitive but there does not exist any x such that $p, q \in \{x, \theta(x)\}^+$, then pq is θ -primitive.*

Corollary 22. *Let p, q be non-empty θ -palindromes such that pq is primitive. Then some word in $\{p, q\}^+$ is θ -primitive if and only if pq is θ -primitive.*

Proof. The converse implication is trivial because $pq \in \{p, q\}^+$. The direct implication can be proved by considering its contrapositive, which is immediately given by Proposition 20. \square

Note that in the statement of Corollary 22 we cannot replace the quantifier “some” with “all”. A trivial example is $(pq)^2 \in \{p, q\}^+$, which is not even primitive. We can also provide a non-trivial example as follows:

Example 23. *Let θ be the mirror image over $\{a, b\}^*$, $p = a$, and $q = baaab$. It is clear that $pq = abaaab$ is θ -primitive. On the other hand, $qppp = (baaa)^2 \in \{p, q\}^+$ is not even primitive.*

Corollary 22 gives a further corollary about the case in which a word obtained from a θ -primitive word by cyclic permutation remains θ -primitive.

Corollary 24. *For two non-empty θ -palindromes p, q , if pq is θ -primitive, then qp is θ -primitive.*

Proof. Since pq is θ -primitive, it is primitive and hence its conjugate qp is also primitive. Applying Corollary 22 to qp gives the result. \square

Corollary 24 gives a partial answer to one of our questions on the preservation of θ -primitivity under cyclic permutation.

Now let us examine the equation $pq = x_1 \cdots x_n$ from a different perspective to get some results useful in Section 4. Here we see that the assumptions considered in Proposition 20: pq being primitive and both of p, q being a θ -palindrome are critical to obtain $p, q \in \{x, \theta(x)\}^+$.

Lemma 25. *For a θ -primitive word $x \in \Sigma^+$ and $k \geq 2$, let $x_1, x_2, \dots, x_k \in \{x, \theta(x)\}$. If $pz = x_1x_2 \cdots x_k$ for some θ -palindrome p and non-empty word $z \in \Sigma^+$ with $|z| < |x|$, then $x_1 = x_2 = \cdots = x_{k-1}$. Moreover, if z is also a θ -palindrome, then $x_k = x_{k-1}$.*

Proof. Due to the length condition on z , we can let $x_k = yz$ for some non-empty word $y \in \Sigma^+$. Hence we have $p = x_1x_2 \cdots x_{k-1}y$. Since p is a θ -palindrome, $p = \theta(y)\theta(x_{k-1}) \cdots \theta(x_1)$. This means that $\theta(x_{k-1}) \cdots \theta(x_1)$ is a proper infix of $x_1 \cdots x_k$, and we can say that $x_1 = \cdots = x_{k-1}$ using Theorem 12 (we can assume $k \geq 3$, since if $k = 2$ the consequence is trivial).

Now we consider the additional result when $z = \theta(z)$. Without loss of generality, we can assume that $x_1 = x$. So we have $p = x^{k-1}y = \theta(y)\theta(x)^{k-1}$. Since $|y| < |\theta(x)|$, this equation gives $\theta(x) = qy$ for some non-empty word q . Actually q is a θ -palindrome. Indeed, we have $qy \in \text{Suff}(p) = \text{Suff}(x^{k-1}y)$, hence as $|q| < |x|$, $q \in \text{Suff}(x)$. Moreover, by definition, $q \in \text{Pref}(\theta(x))$, therefore $\theta(q) \in \text{Suff}(x)$ and thus q has to be a θ -palindrome.

Thus, if $x_k = \theta(x)$, then $\theta(x) = qy = yz$ and hence $\theta(x)$ could not be θ -primitive due to Proposition 8, raising a contradiction. \square

For two θ -palindromes p, q , a θ -primitive word x , and $x_1, \dots, x_k \in \{x, \theta(x)\}$ ($k \geq 1$), if $|q| < |x|$, then the equation $pq = x_1 \cdots x_k$ turns into $pq = x^k$ due to Lemma 25 and its solution is $x = p'q$ for some θ -palindrome p' such that $p = x^{k-1}p'$. If we replace q in this equation with a word z , which is not assumed to be a θ -palindrome, and if $k \geq 3$, then we can still find an intriguing non-trivial solution to the equation $pz = x^{k-1}\theta(x)$.

Example 26. *Let p be a θ -palindrome, x be a θ -primitive word, and $z \in \Sigma^+$ with $|z| < |x|$. For some $i \geq 0$, $j \geq 1$, $k \geq 3$, and θ -palindromes r, t such that rt is primitive, we can see that $x = [r(tr)^i]^2(tr)^j$, $p = x^{k-1}r(tr)^i$, and $z = (tr)^j r(tr)^i$ satisfy $pz = x^{k-1}\theta(x)$.*

Note that r and t in this example are given by Proposition 7. Further research on the properties of words in $\{r(tr)^i, (tr)^j\}^*$ may shed light on the properties of θ -primitive words. In Section 4.2, we will provide some results along this line, such as the ones in Propositions 34 and 35.

4. Extended Lyndon-Schützenberger equation

As an application of the results obtained in Section 3, we address some open cases of the extended Lyndon-Schützenberger equation in this section.

For $u, v, w \in \Sigma^+$, the ExLS equation under consideration is of the form

$$u_1 \cdots u_\ell = v_1 \cdots v_n w_1 \cdots w_m,$$

where $u_1, \dots, u_\ell \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$, for $\ell, n, m \geq 2$. The open cases are $\ell \in \{2, 3, 4\}$ and $m, n \geq 3$ (see Table 1). It suffices to consider the case when both v and w are θ -primitive; otherwise we simply replace them with their θ -primitive roots and increase the parameters n and m . The words $v_1 \cdots v_n$ and $w_1 \cdots w_m$ being symmetric with respect to their roles in the equation, it is also legitimate to assume that $|v_1 \cdots v_n| \geq |w_1 \cdots w_m|$.

Throughout Subsections 4.1 to 4.4, we prove that the triple $(4, \geq 3, \geq 3)$ imposes θ -periodicity. First of all, in Subsection 4.1, the problem which we actually work on is formalized as Problem 1, and we solve some special instances of ExLS equation to which the application of the generalized Fine and Wilf's theorem (Theorem 5) immediately proves the existence of a word t satisfying $u, v, w \in \{t, \theta(t)\}^+$. We call such instances *trivial ExLS equations*. In Subsection 4.2, we provide additional conditions which can be assumed for non-trivial ExLS equations. Several lemmas and propositions are also proved there. They are interesting in their own and our proof techniques for them probably include various applications beyond the investigation on the non-trivial ExLS equations in Subsection 4.3 (the case when $u_2 = u_1$) and Subsection 4.4 (the case when $u_2 \neq u_1$). In each of these subsections, we analyze four cases depending on the values of u_3 and u_4 one at a time. All of these proofs merely consist of direct applications of the results obtained so far and in Subsection 4.2.

In Subsection 4.5, we prove that for $n, m \geq 2$, the triple $(3, n, m)$ does not impose θ -periodicity. We provide several (parametrized) examples which verify that for some specific values of n, m , the triple $(3, n, m)$ does not impose θ -periodicity. Our survey will expose complex behaviors of $(3, n, m)$ ExLS equations.

4.1. Problem setting for the ExLS equation $\ell = 4$

Taking the assumptions mentioned above into consideration, the problem which we are addressing is described as follows:

Problem 1. *Let $u, v, w \in \Sigma^+$ and integers $n, m \geq 3$. Let $u_1, u_2, u_3, u_4 \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$. Does the equation $u_1 u_2 u_3 u_4 =$*

$v_1 \cdots v_n w_1 \cdots w_m$ imply $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$ under all of the following conditions?

1. v and w are θ -primitive,
2. $|v_1 \cdots v_n| \geq |w_1 \cdots w_m|$,
3. $u_1 = u$, $v_1 = v$, and $w_m = w$,
4. $|v|, |w| < |u|$.

The condition 2 means that $2|u| \leq n|v|$. Besides, the condition 4 follows from the conditions 1 and 2 as shown in the next lemma.

Lemma 27. *Let $u, v, w \in \Sigma^+$ such that v, w are θ -primitive. If $u_1 u_2 u_3 u_4 = v_1 \cdots v_n w_1 \cdots w_m$ for some $n, m \geq 3$, $u_1, u_2, u_3, u_4 \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$, then $|v| < |u|$ and $|w| < |u|$.*

Proof. Due to Condition 2, $|v_1 \cdots v_n| \geq |w_1 \cdots w_m|$. This means that $m|w| \leq 2|u|$, which in turn implies $|w| \leq \frac{2}{3}|u|$ because $m \geq 3$. Thus $|w| < |u|$.

Now suppose that the ExLS equation held with $|v| \geq |u|$. Then $v_1 \cdots v_n$ is a prefix of $u_1 u_2 u_3 u_4$ of length at least $3|v| \geq 2|v| + |u|$, and hence $u, v \in \{t, \theta(t)\}^+$ for some θ -primitive word $t \in \Sigma^+$ due to Theorem 6. Unless $|v| = |u|$, we reach the contradiction that v would not be θ -primitive. Even if $|v| = |u|$, we have $u_4 = w_1 \cdots w_m$. Therefore $v_1 = u_1$ could not be θ -primitive. \square

The next lemma reduces the number of steps required to prove a positive answer to Problem 1.

Lemma 28. *Under the setting of Problem 1, if $u, v \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$, then $w \in \{t, \theta(t)\}^+$.*

In fact, we can say more strongly that if two of u, v, w are proved to be in $\{t, \theta(t)\}^+$ for some t , then the other one is also in this set.

First of all, we distinguish the case in which the existence of such t that $u, v, w \in \{t, \theta(t)\}^+$ is trivial due to the generalized Fine and Wilf theorem (Theorem 5).

Theorem 29. *Under the setting of Problem 1, if there exists an index i , $1 \leq i \leq n$, such that $u_1 u_2 = v_1 \cdots v_i$, then $u, v, w \in \{t, \theta(t)\}^+$ for some word $t \in \Sigma^+$.*

Proof. Since v is assumed to be θ -primitive, Theorem 5 implies $u \in \{v, \theta(v)\}^+$. Then $w \in \{v, \theta(v)\}^+$ due to Lemma 28 (in fact, $w \in \{v, \theta(v)\}$ because w is also assumed to be θ -primitive). \square

If a given $(4, n, m)$ ExLS equation satisfies the condition in Theorem 29, then we say that this equation is *trivial*. Before initiating our study on non-trivial ExLS equations, we provide one important condition which makes the equation trivial according to the generalized Fine and Wilf theorem (Theorem 6).

Proposition 30. *Under the setting of Problem 1, if $n|v| \geq 2|u| + |v|$, then the equation is trivial.*

Proof. We can employ Theorem 6 to obtain $u, v \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$. In fact, t is either v or $\theta(v)$ because v is assumed to be θ -primitive. Hence we can find such i stated in Theorem 29, and by definition this equation is trivial. \square

4.2. Non-trivial $(4, \geq 3, \geq 3)$ ExLS equations and related combinatorial results

Now we shift our attention to the non-trivial $(4, \geq 3, \geq 3)$ ExLS equation. What we will actually prove here is that under the setting of Problem 1, any non-trivial equation cannot hold. Along with Theorem 29, this implies that $(4, \geq 3, \geq 3)$ imposes θ -periodicity.

From this theorem and Proposition 30, the equation is non-trivial if and only if $(n - 1)|v| < 2|u| < n|v|$. Thus, the next proposition, which was proposed in [3] to decrease the amount of case analyses for the $(5, \geq 3, \geq 3)$ ExLS equation, is still available for the investigation of non-trivial $(4, \geq 3, \geq 3)$ ExLS equations.

Proposition 31 ([3]) *Let $u, v \in \Sigma^+$ such that v is θ -primitive, $u_2, u_3 \in \{u, \theta(u)\}$, and $v_2, \dots, v_n \in \{v, \theta(v)\}$ for some integer $n \geq 3$. If $vv_2 \cdots v_n \in \text{Pref}(uu_2u_3)$ and $(n - 1)|v| < 2|u| < n|v|$, then there are only two possible cases.*

1. $u_2 = \theta(u)$: and $v_2 = \cdots = v_n = v$ with $u\theta(u) = (pq)^{n-1}p$ and $v = pq$ for some non-empty θ -palindromes p, q .
2. $u_2 = u$: n is even, $v_2 = \cdots = v_{n/2} = v$, and $v_{n/2+1} = \cdots = v_n = \theta(v)$ with $v = r(tr)^i(rt)^{i+j}r$ and $u = v^{n/2-1}r(tr)^i(rt)^j$ for some $i \geq 0, j \geq 1$, and non-empty θ -palindromes r, t such that rt is primitive.

This proposition helps in proving that non-trivial $(4, \geq 3, \geq 3)$ ExLS equations verify the one more condition that $|v| \neq |w|$ as shown in the next proposition.

Proposition 32. *Non-trivial ExLS equations under the setting of Problem 1 imply $|v| \neq |w|$.*

Proof. Suppose that the equation were non-trivial with $|v| = |w|$. Combining $|v| = |w|$ and the non-trivial length condition together implies $m = n - 1$ and furthermore the border between u_2 and u_3 splits v_n into exactly halves. Hence if $u_3 = \theta(u_2)$, then $v_n = x\theta(x)$ for some $x \in \Sigma^+$, contradicting the θ -primitivity of v . Besides, due to the condition 4 of Problem 1, if $u_4 = \theta(u_1)$, then $w = \theta(v)$, and hence $u_1u_2u_3u_4 \in \{v, \theta(v)\}^+$. Taking $(n - 1)|v| < 2|u| < n|v|$ into account, this implies that v is not θ -primitive, raising a contradiction. Therefore, the only possible solutions verify $u_3 = u_2$ and $u_4 = u_1 = u$.

If $u_2 = u_3 = u$, then according to Proposition 31, n is even, and by substituting the representations of u and v given there into $u^4 = v^{n/2}\theta(v)^{n/2}w_1 \cdots w_m$, we obtain

that $w_1 \cdots w_m = (tr)^j [r(tr)^i r(tr)^{i+j}]^{n/2-1} [r(tr)^{i+j} r(tr)^i]^{n/2-1} (rt)^j$, which is a θ -palindrome of even length. Since w is θ -primitive, m has to be even (Lemma 4). It is however impossible because $m = n - 1$ and n is even.

If $u_2 = u_3 = \theta(u)$, then Proposition 31 gives $v = pq$ and $u_1 u_2 = u\theta(u) = (pq)^{n-1} p$ for some θ -palindromes $p, q \in \Sigma^+$. Note that the left side of the ExLS equation is as long as its right side ($4|u| = n|v| + m|w| = (2n - 1)|pq|$). Substituting $2|u| = (n - 1)|pq| + |p|$ into this yields $|p| = |q|$ and it in turn implies that both p and q are of even length. Let $p = p'\theta(p')$ and $q = q'\theta(q')$ for some $p', q' \in \Sigma^+$ of the same length. Then $u_1 = u$ ends with either $\theta(p')qp'$ or $\theta(q')pq'$, and so w_m is either of them. However, neither is θ -primitive. This contradiction proves that the equation is trivial. \square

Supposing that some non-trivial ($4, \geq 3, \geq 3$) ExLS equation held, the next claim would follow from this proposition. Although our conclusion in this section will prove that this claim cannot hold, the equation proposed there, $u_3 u_4 = qw_1 \cdots w_m$, or more generally the relation $qw_1 \cdots w_m \in \{u, \theta(u)\}^{\geq 2}$ provides in its own right challenging themes.

Claim 1. *Under the setting of Problem 1, if the ExLS equation were non-trivial, then we would have $u_3 u_4 = qw_1 \cdots w_m$ for some non-empty θ -palindrome q .*

Proof. According to the presentations of u and v given in Proposition 31, if $u_2 = \theta(u)$, then $u\theta(u)q = v^n$ and hence $u_3 u_4 = qw_1 \cdots w_m$; otherwise, $uu[r(tr)^i]^2 = v^{n/2}\theta(v)^{n/2}$ so that $u_3 u_4 = [r(tr)^i]^2 w_1 \cdots w_m$. Since q, r, t are θ -palindromes, this claim holds. \square

As we shall see soon in Claim 2, the next lemma is of use when considering non-trivial ExLS equations with $u_3 \neq u_4$, that is, $u_3 u_4$ being a θ -palindrome.

Lemma 33. *Let p, q be non-empty θ -palindromes and let w be a θ -primitive word. For some $k \geq 1$ and words $w_1, \dots, w_k \in \{w, \theta(w)\}$, if $p = qw_1 \cdots w_k$ holds, then either $p, q \in \{w, \theta(w)\}^+$ or $w_1 = \cdots = w_k$.*

Proof. First we prove that $q \in \text{Suff}((w_1 \cdots w_k)^+)$. Since $w_1 \cdots w_k \in \text{Suff}(p)$, p being a θ -palindrome implies $\theta(w_1 \cdots w_k) \in \text{Pref}(p)$. Thus if $|q| \leq k|w|$, then $q \in \text{Pref}(\theta(w_1 \cdots w_k))$, that is, $q \in \text{Suff}(w_1 \cdots w_k)$ and we are done. Otherwise, $w_1 \cdots w_k \in \text{Suff}(q)$ so that $(w_1 \cdots w_k)^2 \in \text{Suff}(p)$. By repeating this process, eventually we will find some integer $i \geq 1$ such that $q \in \text{Suff}((w_1 \cdots w_k)^i)$.

If $q \in \{w, \theta(w)\}^+$, then obviously $p \in \{w, \theta(w)\}^+$. Otherwise, let $q = w'w_{j+1} \cdots w_k(w_1 \cdots w_k)^i$ for some $1 \leq j \leq k$ and $i \geq 0$, where w' is a non-empty proper suffix of w_j . Then, $p = w'w_{j+1} \cdots w_k(w_1 \cdots w_k)^{i+1}$ overlaps in a non-trivial way with $p = \theta(p) = (\theta(w_k) \cdots \theta(w_1))^{i+1} \theta(w_k) \cdots \theta(w_{j+1}) \theta(w')$, and Theorem 12 implies that $w_1 = \cdots = w_k$. \square

Claim 2. Under the setting of Problem 1, if the ExLS equation were non-trivial and $u_3 \neq u_4$, then $w_1 = \dots = w_m = w$ and $u_3u_4 \in \text{Suff}(w^+)$.

Proof. We have $u_3u_4 = xw_1 \dots w_m$ for some non-empty θ -palindrome $x \in \Sigma^+$ due to Proposition 31. As suggested before, we can employ Lemma 33 to get either $x, u_3u_4 \in \{w, \theta(w)\}^+$ or $w_1 = \dots = w_m$. In the first case, Theorem 5 implies $u \in \{w, \theta(w)\}^+$ because w is assumed to be θ -primitive. Then the ExLS equation in turn implies that $v_1 \dots v_n \in \{w, \theta(w)\}^+$ and hence $v \in \{w, \theta(w)\}$ for the same reason. As a result the equation would be trivial. Consequently $w_1 = \dots = w_m$. \square

The main strategy used in the analyses of non-trivial ExLS equations is to split $w_1 \dots w_m$ into smaller components which are still in $\{w, \theta(w)\}^+$, until we reach a contradiction. The split is mainly achieved by Propositions 16 and 20. Note that the word to which Proposition 20 is applied must be primitive. The next two lemmas work for this purpose in Subsection 4.3, but we provide them in more general form. An interesting point is that Lyndon and Schützenberger’s original result (Theorem 1) plays an essential role in their proofs; hence for the ExLS equation.

Proposition 34. Let $r, t \in \Sigma^+$ such that rt is primitive. For any $i \geq 0, j, k \geq 1$, and $n \geq 2$, $(tr)^j[(r(tr)^i)^n(tr)^j]^k$ is primitive.

Proof. Suppose that the given word were not primitive; namely, for some $\ell \geq 2$ and a primitive word x , let $(tr)^j[(r(tr)^i)^n(tr)^j]^k = x^\ell$. Catenating $(r(tr)^i)^n$ to the left to the both sides of this equation gives $[(r(tr)^i)^n(tr)^j]^{k+1} = (r(tr)^i)^n x^\ell$. As $k \geq 1$ and $n, \ell \geq 2$, we can apply Theorem 1 to this equation to obtain $\rho((r(tr)^i)^n(tr)^j) = \rho(r(tr)^i) = x$. Using Lemma 9, one can obtain $\rho((tr)^j) = x$, and furthermore, $\rho(tr) = x$. Combining this with $\rho(r(tr)^i) = x$ gives us $\rho(r) = \rho(t)$ and hence rt would not be primitive, which contradicts the hypotheses. \square

Proposition 35. Let $r, t \in \Sigma^+$ such that rt is primitive. For any $i \geq 0, j, k, m \geq 1$, $(tr)^j[(r(tr)^i)^m(tr)^j]^{k-1}(r(tr)^i)^{m-1}(rt)^j$ is primitive.

Proof. Suppose that we had $(tr)^j[(r(tr)^i)^m(tr)^j]^{k-1}(r(tr)^i)^{m-1}(rt)^j = x^\ell$ for some primitive word x and $\ell \geq 2$. Catenating $(r(tr)^i)^{m+1}$ to the right to the both sides of this equation gives $[(tr)^j(r(tr)^i)^m]^{k+1} = x^\ell(r(tr)^i)^{m+1}$. Now as in the proof of Proposition 34, we reach the contradicting conclusion that rt is not primitive. \square

There are some results which can be used for the splitting strategy, once we apply Proposition 31 to non-trivial ExLS equations with $u_1 \neq u_2$, which will be considered in Subsection 4.4. As before, they are provided in more general form than required for the purpose.

Lemma 36. Let $z, w \in \Sigma^+$ with $|z| < |w|$ and let p be a θ -palindrome. If $zp = w^n$ for some $n \geq 2$, then $z = \theta(z)$.

Proof. Let $w = zy$ for some $y \in \Sigma^+$. Then $p = y(zy)^{n-1}$, from which we can obtain $y = \theta(y)$ and $z = \theta(z)$ because $p = \theta(p)$ and $n - 1 \geq 1$. \square

Proposition 37. *Let x be a θ -primitive word, $u \in \Sigma^+$, and q be a non-empty θ -palindrome. If for some $n \geq 2$ and $\ell \geq 1$, $u[\theta(u)q^n u]^\ell \in \{x, \theta(x)\}^{\geq 2}$, then $u, q \in \{x, \theta(x)\}^+$.*

Proof. Let $u[\theta(u)q^n u]^\ell = x_1 \cdots x_m$ for some $m \geq 2$ and $x_1, \dots, x_m \in \{x, \theta(x)\}$. Let $u = x_1 \cdots x_{k-1} z_1$ and $[\theta(u)q^n u]^\ell = z_2 x_{k+1} \cdots x_m$ for some $1 \leq k \leq m$ with $x_k = z_1 z_2$ and $z_1 \neq \lambda$, i.e. $|z_2| < |x|$. If $z_2 = \lambda$, then $u, [\theta(u)q^n u]^\ell \in \{x, \theta(x)\}^+$. Lemma 10 implies $\theta(u)q^n u \in \{x, \theta(x)\}^+$ and the same lemma further gives $q^n \in \{x, \theta(x)\}^+$, that is, $q \in \{x, \theta(x)\}^+$.

Now we prove that z_2 cannot be non-empty. Without loss of generality, we assume $x_m = x$. So suppose $z_2 \neq \lambda$ ($0 < |z_1| < |x|$). We can apply Lemma 25 to $z_1[\theta(u)q^n u]^\ell = x_k \cdots x_m$ to get $x_{k+1} = \cdots = x_m = x$ because $[\theta(u)q^n u]^\ell$ is a θ -palindrome and $|z_1| < |x|$. Thus if $|x| \leq |u|$, then $|z_2| < |u|$ and so $[\theta(u)q^n u]^\ell = z_2 x^{k-1}$ gives $x \in \text{Suff}(u)$ and hence $\theta(x) \in \text{Pref}(\theta(u))$. These further imply that $x \in \text{Suff}(x_1 \cdots x_{k-1} z_1)$ and $\theta(x) \in \text{Pref}(z_2 x_{k+1} \cdots x_m)$. Thus $x\theta(x)$ is a proper infix of $x_{k+1} x_k x_{k-1}$, which is in contradiction with the θ -primitivity of x by Theorem 12.

Therefore, $|x| > |u|$, which means $k = 1$, that is, we have $x_2 = \cdots = x_m = x$. Note that $x \neq \theta(x)$ must hold because of $z_2 x^{m-1}$ being a θ -palindrome, $0 < |z_2| < |x|$ and x is primitive (and cannot be a proper infix of its square). If $x_1 = \theta(x)$, then $u \in \text{Pref}(\theta(x)) \cap \text{Suff}(x)$ holds and so $u = \theta(u)$. Now Lemma 25 would imply $x_1 = x$, which contradicts $x \neq \theta(x)$. Otherwise ($x_1 = x$), $u[\theta(u)q^n u]^\ell = x^m$ and from this Lemma 36 derives $u = \theta(u)$. Then we have $u(uq^n u)^\ell = x^m$; in other words, $(uq^n u)^{\ell+1}$ and x^m share a suffix of length at least $\eta = \max(m|x|, \ell|uq^n u|)$. If $\ell \geq 2$, then $\eta \geq |x| + |uq^n u|$, and the Fine and Wilf theorem implies $\rho(uq^n u) = x$. With $u(uq^n u)^\ell = x^m$, this implies $\rho(u) = x$. However, this contradicts $|u| < |x|$. If $\ell = 1$, then $uuq^n u = x^m$. Using cyclic permutation, we obtain $u^3 q^n = x'^m$, where x' is a conjugate of x . This is of the form of LS equation, and Theorem 1 concludes $\rho(u) = \rho(q) = x'$. Now we reached the same contradiction because $|x'| = |x|$. \square

Lemma 38. *Let w be a θ -primitive word, and $w_1, \dots, w_m \in \{w, \theta(w)\}$ for some $m \geq 2$. Let $u, q \in \Sigma^+$ such that q is a θ -palindrome with $|q| < |u|$. If $u^2 = qw_1 \cdots w_m$, then either $u, q \in \{w, \theta(w)\}^+$ or $u = qr$ for some non-empty θ -palindrome r .*

Proof. It is trivial that the case $u, q \in \{w, \theta(w)\}^+$ is possible. Hence assume that $u, q \notin \{w, \theta(w)\}^+$. Without loss of generality, we can also assume that $w_m = w$. Let $u = qr$ for some $r \in \Sigma^+$. Then $rqr = w_1 \cdots w_m$. We prove that r is a θ -palindrome. Let $r = w_1 \cdots w_{k-1} z_1 = z_2 w_{m-k+2} \cdots w_m$ for some $k \geq 1$, where $z_1 \in \text{Pref}(w_k)$ and $z_2 \in \text{Suff}(w_{m-k+1})$ with $|z_1| = |z_2| < |w|$. If $z_1 = \lambda$, then $r \in \{w, \theta(w)\}^+$ and then $rqr = w_m \cdots w_1$ implies $q \in \{w, \theta(w)\}^+$ by Lemma 10, but this contradicts the assumption. Thus $z_1 \neq \lambda$. Then we have two cases, $k \geq 2$ and $k = 1$. Lemma 11

(for $k = 2$) or Theorem 12 (for $k \geq 3$) works to give $w_1 = \dots = w_{k-1} = \theta(w)$ and $w_{m-k+2} = \dots = w_m = w$. Thus, $z_2 = \theta(z_1)$ and hence $r = \theta(r)$. Even for $k = 1$, if $w_1 \neq w_m$, then $r \in \text{Pref}(\theta(w)) \cap \text{Suff}(w)$ so that $r = \theta(r)$. Otherwise $w = rq_p = q_s r$ for some $q_p \in \text{Pref}(q)$ and $q_s \in \text{Suff}(q)$. Since $q = \theta(q)$, $q_s = \theta(q_p)$ so that we have $rq_p = \theta(q_p)r$. According to Proposition 7, $r = \theta(r)$. \square

Proposition 39. *Let w be a θ -primitive word, and $w_1, \dots, w_m \in \{w, \theta(w)\}$ for some odd integer $m \geq 3$. Let $u, q \in \Sigma^+$ such that q is a θ -palindrome with $|q| < |u|$. If $u^2 = qw_1 \dots w_m$, then $w = \theta(w)$. If additionally $|u| \geq 2|q|$ holds, then $\rho(u) = \rho(q) = w$.*

Proof. Lemma 38 implies that either $q, u \in \{w, \theta(w)\}^+$ or $u = qr$ for some non-empty θ -palindrome r . In the former case, let $u \in \{w, \theta(w)\}^k$ for some $k \geq 1$ and we can see $q \in \{w, \theta(w)\}^{2k-m}$ and $2k - m$ is odd because m is odd. Then $q = \theta(q)$ implies $w = \theta(w)$, and hence $u, q \in w^+$. In the latter case, we have $rqr = w_1 \dots w_m$. This implies $w_{(m+1)/2} = \theta(w_{(m+1)/2})$ (i.e., $w = \theta(w)$) because rqr is a θ -palindrome and m is odd.

Now we consider the additional hypothesis $|u| \geq 2|q|$. Since $2|u| = |q| + m|w|$, $|u| = (|q| + m|w|)/2 \geq 2|q|$, which leads to $|q| \leq \frac{1}{3}m|w|$. As seen above, $rqr = w^m$, hence $|r| = (m|w| - |q|)/2 \geq \frac{1}{3}m|w| \geq |w|$ as $m \geq 3$. With this, the equation $rqr = w^m$ gives $r = w^k w'_p = w'_s w^k$ for some $k \geq 1$, $w'_p \in \text{Pref}(w)$, and $w'_s \in \text{Suff}(w)$. Since w is primitive, w'_p and w'_s have to be empty. Consequently $\rho(r) = \rho(q) = w$ and hence $\rho(u) = w$ by using Lemma 10. \square

4.3. ExLS equation of the form $u^2 u_3 u_4 = v_1 \dots v_n w_1 \dots w_m$

In this subsection, we prove that an ExLS equation of the form $u^2 u_3 u_4 = v_1 \dots v_n w_1 \dots w_m$ implies that $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$. We have already seen that for this purpose it suffices to show that any non-trivial equation of this form cannot hold. Recall that we assumed $u_1 = u$, $v_1 = v$, and $w_m = w$, and that Proposition 32 allows us to assume $|v| \neq |w|$.

We can apply Proposition 31 to the non-trivial equation to obtain that n is an even integer except 2, $v_1 = \dots = v_{n/2} = v$ and $v_{n/2+1} = \dots = v_n = \theta(v)$ (i.e., $v_1 \dots v_n$ is a θ -palindrome), $u = [r(tr)^i r(tr)^{i+j}]^{n/2-1} r(tr)^i (rt)^j$, and $v = r(tr)^i r(tr)^{i+j}$ for some $i \geq 0$, $j \geq 1$, and non-empty θ -palindromes r, t such that rt is primitive. Actually rt has to be θ -primitive due to Corollary 22 because $v \in \{r, t\}^+$ is assumed to be θ -primitive. Let us now study all possible values of $u_3 u_4$.

Proposition 40. *Under the setting of Problem 1, if $u_1 u_2 u_3 u_4 = u^4$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. According to the representations of u and v in terms of r and t , we obtain

$$w_1 \dots w_m = (tr)^j [(r(tr)^i)^2 (tr)^j]^{n/2-1} [(rt)^j (r(tr)^i)^2]^{n/2-1} (rt)^j .$$

This expression is a θ -palindrome of even length and hence m has to be even (Lemma 4). Therefore, $w_1 \cdots w_{m/2} = [(tr)^j (r(tr)^i)^2]^{n/2-1} (tr)^j$, and this was proved to be primitive in Proposition 34. Moreover, its right hand side is the catenation of two θ -palindromes $p_1 = (tr)^j [r(tr)^i r(tr)^{i+j}]^{n/2-2} r(tr)^i (rt)^j$ and $p_2 = r(tr)^i$. Proposition 20 gives $p_2 = r(tr)^i \in \{w, \theta(w)\}^+$. Furthermore, applying Proposition 16 to $p_1 p_2 = (tr)^j [r(tr)^i r(tr)^{i+j}]^{n/2-2} r(tr)^i \cdot p_2 \cdot (tr)^j$ gives $(tr)^j \in \{w, \theta(w)\}^+$. Finally Lemma 10 derives $r, t \in \{w, \theta(w)\}^+$ from $r(tr)^i, (tr)^j \in \{w, \theta(w)\}^+$, but this contradicts the θ -primitivity of rt . As a result, there are no solutions to the non-trivial equation. \square

Proposition 41. *Under the setting of Problem 1, if $u_1 u_2 u_3 u_4 = u^3 \theta(u)$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. Since u_4 is $\theta(u)$ instead of u , we have $w_1 \cdots w_m = x^2 (r(tr)^i)^2$, where $x = (tr)^j [(r(tr)^i)^2 (rt)^j]^{n/2-1} r(tr)^i (rt)^j$. Claim 2 gives that $w_1 = \cdots = w_m = w$, and hence $w^m = x^2 (r(tr)^i)^2$. This is a classical LS equation; thus Theorem 1 is applicable to conclude that $\rho(x) = \rho(r(tr)^i)$. However, this contradicts the primitivity of x obtained in Proposition 35 because $|x| > |r(tr)^i|$. \square

Proposition 42. *Under the setting of Problem 1, if $u_1 u_2 u_3 u_4 = u^2 \theta(u) u$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. Since $u_3 \neq u_4$, $w_1 = \cdots = w_m = w$ due to Claim 2. Using the representations of u and v by r and t , we can see that $u_3 u_4 = \theta(u) u$ is equal to both sides of the following equation:

$$(tr)^j r(tr)^i [(rt)^j (r(tr)^i)^2]^{n/2-1} [(r(tr)^i)^2 (tr)^j]^{n/2-1} r(tr)^i (rt)^j = (r(tr)^i)^2 w^m .$$

By catenating $(r(tr)^i)^4$ to the left of both sides, we get $(r(tr)^i)^6 w^m = x^2$, where $x = (r(tr)^i)^2 [(r(tr)^i)^2 (tr)^j]^{n/2-1} r(tr)^i (rt)^j$. Then, Theorem 1 implies that $\rho(x) = \rho(r(tr)^i) = w$. Since x contains $r(tr)^i$ as its infix, the share of primitive root between x and $r(tr)^i$ gives $\rho(r(tr)^i) = \rho((rt)^j)$. We deduce from this using Lemma 9 that rt would not be primitive, which contradicts our hypothesis. \square

Proposition 43. *Under the setting of Problem 1, if $u_1 u_2 u_3 u_4 = u^2 \theta(u)^2$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. Recall that $v_1 \cdots v_n$ is a θ -palindrome. Since $u^2 \theta(u)^2$ is a θ -palindrome, $\theta(w_1 \cdots w_m)$ is one of its prefixes and the assumption $|w_1 \cdots w_m| < |v_1 \cdots v_n|$ implies that $\theta(w_1 \cdots w_m) \in \text{Pref}(v_1 \cdots v_n)$. Hence $w_1 \cdots w_m \in \text{Suff}(v_1 \cdots v_n)$ and now we have $(w_1 \cdots w_m)^2 \in \text{Suff}(u^2 \theta(u)^2)$.

We prove that this suffix is long enough to apply the extended Fine and Wilf theorem. Since $(n-1)|v| < 2|u|$ and $n \geq 4$, we have $|v| < \frac{2}{3}|u|$ and, in turn, $n|v| < 2|u| + \frac{2}{3}|u| = \frac{8}{3}|u|$. From this we obtain $m|w| > \frac{4}{3}|u|$. Then, $2m|w| - (|w| + 2|u|) > (2m-1)|w| - \frac{3}{2}m|w| = (\frac{1}{2}m-1)|w| > 0$ since $m \geq 3$. Thus, $u^2 \theta(u)^2$ and

$(w_m \cdots w_1)^2$ share a suffix of length at least $2|u| + |w|$ and Theorem 6 concludes that $u \in \{w, \theta(w)\}^+$ because w is θ -primitive. Now it is clear that also $v \in \{w, \theta(w)\}^+$, but in fact $v \in \{w, \theta(w)\}$ must hold because v is also θ -primitive. However this contradicts the assumption that $|v| \neq |w|$. \square

4.4. ExLS equation of the form $u\theta(u)u_3u_4 = v_1 \cdots v_n w_1 \cdots w_m$

Note that in the following propositions, we consider only the non-trivial equations; hence Proposition 32 allows to assume $|v| \neq |w|$.

Using Proposition 31, $u\theta(u) = (pq)^{n-1}p$ and $v_1 = \cdots = v_n = v = pq$ for some non-empty θ -palindromes p, q . Unlike the case considered before, in the current case n can be odd. In fact, if n is odd, then $u = (pq)^{(n-1)/2}y$, where $p = y\theta(y)$ for some $y \in \Sigma^+$; while if n is even, then $u = (pq)^{n/2-1}px$, where $q = x\theta(x)$ for some $x \in \Sigma^+$. Again, we consider the four cases associated with the four possible values of u_3u_4 . The last two, $u_3 = u_4 = u$ and $u_3 = u_4 = \theta(u)$, are merged and studied in two separate propositions depending on the parity of m instead.

Proposition 44. *Under the setting of Problem 1, if $u_1u_2u_3u_4 = u\theta(u)u\theta(u)$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. In this setting, $u_3u_4 = u\theta(u) = qw_1 \cdots w_m$. Since both $u\theta(u)$ and q are θ -palindromes, we can employ Claim 2 to obtain $w_1 = \cdots = w_m = w$. Now the equation turns into the LS equation $(u\theta(u))^2 = v^n w^m$, and hence $\rho(v) = \rho(w)$ due to Theorem 1. Both v and w being primitive, this contradicts the assumption $|v| \neq |w|$ and consequently the existence of non-trivial solutions. \square

Proposition 45. *Under the setting of Problem 1, if $u_1u_2u_3u_4 = u\theta(u)\theta(u)u$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. Recall that $u\theta(u) = (pq)^{n-1}p$. Claim 2 implies that $\theta(u)u = qw^m$ with $q = w'w^{k-1}$ for some $1 \leq k \leq m$ and a non-empty proper suffix w' of w .

Case 1 (n is odd): Then we have $\theta(u)u = qw^m = x_sx$, where $x_s = \theta(y)q(pq)^{(n-1)/2-1}y$ and $x = \theta(y)(pq)^{(n-1)/2}y$; note that $x_s \in \text{Suff}(x)$. One can easily calculate that $|w| = \frac{1}{m}[n|p| + (n-2)|q|]$ and $|x_s| = \frac{1}{2}(n-1)(|p| + |q|)$, and hence $|x_s| - |w| = \frac{(m-2)(n-1)-2}{2m}|p| + \frac{(m-2)(n-1)+2}{2m}|q|$, which is positive because $n, m \geq 3$. Thus we can say that x^2 and w^{m+k} share a prefix of length at least $|x| + |w|$ so that by the Fine and Wilf theorem, $\rho(x) = \rho(w) = w$. Starting from $\theta(y)qyw^m = \theta(y)yx_sx = x^2$, we can verify that $2|x| - m|w| = |pq|$, that is, $|pq|$ is a multiple of $|w|$. The suffix of x of length $|pq|$ is $\theta(y)qy$, which is w^j for some $j \geq 2$ because $|pq| = |v| \neq |w|$. Therefore, this conjugate of v is not primitive, either. This is a contradiction with the θ -primitivity of v .

Case 2 (n is even): In this case, $u = (pq)^{n/2-1}px$ for some $x \in \Sigma^+$ such that $q = x\theta(x)$. Substituting this into $\theta(u)u = qw^m$ gives

$$[\theta(x)px]^{n/2-1}\theta(x)p^2x[\theta(x)px]^{n/2-1} = x\theta(x)w^m. \tag{2}$$

From this equation, we can obtain $x = \theta(x)$ and hence $px = xz$ for some $z \in \Sigma^+$. If $|x| \geq |p|$, then Lemma 2 implies $\rho(x) = \rho(p)$ so that $v = pq = px^2$ would not be primitive. Hence $|x| < |p|$ must hold and under this condition, the solution of $px = xz$ is given by $p = xy$ and $z = yx$ for some $y \in \Sigma^+$. Since $p = \theta(p)$, we have $p = xy = \theta(y)x$. Proposition 7 gives $x = r(tr)^i$ and $y = (tr)^j$ for some $i \geq 0$, $j \geq 1$, and θ -palindromes r, t such that rt is primitive. Both of r and t should be non-empty; otherwise, $\rho(p) = \rho(x)$ and $v = pq = px^2$ would not be primitive. Substituting these into Eq. (2) yields the following equation.

$$[(tr)^j r(tr)^i [r(tr)^i r(tr)^{i+j} r(tr)^i]^{n/2-1}]^2 = w^m.$$

Since w is θ -primitive, this equation means that m has to be even. Then $w^{m/2} = (tr)^j r(tr)^i [r(tr)^i r(tr)^{i+j} r(tr)^i]^{n/2-1}$. By catenating $(r(tr)^i)^2$ from the left to the both sides of this equation, we obtain an LS equation $[r(tr)^i]^2 w^{m/2} = [r(tr)^i r(tr)^{i+j} r(tr)^i]^{n/2}$. Theorem 1 gives $\rho(r(tr)^i) = \rho(r(tr)^i r(tr)^{i+j} r(tr)^i)$ and Lemma 9 reduces it to $\rho(r) = \rho(t)$, but this contradicts the primitivity of $pq = r(tr)^{i+j} (r(tr)^i)^2$. \square

Proposition 46. *Under the setting of Problem 1, if $u_1 u_2 = u\theta(u)$, $u_3 = u_4$, and m is odd, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. We have $u_3 u_4 = q w_1 \cdots w_m$. Since $u_3 = u_4$ and $|q| < |u|$, we can employ Proposition 39 to obtain $w = \theta(w)$. Moreover, when $n \geq 5$, we have $|u| \geq 2|q|$ and the proposition also gives $\rho(u_3) = \rho(q) = w$. Since $w = \theta(w)$, we can see that $\rho(u) = w$. Then $\rho(p) = w$ because $\rho(u) = \rho(q) = w$ and $pq \in \text{Pref}(u)$. However, $\rho(p) = \rho(q)$ means that $v = pq$ would not be even primitive. Therefore in the following let n be either 3 or 4.

First we consider the case when $u_3 = u$. Then we have either $(pqy)^2 = qw^m$ (when $n = 3$) where $p = y\theta(y)$, or $(pqp)^2 = qw^m$ (when $n = 4$) where $q = x\theta(x)$, for some $x, y \in \Sigma^+$. In both cases, if $|p| \leq |q|$, Lemma 2 can be applied and we have $\rho(p) = \rho(q)$, so $v = pq$ would not be even primitive. Hence $|p| > |q|$ must hold, but then $|u| \geq 2|q|$ and then Proposition 39 implies $\rho(p) = \rho(q)$.

Next we consider the case when $u_3 = \theta(u)$ and $n = 3$. Then $\theta(u) = \theta(y)qp$ so that $\theta(y)qp\theta(y)qp = qw^m$. Let $\theta(y)q = qz$ for some z with $|y| = |z|$. Using $pq = y\theta(y)q = yqz$, from $\theta(y)qp\theta(y)qp = qw^m$ we can obtain $zyqzzy\theta(y) = w^m$. Since $w = \theta(w)$, this equation gives $z = y = \theta(y)$. Then $\theta(y)q = qz$ turns into $yq = qy$ and hence $\rho(y) = \rho(q)$ by Theorem 3. This however implies that $v = y\theta(y)q$ would not be θ -primitive.

Finally we consider the case when $u_3 = \theta(u)$ and $n = 4$. Then we have $[\theta(x)pqp]^2 = qw^m$, which gives $x = \theta(x)$ because $q = x\theta(x)$. Then $\theta(u)^2 = x^2 w^m$, which is an LS equation and Theorem 1 implies $\rho(\theta(u)) = \rho(x) = w$. However since $x^2 p = qp \in \text{Suff}(\theta(u))$, we also get $\rho(p) = w$ (otherwise w would be a proper infix of its square in x^2). This leads to the usual contradiction that $v = px^2$ would not be primitive. \square

Proposition 47. *Under the setting of Problem 1, if $u_1u_2 = u\theta(u)$, $u_3 = u_4$, and m is even, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. As before we consider only non-trivial equation so that we have $u_3u_4 = qw_1 \cdots w_m$ and $|v| \neq |w|$. Lemma 38 gives two cases, but actually it suffices to consider the case when $u = qr$ for some non-empty θ -palindrome r .

First we consider the case when $u_3 = u$ and n is even. Then $[(pq)^{n/2-1}px]^2 = qw_1 \cdots w_m$, where $q = x\theta(x)$ for some $x \in \Sigma^+$. If $|p| \leq |q|$, then $pq = qp$ and v would not be even primitive. Hence let $p = qz_1$ for some $z_1 \in \Sigma^+$. Then $r = z_1x\theta(x)(pq)^{n/2-2}x\theta(x)z_1x$. Since $r = \theta(r)$, this equation gives $z_1x = \theta(z_1x)$ and $x = \theta(x)$. Thus we have $z_1x = x\theta(z_1)$ and $p = x^2z_1 = \theta(z_1)x^2$. Then $x^3z_1 = x\theta(z_1)x^2 = z_1x^3$ so that $\rho(x) = \rho(z_1)$ by Theorem 3. However, this result contradicts the primitivity of $v = pq = x^2z_1x^2$.

The second case is when $u_3 = u$ and n is odd. We have $[(pq)^{(n-1)/2}y]^2 = qw_1 \cdots w_m$, where $p = y\theta(y)$. From this equation, q is of even length so let $q = x\theta(x)$. If $|p| \leq |q|$, then we can apply Lemma 2 to the equation above to prove that $\rho(p) = \rho(q)$, which contradicts the primitivity of v . Thus we can let $y = xz_2$ for some $z_2 \in \Sigma^+$. Then $[(xz_2\theta(z_2)\theta(x)x\theta(x))^{(n-1)/2}xz_2]^2 = x\theta(x)w_1 \cdots w_m$. We can easily check that $w_{m/2+1} \cdots w_m = z_2[\theta(z_2)\theta(x)x\theta(x)xz_2]^{(n-1)/2}$. According to Proposition 37, we can deduce from this that $z_2, \theta(x)x \in \{w, \theta(w)\}^+$ and this further implies $x \in \{w, \theta(w)\}^+$. However then $v = pq = xz_2\theta(z_2)\theta(x)x\theta(x)$ would not be θ -primitive.

Thirdly we consider the case when $u_3 = \theta(u)$ and n is even. We have $[\theta(x)p(qp)^{n/2-1}]^2 = x\theta(x)w_1 \cdots w_m$, and this equation immediately gives $x = \theta(x)$. Then $p(qp)^{n/2-1}xp(qp)^{n/2-1} = xw_1 \cdots w_m$. Since the left-hand side and x are θ -palindromes, we have either $x \in \{w, \theta(w)\}^+$ or $w_1 = \cdots = w_m = w$ by Lemma 33. In the former case, $\theta(u)^2 = x^2w_1 \cdots w_m \in \{w, \theta(w)\}^+$ and hence $\theta(u), u \in \{w, \theta(w)\}^+$ (Lemma 10). Then $v^n = u\theta(u)x\theta(x) \in \{w, \theta(w)\}^+$, and hence $v \in \{w, \theta(w)\}$ because of Lemma 10 and the θ -primitivity of v, w . However, this contradicts the assumption $|v| \neq |w|$. In the latter case, we have $\theta(u)^2 = x^2w^m$ and hence $\rho(\theta(u)) = \rho(x) = w$ (Theorem 1). However since $qp = x^2p \in \text{Suff}(\theta(u))$, we reach the contradictory result $\rho(p) = w$.

The final case is when $u_3 = \theta(u)$ and n is odd. Then $[\theta(y)(qp)^{(n-1)/2}]^2 = qw_1 \cdots w_m$, where $p = y\theta(y)$ for some $y \in \Sigma^+$. Let $\theta(y)q = qz_4$ for some z_4 with $|y| = |z_4|$. Then $r = z_4(y\theta(y)q)^{(n-1)/2}y\theta(y)$, which is a θ -palindrome so that $z_4 = y = \theta(y)$. Now we can transform $\theta(y)q = qz_4$ into $yq = qy$ and hence $\rho(y) = \rho(q)$ (Theorem 3). However, then $v = y\theta(y)q$ would not be θ -primitive. \square

Combining the results obtained in this section, we can give a positive answer to Problem 1. Furthermore, with the result proved in [3] (also see Table 1), this positive answer concludes the following theorem, the strongest positive result we obtain on the ExLS equation.

Theorem 48. *Let $u, v, w \in \Sigma^+$ and let $u_1, \dots, u_\ell \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$. For $\ell \geq 4$ and $n, m \geq 3$, the equation $u_1 \cdots u_\ell = v_1 \cdots v_n w_1 \cdots w_m$ implies $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

4.5. The case $\ell \leq 3$ of the ExLS equation

We conclude this section with some examples which prove that an extended Lyndon-Schützenberger theorem cannot be stated for $\ell = 2$, and for some particular cases when $\ell = 3$.

Example 49. *Let $\Sigma = \{a, b\}$ and θ be an antimorphic involutions on Σ^* defined as $\theta(a) = a$ and $\theta(b) = b$. Let $v = a^{2m}b^2$ and $w = aa$ (i.e., $w = \theta(w)$) for some $m \geq 1$. Then $v^n w^m = (a^{2m}b^2)^n a^{2m}$. By letting either $u = (a^{2m}b^2)^{n/2} a^m$ if n is even or $u = (a^{2m}b^2)^{(n-1)/2} a^{2m}b$ otherwise, we have $u\theta(u) = v^n w^m$. Nevertheless, there cannot exist a word t such that $u, v, w \in \{t, \theta(t)\}^+$ because v contains b , while w does not. In conclusion, for arbitrary $n, m \geq 2$, $(2, n, m)$ does not impose θ -periodicity.*

Next we examine briefly the $(3, n, m)$ ExLS equation. The actual problem which we address is formalized as follows:

Problem 2. *Let $u, v, w \in \Sigma^+$ and integers $n, m \geq 3$. Then, let $u_1, u_2, u_3 \in \{u, \theta(u)\}$, $v_1, \dots, v_n \in \{v, \theta(v)\}$, and $w_1, \dots, w_m \in \{w, \theta(w)\}$. Does the equation $u_1 u_2 u_3 = v_1 \cdots v_n w_1 \cdots w_m$ imply $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$ under all of the following conditions?*

1. v and w are θ -primitive,
2. $|v_1 \cdots v_n| \geq |w_1 \cdots w_m|$,
3. $u_1 = u$, $v_1 = v$, and $w_m = w$.

As shown from now by examples, the general answer is “No”. More significant is the fact that depending on the values of variables u_2, u_3 and on the lengths of $v_1 \cdots v_n$ and $w_1 \cdots w_m$, the $(3, n, m)$ ExLS equation exhibits very complicated behavior.

First we present a parameterized example to show that for arbitrary $m \geq 2$, $(3, 3, m)$ does not impose θ -periodicity.

Example 50. *Let $\Sigma = \{a, b\}$ and θ be the mirror image over Σ^* . For $u = (abb)^{2m-1}ab$, $v = (abb)^{m-1}ab$, and $w = (bba)^3$, we have $u^2\theta(u) = v\theta(v)^2w^m$ for any $m \geq 2$. Nevertheless, there does not exist a word $t \in \Sigma^+$ satisfying $u, v, w \in \{t, \theta(t)\}^+$.*

In this example, the border between $v\theta(v)^2$ and w^m is located at u_2 . Intriguingly, as long as $u_1 u_2 u_3 = uu\theta(u)$ we cannot shift the border to u_3 without imposing $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.

Proposition 51. *For any $n, m \geq 3$, if $uu\theta(u) = v_1 \cdots v_n w_1 \cdots w_m$ and $n|v| > 2|u|$, then $u, v, w \in \{t, \theta(t)\}^+$ for some $t \in \Sigma^+$.*

Proof. It suffices to consider the case when $(n-1)|v| < 2|u| < n|v|$, otherwise Theorem 6 applies. As done in the analyses on the ExLS equation with $\ell = 4$, we can assume that both v and w are θ -primitive. Then, using Proposition 31, we obtain that n is even, $u = [r(tr)^i r(tr)^i (tr)^j]^{n/2-1} r(tr)^i (rt)^j$ and $v = r(tr)^i r(tr)^i (tr)^j$ for some $i \geq 0, j \geq 1$, and two non-empty θ -palindromes r, t such that rt is primitive. Moreover, $\theta(u) = (tr)^j r(tr)^i [(rt)^j r(tr)^i r(tr)^i]^{n/2-1} = r(tr)^i r(tr)^i w_1 \cdots w_m$. Hence if $i \geq 1$, then $tr = rt$, which contradicts the primitivity of rt (Theorem 3). Thus we have

$$(tr)^j r [(rt)^j r^2]^{n/2-1} = r^2 w_1 \cdots w_m. \quad (3)$$

If $|t| \leq |r|$, then $t \in \text{Pref}(r)$ from which $rt \in \text{Pref}(r^2 w_1 \cdots w_m)$, and finally $rt = tr$, contradicting the primitivity of rt again. If $|r| < |t| \leq 2|r|$, then we can write $rrs = tr$ for some $s \in \Sigma^+$ such that $|r| + |s| = |t|$. Since $s \in \text{Suff}(r)$ and r is a θ -palindrome, $\theta(s) \in \text{Pref}(r)$, i.e., $r = \theta(s)r_1$ for some $r_1 \in \Sigma^+$. Then, $rrs = r\theta(s)r_1s = tr$, so $r\theta(s) = t$ because their length is the same. Since $\theta(s) \in \text{Suff}(t)$ and t is a θ -palindrome, it holds that $s \in \text{Pref}(t)$ and $rrs \in \text{Pref}(rrt)$. Therefore, rrt and tr share a prefix of length $|t| + |r|$ so that Theorem 3 concludes that $\rho(r) = \rho(t)$, contradicting the primitivity of rt .

Thus both $i = 0$ and $|t| > 2|r|$ must hold. Eq. (3) implies that $r^2 \in \text{Pref}(t)$, that is, $r^2 \in \text{Suff}(t)$ (t is a θ -palindrome), and hence $r^4 \in \text{Suff}((rt)^j r^2)$. So we can let $r^4 = z_1 w_{k+1} \cdots w_m$ for some $k \geq 1$ and $z_1 \in \text{Suff}(w_k)$. If $z_1 = \lambda$, then this equation gives $r \in \{w, \theta(w)\}^+$ because w is assumed to be θ -primitive due to Theorem 5. Then Eq. (3) means $(tr)^j r [(rt)^j r^2]^{n/2-1} \in \{w, \theta(w)\}^+$. Using Proposition 16, we obtain $t \in \{w, \theta(w)\}^+$, but this contradicts the θ -primitivity of v . Otherwise, catenating r^2 from the left to the both sides of Eq. (3) gives us $r [(rt)^j r^2]^{n/2} = z_1 w_{k+1} \cdots w_m w_1 \cdots w_m$. Note that the left hand side of this equation is a θ -palindrome so that Lemma 25 implies $w_1 = \cdots = w_m = w$. Now catenating r in the same way to Eq. (3) gives $[(rt)^j r^2]^{n/2} = r^3 w^m$. This is in the form of LS equation and Theorem 1 implies $\rho((rt)^j r^2) = \rho(r) = w$ because w is primitive. From this we further deduce that $\rho(t) = w$. However, then rt would not be primitive. \square

Once we change $u_1 u_2 u_3$ from $u^2 \theta(u)$ to $u \theta(u)^2$, it becomes possible to construct a parameterized example for $(3, 3, m)$ with the border between $v_1 \cdots v_n$ and $w_1 \cdots w_m$ on u_3 , though it works only when m is a multiple of 3.

Example 52. Let $\Sigma = \{a, b\}$ and θ be the mirror image over Σ^* . For $i, j \geq 0$, let $u = (ab)^{i+j+1} (ba)^{2i+2j+2} b (ab)^j$, $v = (ab)^{i+j+1} (ba)^{i+2j+1} b$, and $w = ab$. Then $u \theta(u)^2 = v^3 w^{2(i+j+1)} \theta(w)^{i+j+1}$, but we cannot find such t that $u, v, w \in \{t, \theta(t)\}^+$.

Next we increase n to 4, and prove that still we can construct a parameterized example of the $(3, 4, 2i)$ ExLS equation.

Example 53. Let $\Sigma = \{a, b\}$ and θ be the mirror image over Σ . For $i \geq 1$, let $u = a^4 (ba^3)^i (a^3 b)^i$, $v = a^4 (ba^3)^{i-1} ba^2$, and $w = ba^3$. Then we have $u^3 =$

$v^2\theta(v)^2w^i\theta(w)^i$, but there does not exist a word $t \in \Sigma^+$ satisfying $u, v, w \in \{t, \theta(t)\}^+$.

The cases $(3, n, m)$ when $n = 4$ and m is odd, as well as when $m, n \geq 5$, remain open.

Table 2. Updated summary on the results regarding the extended Lyndon-Schützenberger equation

l	n	m	θ -periodicity	
≥ 4	≥ 3	≥ 3	YES	Theorem 48
3	≥ 5	≥ 5	?	
3	4	odd	?	
3	4	even	NO	Example 53
3	3	≥ 3	NO	Example 50
	one of them is 2		NO	Example 49

5. Conclusion

In this paper, we proved several consequences of the overlap between pseudo-primitive words. They made it possible to prove that, for a given antimorphic involution θ and words $u, v, w \in \Sigma^+$, if $l \geq 4$ and $n, m \geq 3$, then the ExLS equation $u_1 \cdots u_\ell = v_1 \cdots v_n w_1 \cdots w_m$ implies that $u, v, w \in \{t, \theta(t)\}^+$ for some t . This is the strongest result obtained so far on the ExLS equation. Our case analyses on $(3, \geq 3, \geq 3)$ ExLS equations demonstrated that these tools may not be sufficient to provide a complete characterization of ExLS equations. Further investigation on the overlaps of θ -primitive words, reduction schemes from ExLS equations to LS equations, and the weak defect effect seems promising and required to fill the gap in Table 2.

Acknowledgments

This research was supported by Natural Sciences and Engineering Research Council of Canada Discovery Grant R2824A01, and Canada Research Chair Award to L.K.

References

- [1] K. I. Appel and F. M. Djourup. On the equation $z_1^n z_2^n \cdots z_k^n = y^n$ in a free semigroup. *Transactions of the American Mathematical Society*, 134:461–470, 1968.
- [2] C. Choffrut and J. Karhumäki. Combinatorics of words. In *Handbook of Formal Languages*, volume 1, pages 329–438. Springer, 1997.
- [3] E. Czeizler, E. Czeizler, L. Kari, and S. Seki. An extension of the Lyndon-Schützenberger result to pseudoperiodic words. In *Proceedings of DLT 2009*, volume 5583 of *LNCS*, pages 183–194. Springer, 2009.

- [4] E. Czeizler, L. Kari, and S. Seki. On a special class of primitive words. In *Proceedings of MFCS 2008*, volume 5162 of *LNCS*, pages 265–277. Springer, 2008.
- [5] A. de Luca and A. De Luca. Pseudopalindrome closure operators in free monoids. *Theoretical Computer Science*, 362:282–300, 2006.
- [6] N. J. Fine and H. S. Wilf. Uniqueness theorem for periodic functions. *Proceedings of the American Mathematical Society*, 16(1):109–114, February 1965.
- [7] T. Harju and D. Nowotka. The equation $x^i = y^j z^k$ in a free semigroup. *Semigroup Forum*, 68:488–490, 2004.
- [8] T. Harju and D. Nowotka. On the equation $x^k = z_1^{k_1} z_2^{k_2} \dots z_n^{k_n}$ in a free semigroup. *Theoretical Computer Science*, 330(1):117–121, 2005.
- [9] L. Kari, R. Kitto, and G. Thierrin. Codes, involutions and DNA encoding. In *Formal and Natural Computing*, volume 2300 of *LNCS*, pages 376–393. Springer, 2002.
- [10] L. Kari and K. Mahalingam. Watson-Crick conjugate and commutative words. In *Proc. of DNA 13*, volume 4848 of *LNCS*, pages 273–283, 2008.
- [11] A. Lentin. Sur l'équation $a^m = b^n c^p d^q$ dans un monoïde libre. *Comptes Rendus de l'Académie des Sciences Paris*, 260:3242–3244, 1965.
- [12] M. Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1983.
- [13] R. Lyndon and M.-P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *Michigan Mathematical Journal*, 9:289–298, 1962.
- [14] J. Mañuch. *Defect Theorems and Infinite Words*. PhD thesis, University of Turku, Department of Mathematics, 2002.
- [15] G. Păun, G. Rozenberg, and T. Yokomori. Hairpin languages. *International Journal of Foundations of Computer Science*, 12(6):849–857, 2001.
- [16] H. J. Shyr. *Free Monoids and Languages*. Hon Min book company, Taichung, Taiwan, 3 edition, 2001.