Binary equality words with two b's

Štěpán Holub, Jiří Sýkora

Abstract. Deciding whether a given word is an equality word of two nonperiodic morphisms is also known as the dual Post correspondence problem. Although the problem is decidable, there is no practical decision algorithm. Already in the binary case, the classification is a large project dating back to 1980s. In this paper we give a full classification of binary equality words in which one of the letters has two occurrences.

Keywords: equality languages; dual Post correspondence problem; periodicity forcing

Classification: 68R15

1. Introduction

Equality sets of morphisms have been of interest for over seventy years. In 1946, E. L. Post published (see [19]) one of the most famous undecidable problems, which is now known as the Post correspondence problem (PCP). In algebraic terms, we ask whether there exists an equality word for two morphisms g and h. More specifically, we have two morphisms g and h from $\{a_1, a_2, \ldots, a_N\}^*$ to Σ^* and we ask whether there exists a word $w \in \{a_1, a_2, \ldots, a_N\}^+$ such that $g(w) = h(w)$. While PCP is undecidable, its binary version, i.e. when $N = 2$, was proved to be decidable, see [6] (complete proof in [10]), even in polynomial time, see [11]. This naturally led to interest in binary equality sets — the sets of all equality words for binary morphisms. They were first intensively studied in 1980 in [3], but the classification remains incomplete even today. The cases when one or both of the morphisms are periodic are relatively easy (see [14] and [7]). In case when both of the morphisms are nonperiodic, their equality set is generated by at most two words (see [13]). The equality sets with exactly two generators were described in [12].

Therefore, it remains to consider situations when two nonperiodic morphisms have an equality set generated by a single word. J. Hadravová and Š. Holub exam-

This work was supported by the Charles University, project GA UK No. 730414 and by the Czech Science Foundation grant number 13-01832S.

DOI 10.14712/1213-7243.2015.247

ined this situation thoroughly; their latest results were summarized in J. Hadravová's PhD thesis (see [8]). In [9], the authors proved that the equation $x^i y^j x^k = u^i v^j u^k$ if $j \geq 3$ and $i + k \geq 3$ has only periodic solutions. That implies that the word $a^i b^j a^k$ with $j \geq 3$ and $i + k \geq 3$ cannot be an equality word for nonperiodic morphisms. They found a nonperiodic solution for $j = 2$ and $i = k + 1$, but also conjectured that there are no nonperiodic solutions for $j = 2$ when $|i - k| \geq 2$. In this paper, we show that their conjecture holds. Using this key result, we are able to classify all binary equality words in which one of the letters occurs exactly twice.

The paper is organized as follows. After Preliminaries, in Section 3, we give a concise exposition of an important result about bi-infinite words, needed for our results. Section 4 contains more specific auxiliary lemmas. Our main partial classification theorem is stated and proved in Section 5. Using that theorem, we are able to complete the classification of all binary equality words with two b's in Section 6.

2. Preliminaries

We use standard notation of combinatorics on words. Throughout the paper, Σ will denote the binary alphabet $\{a, b\}$. Every nonempty word u has its (uniquely determined) *primitive root*, denoted by p_u, i.e. the shortest word v such that $u = v^i$ for some $i \in \mathbb{N}$. A word that is equal to its primitive root is called *primitive*. It is well known that two nonempty words u and v commute if and only if $p_u = p_v$. We denote by $|u|$ the *length of* u, i.e. its number of letters, and by $|u|_a$ the number of letters a contained in u. Words u, v are *conjugate* if there exist words w_1 and w_2 such that $u = w_1 w_2$ and $v = w_2 w_1$. We denote by $u \leq_p v$ or $u \leq_s v$ the fact that u is a prefix of v or a suffix of v, respectively. The maximal common prefix and suffix of u and v are denoted by $u \wedge v$ and $u \wedge_s v$, respectively. The symbol u^ω denotes the (one-way) infinite word obtained by an infinite concatenation of copies of u.

We say that a morphism g is *periodic* if there exists a word u such that $g(v) \in u^*$ for each v on which g is defined. Let g, h be two morphisms. A nonempty word v such that $g(v) = h(v)$ is called an *equality word of g and h*. We say that v is a *binary equality word* if there exist two distinct nonperiodic morphisms g and h defined on Σ^* such that v is an equality word of g and h. The set of all equality words of g and h is called their *equality set* and we denote it by $\mathrm{Eq}(g, h)$. A word $v \in A^*$ is *periodicity forcing* if the equality $g(v) = h(v)$ is satisfied only if both g and h are periodic or $g = h$. Note the asymmetry between definitions of periodicity forcing words and binary equality words which leaves aside words that would force just one morphism to be periodic. In the binary case, however, if $g(w) = h(w)$, and just one of the morphisms is periodic, then $w = a^i b a^j$ (see Lemma 7 or [12]), which also allows both morphisms to be nonperiodic. Therefore, any binary word is either periodicity forcing or binary equality word.

It is well known that if two words satisfy a nontrivial relation, then they commute. This fact actually holds also for the free group $F(\Sigma)$ as follows.

Lemma 1. *Let $x, y \subset F(\Sigma)$. If x and y are not free generators of the subgroup $G = \langle x, y \rangle$ of $F(\Sigma)$, then x and y commute.*

For the proof, cf. for example [2, Chapter III, Theorem 9].

Another important lemma is the following one. Its proof can be found in [20, Chapter 6, Theorem 6.1].

Lemma 2 (Periodicity lemma)**.** *Let u and v be primitive words. If the words u^ω and v^ω have a common factor of length at least $|u| + |v| - 1$, then u and v are conjugate.*

Remark. Note that if u and v from the Periodicity lemma are prefix or suffix comparable, then they are equal.

The following two lemmas describe well-known facts about primitive words (see e.g. [16, Chapter 12, Proposition 12.1.3]):

Lemma 3. *Let $w \in \Sigma^*$ be a word. If $w^i = uwv$ for some $i \in \mathbb{N}$ and some words $u, v \in \Sigma^*$, then $u = p_w^j$ and $v = p_w^k$ for some $j, k \in \mathbb{N}_0$ such that $p_w^{j+k} = w^{i-1}$.*

Lemma 4. *Let $w \in \Sigma^*$ be a primitive word. If $w^i = uwv$ for some $i \in \mathbb{N}$ and some words $u, v \in \Sigma^*$, then $u = w^j$ and $v = w^k$ for some $j, k \in \mathbb{N}_0$ such that $j + k = i - 1$.*

We also use two lemmas about the bound on the length of the maximal common prefix (or suffix) of two different words from a binary code (cf. [20, Chapter 6, Lemma 3.1]):

Lemma 5. *Let $X = \{x, y\} \subseteq \Sigma^*$ and let $\alpha \in xX^*$, $\beta \in yX^*$ be words such that $\alpha \wedge \beta \geq |x| + |y|$. Then x and y commute.*

Lemma 6. *Let $X = \{x, y\} \subseteq \Sigma^*$ and let $\alpha \in X^*x$, $\beta \in X^*y$ be words such that $\alpha \wedge_s \beta \geq |x| + |y|$. Then x and y commute.*

Let $X = \{x, y\}$, where $x, y \in \Sigma^*$ and x and y do not commute. We say that a word $u \in X^*$ is *X-primitive*, if $u = v^i$ implies $u = v$ for all $v \in X^*$. The following lemma can be found in [9] as Lemma 9.

Lemma 7. *Suppose that $x, y \in \Sigma^*$ do not commute and let $X = \{x, y\}$. If there is an X-primitive word $\alpha \in X^*$ and a word $z \in \Sigma^*$ such that $\alpha = z^i$ with $i \geq 2$, then $\alpha = x^k y x^l$ or $\alpha = y^k x y^l$ for some $k, l \geq 0$.*

It should be said that the previous result was first proved by J.-C. Spehner in [21], and consecutively by E. Barbin-Le Rest and M. Le Rest in [1]. Note that the famous result of Lyndon and Schützenberger (see [17]) is a consequence of the previous lemma.

And finally, we need to formulate two well-known facts about conjugate words.

Lemma 8. *Let u and v be conjugate words. Then also p_u and p_v are conjugate. In particular, if u is primitive, then v is primitive as well.*

Lemma 9. *Let $x \neq \varepsilon \neq y$ and z be words satisfying $zx = yz$. Then there exist words s, t such that $s \neq \varepsilon$, $z = (ts)^j t$ for some $j \geq 0$, st is the primitive root of x and ts is the primitive root of y.*

Note that the previous lemma implies that the words x and y are conjugate. We say that they are *conjugate by z*. In this situation, we may assume that z is shorter than x and y or, more precisely, there exists z' shorter than x such that $z'x = yz'$.

3. Bi-infinite words

In the following section, we shall deal with bi-infinite words. The purpose is to use Theorem 3.11 from [18] to prove a useful lemma about conjugate words. This lemma plays a crucial role in the proof of our main theorem. The results of this section, Theorem 10 and Lemma 11, are nice general statements from combinatorics on words, which are probably not as well known as they would deserve. Therefore, we include them with a short introduction, where we establish proper notation and address some of the intricacies of bi-infinite words.

While the concept of bi-infinite words may seem natural, there arise certain problems and ambiguities when we try to formalize it. Intuitively, a bi-infinite word w over the alphabet Σ is an infinite sequence (in both directions) of letters from Σ. We usually write $w = \dots w_{-1} w_0 w_1 \dots$, where $w_i \in \Sigma$. Note that this bi-infinite sequence represents a mapping $w \colon \mathbb{Z} \to \Sigma$. The words w and w' defined as $w'(i) = w(i + k)$ for some $k \in \mathbb{Z}$ are formally different, although they are isomorphic as ordered sequences (of type $\mathbb{Z}$). J. Maňuch calls w and w' two representations of the same word. We respect the fact that they are formally different and call them *equivalent* instead, and write $w \sim w'$.

Similar problems occur when one tries to define a factorization of a bi-infinite word. A *factorization* F is an order preserving (injective) mapping $F \colon \mathbb{Z} \to \mathbb{Z}$ satisfying $F(0) \leq 0$ and $F(1) \geq 1$. Note that a factorization is defined by its range, in particular, unlike for words, we do not need any notion of equivalent factorizations. Therefore, we can also see factorizations as subsets of $\mathbb{Z}$ that have no lower nor upper bounds. A *factorization applied to a bi-infinite word w* is a mapping $F^w \colon \mathbb{Z} \to \Sigma^+$ where the words $F^w(i)$ are defined as follows:

$$\begin{aligned} &\vdots \\ F^w(-1) &= w_{F(-1)} w_{F(-1)+1} \cdots w_{F(0)-1}; \\ F^w(0) &= w_{F(0)} w_{F(0)+1} \cdots w_{F(1)-1}; \\ F^w(1) &= w_{F(1)} w_{F(1)+1} \cdots w_{F(2)-1}; \\ &\vdots \end{aligned}$$

The set of all factors of w as factorized by F, i.e. the range of F^w is denoted by $F^w(\mathbb{Z})$. Let $X = \{\alpha, \beta\}$ be a binary set where $\alpha, \beta \in \Sigma^+$ (this implies $\alpha \neq \beta$). A factorization F of a bi-infinite word w over Σ is an *X-factorization* if $F^w(\mathbb{Z}) \subseteq X$.

Conversely, when we have a bi-infinite sequence of words $S\colon \mathbb{Z} \to \Sigma^+$ we can define its concatenation $\prod S$ as a bi-infinite word w such that

$$\begin{aligned}
&\vdots\\
w(-1) &= S(-1)_{|S(-1)|};\\
w(0) &= S(0)_1;\\
w(1) &= S(0)_2;\\
&\vdots\\
w(|S(0)|-1) &= S(0)_{|S(0)|};\\
w(|S(0)|) &= S(1)_1;\\
&\vdots
\end{aligned}$$

Note that for a bi-infinite word w, the word $w' = \prod F^w$ may be different from w. However, these two words are equivalent. Let $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})$ be a finite sequence of words. We denote by $(\alpha_0, \alpha_1, \ldots, \alpha_{n-1})^{\mathbb{Z}}$ the bi-infinite sequence of words S such that $S(i) = \alpha_{(i \mod n)}$ for every $i \in \mathbb{Z}$. For a single word, we use simply $\alpha^{\mathbb{Z}}$ instead of $(\alpha)^{\mathbb{Z}}$. We also define the equivalence of bi-infinite sequences of words analogously to the equivalence of be-infinite words.

We are now prepared to reformulate Theorem 3.11 from [18] (see also [15, Theorem 2]).

Theorem 10. *Consider a binary set $X = \{\alpha, \beta\}$ with $\alpha, \beta \in \Sigma^+$. Let w be a bi-infinite word over Σ and let F_1 and F_2 be two different X-factorizations such that $F_1^w(\mathbb{Z}) \cup F_2^w(\mathbb{Z}) = X$. Then at least one of the following conditions is satisfied:*

(i) *α and β commute; or*
(ii) *the primitive roots of α and β are conjugate, $w \sim \prod \alpha^{\mathbb{Z}} \sim \prod \beta^{\mathbb{Z}}$, and $F_1^w = \alpha^{\mathbb{Z}}$ and $F_2^w = \beta^{\mathbb{Z}}$, or vice versa; or*
(iii) *there exists an imprimitive word $y = y_0 \ldots y_{n-1} \in X^+$ such that $w \sim \prod y^{\mathbb{Z}}$ and $F_1^w \sim F_2^w \sim (y_0, \ldots, y_{n-1})^{\mathbb{Z}}$.*

Remark. J. Maňuch calls a word w for which there exist factorizations F_1, F_2 as in Theorem 10 *proper X-ambiguous.*

The following lemma is a direct consequence of the previous theorem.

Lemma 11. *Let $A = \{a, b\}$, and $u, v \in A^*$ such that at least one of u, v contains both a and b. Let $g\colon A^* \to \Sigma^*$ be a morphism such that $g(a) = s$, $g(b) = t$ and the words $g(u)$ and $g(v)$ are conjugate. Then u and v are conjugate or s and t commute.*

Proof: In the following proof, we shall assume that u and v are not conjugate and we shall show that then s and t commute. If $s=t$ or one of them is empty, they commute trivially. Thus, suppose $\varepsilon \neq s \neq t \neq \varepsilon$. Note that even the primitive roots of u and v are not conjugate; that would mean either that u and v are conjugate or that $g(u)$ and $g(v)$ have different lengths. Since $g(u)$ and $g(v)$ are conjugate, we get $\prod(g(u))^{\mathbb{Z}} \sim \prod(g(v))^{\mathbb{Z}}$. The bi-infinite word $w=\prod(g(u))^{\mathbb{Z}}$ has two natural $\{s,t\}$-factorizations F_1, F_2 such that $F_1^w \sim (g(u_1),\dots,g(u_{|u|}))^{\mathbb{Z}}$ and $F_2^w \sim (g(v_1),\dots,g(v_{|v|}))^{\mathbb{Z}}$. Since p_u and p_v are not conjugate, $\prod u^{\mathbb{Z}} \nsim \prod v^{\mathbb{Z}}$ by the Periodicity lemma. That implies that factorizations F_1 and F_2 are different and $F_1^w \nsim F_2^w$. We also have $F_1^w(\mathbb{Z}) \cup F_2^w(\mathbb{Z}) = \{s,t\}$, because at least one of the words u and v contains both a and b. Hence, we can use Theorem 10. There, the only situation that can happen is (i). The second case cannot occur because $F_1^w(\mathbb{Z}) = \{s,t\}$ or $F_2^w(\mathbb{Z}) = \{s,t\}$. And the last one is precluded by the fact that $F_1^w \nsim F_2^w$. □

4. Binary equality words with two b's — part 1

Let $g,h\colon \{a,b\}^* \to \Sigma^*$ be two distinct nonperiodic morphisms. Consider the equality set of g and h. As we mentioned in the introduction, $\mathrm{Eq}(g,h)$ is generated by at most two words, and the case with exactly two generators has already been solved:

Theorem 12 ([12]). *Let g and h be nonperiodic binary morphisms, and let $Eq(g,h)$ be generated by two words. Then $Eq(g,h) = \{a^ib, ba^i\}^+$ for some $i \geq 1$ (up to exchange of letters).*

The following example comes from [3, Example 7.1].

Example 13. Let

$$g\colon \quad a \mapsto a \qquad\qquad h\colon \quad a \mapsto a^iba^i$$
$$\quad b \mapsto (ba^{2i})^{i-1}ba^i(ba^{2i})^{i-1}b \qquad\qquad b \mapsto (ba^{2i})^{i-1}b.$$

Then $\mathrm{Eq}(g,h) = \{a^ib, ba^i\}^+$.

Let $w \in \mathrm{Eq}(g,h)$. We focus on the case when $|w|_b = 2$. In [3], it was pointed out that all binary words of length at most four are binary equality words. Also, all binary equality words w with $|w|_b = 2$ and $|w|_a = 3$ were classified as follows:

$$w \in \{a^3b^2, b^2a^3, a^2b^2a, ab^2a^2, ababa, ba^3b\}.$$

Note that the above claim also considers words that do not generate the equality set. For example, $abab$ is a binary equality word but it is not a generator, since it is a square of ab and two morphisms agree on $abab$ if and only if they agree on ab.

The situation becomes unclear for more a's. It has been known (see [8]) that the generator of $\mathrm{Eq}(g,h)$ for nonperiodic binary morphisms can be one of the following words:

$$a^ib^2,\quad b^2a^i,\quad ba^{2i+1}b,\quad a^{i+1}b^2a^i,\quad a^ib^2a^{i+1}.$$

The open question was whether there are some other generators with two b's. We show that words of the form a^ibaba^i are also binary equality words. On the other hand, we prove that there are no other binary equality words with two b's that can generate $\mathrm{Eq}(g,h)$ for nonperiodic morphisms.

We need three useful lemmas. They are closely connected to Lemmas 5.3 and 5.4, and Theorem 6.2 in [3]. In particular, it can be observed that they show that the words (missing in the above list of binary equality words of length five) $abaab$, $baaba$ (Lemma 14) and $aabab$, $babaa$ (Lemma 15) are periodicity forcing.

Lemma 14. *Let x, y, u and v be nonempty words such that $yx^{l+1}yx = vu^{l+1}vu$ or $xyx^{l+1}y = uvu^{l+1}v$, with $l \geq 1$, $x \neq u$. Then all u, v, x and y commute.*

PROOF: Consider the equality $yx^{l+1}yx = vu^{l+1}vu$. The other one is symmetric.

Let $\bar{y} = yx$ and $\bar{v} = vu$. Then $\bar{y}x^l\bar{y} = \bar{v}u^l\bar{v}$. We have $|\bar{y}| \neq |\bar{v}|$ since $x \neq u$. By symmetry, suppose $|\bar{y}| < |\bar{v}|$.

Let $x^l = x_1u^lx_2$, where $\bar{y}x_1 = x_2\bar{y} = \bar{v}$. Then, by Lemma 9, there are words s and t, and integers $i > 0$ and $j \geq 0$ such that st is primitive, $s \neq \varepsilon$, $x_1 = (st)^i$, $x_2 = (ts)^i$ and $\bar{y} = (ts)^jt$.

1. If $|x| \geq |st|$, then ts is a suffix of x, and therefore also of $\bar{y}$. Hence $st = ts$, which implies $t = \varepsilon$ and $x^l = s^iu^ls^i$. If $l > 1$, then s commutes with u by Lemma 7. If $l = 1$, then $s^ius^i = x \leq_s \bar{y} = s^j$, and again s commutes with u by Lemma 4. Then u, x, v and y commute.

2. If $|x| < |st|$, then x is a suffix of st, since x is a suffix of $\bar{y}$. Therefore x is a suffix of x_1 and since x_1 is a prefix of x^l, Lemma 3 implies that x_1 and x commute, a contradiction with $x \neq \varepsilon$ and $|x| < |st|$. □

Lemma 15. *Let x, y, u and v be nonempty words such that $x^{l+1}yxy = u^{l+1}vuv$, where $l \geq 1$ and $x \neq u$. Then all u, v, x and y commute.*

PROOF: Let $\bar{y} = xy$ and $\bar{v} = uv$. Then $x^l\bar{y}\bar{y} = u^l\bar{v}\bar{v}$. Suppose by symmetry that $|\bar{y}| < |\bar{v}|$, hence $|x| > |u|$. If x and u commute, then the claim holds by Lemma 7. Suppose that they do not commute. Since u^{l+1} is a prefix of x^{l+1}, the Periodicity lemma implies that $|u^l| < |x|$. Then $zx^lyxy = \bar{v}\bar{v}$, where z is a suffix of x.

1. If $|\bar{y}\bar{y}| \leq |\bar{v}|$, then xyx is a factor of $\bar{v}$ which is a factor of x^+. Therefore, by Lemma 3, x and $\bar{y}$ commute. Then all words commute by Lemma 7.

2. Let now $|\bar{y}\bar{y}| > |\bar{v}|$ and let $\bar{v} = zx^{l-1}y_1 = y_2\bar{y}$, where $\bar{y} = y_1y_2 = y_3y_1$. Let s and t be words, and $i > 0$ and $j \geq 0$ integers such that st is primitive, $s \neq \varepsilon$, $y_2 = (st)^i$, $y_3 = (ts)^i$ and $y_1 = (ts)^jt$. From $zx^{l-1}y_1 = y_2\bar{y}$, we obtain $zx^l = y_2y_3x$. Moreover, x is a prefix of y_3x since $xyy_2 = y_3xy$. By applying Lemma 3 to the word x^{l+1}, we deduce that ts is the primitive root of x. However, y_2y_3, and hence also $stts$, is a factor of x^{l+1}, which implies that s and t commute, and we are done. □

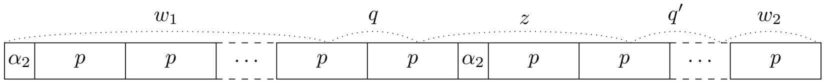

FIGURE 1. Situation in Lemma 17 when $|w_1| < \left|\alpha_2 p^{l-1}\right|$ and $|w_2| < \left|\alpha_2 p^{l-1}\right|$.

Lemma 16. *Let x, y, u and v be nonempty words such that $x^l yxyx = u^l vuvu$, where $l \geq 2$ and $x \neq u$. Then all u, v, x and y commute.*

PROOF: Let $\bar{y} = yx$ and $\bar{v} = vu$. Then $x^l \bar{y}\bar{y} = u^l \bar{v}\bar{v}$. By symmetry, suppose $|\bar{y}| < |\bar{v}|$. Hence there exists w such that $\bar{v} = w\bar{y}$. This leads to $x^l \bar{y} = u^l w\bar{y}w$. Let z be the suffix of x^l satisfying $\bar{y}w = z\bar{y}$. Then, by Lemma 9, there are words s and t and integers $i > 0$ and $j \geq 0$ such that st is primitive, $s \neq \varepsilon$, $w = (st)^i$, $z = (ts)^i$ and $y = (ts)^j t$.

1. If $|x| \geq |st|$, then ts is a suffix of x, and therefore also of $\bar{y}$. Hence $st = ts$, which implies $t = \varepsilon$ and $x^l = u^l s^{2i}$. Since $l > 1$, Lemma 7 implies that s commutes with u and x. Therefore all four words u, x, v and y commute.

2. If $|x| < |st|$, then x is a suffix of st, since x is a suffix of $\bar{y}$. Therefore x is a suffix of w and since $x^l = u^l wz$, Lemma 3 implies that z and x commute, a contradiction with $x \neq \varepsilon$ and $|x| < |st|$. □

5. The equation $x^i y^2 x^k = u^i v^2 u^k$

In this section we formulate our crucial theorem, which states that the word $a^i bba^k$ is periodicity forcing for $|i - k| \geq 2$. Firstly however, we need a few more lemmas. They seem rather technical but have a clear intuitive meaning: they directly exploit the "synchronization property" of primitive words formulated by Lemma 4. The first lemma is the most general one, while the next three are closely connected to it but more specific.

Lemma 17. *Let $p = \alpha_1\alpha_2 = \beta_1\beta_2 = \gamma_1\gamma_2$, where $\alpha_2, \beta_2, \gamma_2 \neq \varepsilon$, be a primitive word, $l \geq 1$, and let $(\alpha_2 p^l)^2 = w_1 qzq' w_2$, where $q = \beta_2\beta_1$, $q' = \gamma_2\gamma_1$, $z \neq \varepsilon$ and $|z| \equiv |\alpha_2| \pmod{|p|}$. Then the following holds:*

- (A) *if $|w_1| < \left|\alpha_2 p^{l-1}\right|$ and $|w_2| < \left|\alpha_2 p^{l-1}\right|$, then $z \in q^* \beta_2 \alpha_1^{-1} \beta_1 q^*$ and $q = q'$;*
- (B) *if $|w_2| \geq \left|\alpha_2 p^{l-1}\right|$, then $z \in q^* w$ where $|w| = |\alpha_2|$ and w is a prefix of q;*
- (C) *if $|w_1| \geq \left|\alpha_2 p^{l-1}\right|$, then $z \in w'(q')^*$ where $|w'| = |\alpha_2|$ and w' is a suffix of q'.*

PROOF: In the first case, q and q' are factors of $\alpha_2 p^l$ and z "lies over the edge" $p\alpha_2$. The fact that $q = q'$ follows immediately from the length of z. Since $\alpha_2 p^l = \alpha_1^{-1}\beta_1 q^l \beta_2$, Lemma 4 implies that z has the required form. If $|w_2| \geq \left|\alpha_2 p^{l-1}\right|$, then qz is a factor of $\alpha_2 p^l$, which is a factor of q^ω. Hence z must have the form from point (B) by Lemma 4. The last case is symmetric. □

The following lemma describes the same situation when the whole word qzq' is a factor of $\alpha_2 p^l$.

Lemma 18. *Let $p = \alpha_1\alpha_2 = \beta_1\beta_2 = \gamma_1\gamma_2$, $\alpha_2, \beta_2, \gamma_2 \neq \varepsilon$, be a primitive word, $l \geq 1$, and let $(\alpha_2 p^l)^2 = uqzq'v$, where $q = \beta_2\beta_1, q' = \gamma_2\gamma_1, z \neq \varepsilon$ and $|z| \equiv |\alpha_2| \pmod{|p|}$. If either $|w_1| \geq |\alpha_2 p^l|$ or $|w_2| \geq |\alpha_2 p^l|$, then $z \in q^*\beta_2\gamma_2^{-1}$.*

Proof: In this situation, the whole qzq' is a factor of $\alpha_2 p^l$. Therefore, the conclusions of both (B) and (C) from Lemma 17 hold as we may consider qzq' to be inside the first or second $\alpha_2 p^l$. Using the notation from (B), we get $w = \beta_2\gamma_2^{-1}$ if $|\beta_2| > |\gamma_2|$ and $w = q\beta_2\gamma_2^{-1}$ otherwise. In either case, $z \in q^*\beta_2\gamma_2^{-1}$. □

Now we describe the situation when $q = q' = \alpha_2\alpha_1$.

Lemma 19. *Let $p = \alpha_1\alpha_2$, $\alpha_2 \neq \varepsilon$, be a primitive word and let $q = \alpha_2\alpha_1$. If qzq is a factor of $(\alpha_2 p^l)^2$ for $l \geq 1$ and for some nonempty word z such that $|z| \equiv |\alpha_2| \pmod{|p|}$, then $z \in q^*\alpha_2 q^*$ or $z \in \alpha_2' q^*$, where $\alpha_1\alpha_2' = q$.*

Proof: We can use Lemma 17 for $\alpha_1 = \beta_1 = \gamma_1$ and $\alpha_2 = \beta_2 = \gamma_2$. Case (A) leads to $z \in q^*\alpha_2 q^*$. Case (B) implies $z \in q^*\alpha_2$. Thus, the problematic case is (C), i.e. $z \in w'q^*$. If $|w_1| \geq |\alpha_2 p^l|$, then Lemma 4 implies that $\alpha_2 = p$ and Lemma 18 suggests that $z \in q^*$. If $|\alpha_2 p^{l-1}| \leq |w_1| < |\alpha_2 p^l|$, then $w' = \alpha_2'$, where $\alpha_1\alpha_2' = \alpha_2\alpha_1$, which is what we wanted to prove. This follows from the facts that the occurrence of z from qzq is within q^ω, this z is preceded by q and $w_1 q = \alpha_2 p^l w_1'$, where $|w_1'| = |\alpha_1|$. □

The final lemma deals with the case $q = q' = p$.

Lemma 20. *Let $p = \alpha_1\alpha_2$, $\alpha_2 \neq \varepsilon$, be a primitive word. If pzp is a factor of $(\alpha_2 p^l)^2$ for $l \geq 1$ and for some nonempty word z such that $|z| \equiv |\alpha_2| \pmod{|p|}$, then $z \in p^*\alpha_2 p^*$ or $z \in p^*\alpha_2'$, where $\alpha_2'\alpha_1 = \alpha_1\alpha_2 = p$.*

Proof: The proof is analogous and symmetric to the proof of Lemma 19. □

Finally, we present four lemmas about commutation of words under certain conditions.

Lemma 21. *Let x, α_1, α_2 and u be words such that $x = \alpha_1\alpha_2$, u is a prefix of x and $x = \alpha_2\alpha_2 u^i$ for some $i > 0$. Then x and u commute.*

Proof: Assume that u and x are nonempty (otherwise the claim holds). The equality $\alpha_1\alpha_2 = \alpha_2\alpha_2 u^i$ implies that $\alpha_1 = \alpha_2 w$ for some w of length $i|u|$. Then $\alpha_2 u^i = w\alpha_2$. Lemma 9 implies that $w = (ts)^l$, $u^i = (st)^l$ and $\alpha_2 = (ts)^j t$ for some $l \geq 1$, $j \geq 0$, and some s, t such that $s \neq \varepsilon$ and ts is the primitive root of u. Then either tts or ts is a prefix of α_1 depending on whether $j = 0$ or not. Since u is a prefix of $x = \alpha_1\alpha_2$, we have either $st \leq_p tts$ or $st = ts$. In either case, s and t commute which means that t is empty, and s is the primitive root of both u and x. □

Lemma 22. *Let x, α_1, α_2 and u be words such that $x = \alpha_1\alpha_2$, u^i is a prefix of x, u is a suffix of x and $\alpha_1\alpha_2 u^{-(i+1)}\alpha_1 = \alpha_2\alpha_1\alpha_2$ for some $i > 1$. Then x and u commute.*

Proof: Assume that u and x are nonempty (otherwise the claim holds). Since $|\alpha_1| = \left|u^{i+1}\alpha_2\right|$, $x = \alpha_1\alpha_2$ and u^i is a prefix of x, we have $\alpha_1 = u^i\tilde{u}\tilde{\alpha}_2$, where $|\tilde{u}| = |u|$ and $|\tilde{\alpha}_2| = |\alpha_2|$. Then $u^i\tilde{u}\tilde{\alpha}_2\alpha_2 u^{-1}\tilde{u}\tilde{\alpha}_2 = \alpha_2 u^i\tilde{u}\tilde{\alpha}_2\alpha_2$. We get that $\tilde{\alpha}_2 = \alpha_2$ and $u^i\tilde{u}\alpha_2\alpha_2 u^{-1}\tilde{u} = \alpha_2 u^i\tilde{u}\alpha_2$.

a) If $|\alpha_2| \geq |u|$, then u is a suffix of α_2 which implies $\tilde{u} = u$ and $u^{i+1}\alpha_2 = \alpha_2 u^{i+1}$. Since also $\alpha_1 = u^{i+1}\alpha_2$, we deduce that x and u commute.

b) Let $|\alpha_2| < |u|$. Since $\alpha_2 u$ is a prefix of u^i, Lemma 5 implies that α_2 and u commute. Then $u^i\tilde{u}\alpha_2\alpha_2 u^{-1}\tilde{u} = \alpha_2 u^i\tilde{u}\alpha_2$ is a nontrivial relation between the primitive roots of u and $\tilde{u}$, which again implies (by Lemma 1) that $u = \tilde{u}$ and we can continue as in a). □

Lemma 23. *Let x, α_1, α_2 and u be words such that $x = \alpha_1\alpha_2$, u^i is a prefix of x, u is a suffix of x and $xu^{-(i+1)}x\alpha_1 = \alpha_2$ for some $i > 1$. Then x and u commute.*

Proof: Assume that u and x are nonempty (otherwise the claim holds). We have $2|\alpha_1| = \left|u^{i+1}\right| - |x| \leq |u|$. Since α_1 is a prefix and also a suffix of x, it is a border of u and we deduce that $x = u^i\alpha_1^{-2}u$. Note that $\alpha_2 = \alpha_1^{-1}x$. Therefore $u^i\alpha_1^{-2}\alpha_1^{-2}u\alpha_1 = \alpha_1^{-1}u^i\alpha_1^{-2}u$. This is a nontrivial relation, therefore α_1 and u commute by Lemma 1. Consequently, u and x commute as well. □

Lemma 24. *Let x, α_1 and u be words such that u^i is a prefix of x, u is a suffix of x and $u^{i+1} = \alpha_1\alpha_1 x$ for some $i > 1$. Then x and u commute.*

Proof: Assume that u and x are nonempty (otherwise the claim holds). Since $x = ux'u$ for some x', x and u commute by the Periodicity lemma, because the words u^ω and $(ux')^\omega$ have a common factor of length $|u| + |ux'|$. □

The following theorem is one of our main results. It shows that certain words with two b's are periodicity forcing.

Theorem 25. *Let x, y, u and v be words such that they satisfy $x^i y^2 x^k = u^i v^2 u^k$ for some $i, k \in \mathbb{N}$ and $x \neq u$. If $|i - k| \geq 2$, then all the words x, y, u and v commute.*

It is important to note that this theorem, as well as other above results about symmetric equations, are straightforwardly related to periodicity forcing words. Indeed, Theorem 25 implies that the word $a^i bba^k$ with $|i - k| \geq 2$ is periodicity forcing; it is enough to set $g(a) = x$, $g(b) = y$, $h(a) = u$ and $h(b) = v$.

Remark. Note that the condition $i, k \geq 1$ is necessary. For example, the equation $x^2y^2 = u^2v^2$ has a solution $x = aab, y = a, u = a$ and $v = baa$. The condition $|i - k| \geq 2$ is necessary as well. If $i = k + 1$, we have the following solution (see [9]):

$$x = a^{2k+1}(ba^k)^2, \qquad u = a,$$
$$y = ba^k, \qquad v = (a^k b)^2 (a^{3k+1} ba^k b)^k.$$

Proof: If one of the words x, y, u and v is empty, the theorem holds by Lemma 7. For example, if $v = \varepsilon$, we get $x^i y^2 x^k = u^{i+k}$. Since, $i + k \geq 2$, x and y commute by Lemma 7. Then also u commutes with x and y and all these words commute with the empty word v. Thus we may assume that x, y, u and v are nonempty. Without loss of generality, we also assume that $|x| \geq |u|$ and $i > k+1$. It is enough to prove that x and u commute. Then, $p_x = p_u$ and we obtain $p_x^{in} y^2 p_x^{kn} = v^2$ for some $n \geq 1$. Hence we are done by Lemma 7. If $(i + k - 1)|u| \geq |p_x|$, p_x^ω and u^ω have a common factor of length at least $|p_x| + |u|$, x and u commute by the Periodicity lemma. We therefore suppose

$$(i + k - 1)|u| < |p_x|,$$

which implies

$$\frac{i-k}{2}|u| < \frac{|p_x|}{2}.$$

The equality is equivalent to $yx^k u^{-(k+i)} x^i y = v'v'$, where v' is a conjugate of v. Let α be the prefix of x^i of length

$$\frac{i-k}{2}|x| + \frac{i+k}{2}|u|.$$

Then

$$yx^k u^{-(k+i)} \alpha = \alpha^{-1} x^i y,$$

that is, $x^k u^{-(k+i)} \alpha$ is conjugate (by y) with $\alpha^{-1} x^i$. We can assume, without loss of generality, that y is shorter than the two words. Let $p_x = \alpha_1 \alpha_2$, where $|\alpha_1| < |p_x|$ and $\alpha = p_x^c \alpha_1$ for some $c \geq 0$. Let y' be such that

$$x^k u^{-(k+i)} \alpha = y'y, \qquad \alpha^{-1} x^i = yy'.$$

Note that $yy' = \alpha_2 p_x^l$ for some $l \geq 0$. Also note that the words yy' and $y'y$ have the same length, i.e. $\left|\alpha_2 p_x^l\right| = |p_x^{c_1}| + |\alpha_1| - (i + k)|u|$ for some c_1. From that, we obtain

$$|\alpha_1| - (i + k)|u| \equiv |\alpha_2| \pmod{|p_x|}.$$

Finally, note that if α_1 and α_2 commute, then α_1 commutes with p_x, i.e. it is an empty word. Then both α and $\alpha^{-1} x^i$ are powers of p_x, which implies that $x^k u^{-(k+i)} \alpha$ is a power of a conjugate of p_x. Bearing in mind that α is a power of p_x, we deduce that $u^{k+i} = p_x$, i.e. x and u commute. The fact that x and u commute if α_1 and α_2 commute will be often used in the proof.

1. *Case $i-k \geq 3$ and $k \geq 2$.* In this case, since $y'y$ and yy' are conjugate, we get that $xxu^{-(k+i)}\alpha$ is a factor of $(yy')^2 = (\alpha_2 p_x^l)^2$. We also have $\left|u^{-i}\alpha\right| > |p_x|$ because

$$\left|u^{-i}\alpha\right| = \frac{i-k}{2}(|x| - |u|).$$

Let us put $q' = \alpha_2\alpha_1$, $\alpha' = \alpha(q')^{-1}$ and $z = xu^{-(k+i)}\alpha'$. Then $p_x z q'$ is a factor of $(\alpha_2 p_x^l)^2$ and $|z| \equiv |\alpha_2| \pmod{|p_x|}$. Hence, we can use Lemma 17, where $\beta_1 = \varepsilon$, $\beta_2 = p_x$, $\gamma_1 = \alpha_1$ and $\gamma_2 = \alpha_2$.

1. I) If Case (A) applies, we get $q' = p_x$, i.e. $\alpha_1 = \varepsilon$ and we are done.

1. II) In Case (B), $z \in p_x^* w$, where w is a prefix of p_x of length $|\alpha_2|$. Then there are two possibilities.

1. II. A) Firstly, we may have $|w_2| \geq \left|\alpha_2 p_x^l\right|$. Lemma 18 implies that $w = p_x\alpha_2^{-1} = \alpha_1$. Thus, we obtain $z \in p_x^*\alpha_1$. That means $xu^{-(k+i)}p_x^{c-1} = p_x^m$ for some m. Therefore, x and u commute by Lemma 1.

1. II. B) Another option is that $|w_2| < \left|\alpha_2 p_x^l\right|$. Then $z \in p_x^* w$, where $|w| = |\alpha_2|$ and $wq' = p_x\alpha_2$. Thus $\alpha_2\alpha_1 \leq_s \alpha_1\alpha_2\alpha_2$ and α_1 and α_2 commute by Lemma 6.

1. III) In Case (C), there are also two possible subcases.

1. III. A) The first of them, when $|w_1| \geq \left|\alpha_2 p_x^l\right|$, is the same as 1. II. A).

1. III. B) The last possibility is when $|w_1| < \left|\alpha_2 p_x^l\right|$. Then $p_x w'$ is a suffix of $p_x q'$. That leads to $p_x w' = \alpha_2\alpha_2\alpha_1$ and $\alpha_1\alpha_2 \leq_p \alpha_2\alpha_2\alpha_1$. Hence α_1 and α_2 commute by Lemma 5.

2. *Case $i-k$ is even.* We can write $i - k = 2n$ for some $n \geq 1$. Then $\alpha = x^n u^{k+n}$, $y'y = x^k u^{-(2k+2n)} x^n u^{k+n}$ and $yy' = u^{-(k+n)} x^{k+n}$. Since we assume $|x| > (i+k-1)\,|u| = (2k+2n-1)\,|u|$, we can write $x = u^{k+2n} x_1 u^{k-1} = u^{k+2n-1} x_2 u^k$. The words x_1 and x_2 are nonempty and satisfy $ux_1 = x_2 u$; in other words they are conjugate by u. Thus, by Lemma 9, there are words s and t such that

$$x_1 = (st)^{m_1}, \qquad x_2 = (ts)^{m_1}, \qquad \text{and} \qquad u = (ts)^{m_2} t \tag{1}$$

for some $m_1 \geq 1, m_2 \geq 0$. We have

$$\begin{aligned} y'y &= (u^{k+2n-1} x_2 u^k)^{k-1} u^{k+2n-1} x_2 x_1 u^{k-1} (u^{k+2n} x_1 u^{k-1})^{n-1} u^{k+n}, \\ yy' &= u^n x_1 u^{k-1} (u^{k+2n} x_1 u^{k-1})^{k+n-1}. \end{aligned}$$

In these equalities, we can replace u, x_1, x_2 by the expressions from (1) accordingly. Note that (using the notation from Lemma 11) there are words w_1 and w_2 over the alphabet A such that $g(w_1) = y'y$ and $g(w_2) = yy'$. Moreover, w_1 contains aa as a cyclic factor, whereas w_2 does not. Hence, Lemma 11 implies that s and t commute, which means x and u commute as well.

3. *Case $k = 1$, $i-k$ is odd and $i \geq 6$.* In this case, we have

$$\left|u^{-i}\alpha\right| = \frac{i-k}{2}(|x| - |u|) > 2|p_x|.$$

Denote by z the word $xu^{-(i+1)}\alpha q^{-2}$ where $q = \alpha_2\alpha_1$. If we put $\beta = u^{-i}\alpha q^{-2}$, we can write $z = xu^{-1}\beta$, where $\beta = \beta' q^r$ for some β' such that $|\beta'| < |q|$ and some $r \geq 0$. Note that either a) $\beta' = u^{-i}\alpha_1$ or b) $\beta' = u^{-i}p_x\alpha_1 = u^{-i}\alpha_1 q$ depending on whether α_1 is longer than u^i or not. The word qzq is a factor of $(\alpha_2 p_x^l)^2$. Since $|\alpha_1| - (i+k)|u| \equiv |\alpha_2| \pmod{|p_x|}$, we deduce $|z| \equiv |\alpha_2| \pmod{|p_x|}$. We can use Lemma 19 which leads to two main options: either $z \in q^*\alpha_2 q^*$ or $z \in \alpha_2' q^*$, where $\alpha_1\alpha_2' = q$.

3. I) Let first

$$xu^{-(i+1)}\alpha q^{-2} \in q^*\alpha_2 q^*,$$

i.e. $xu^{-1}\beta' q^r = q^m\alpha_2 q^n$ for some $m, n \geq 0$. If x is not primitive, we get either $\alpha_1\alpha_2 = \alpha_2\alpha_1$ if $m > 0$, or $\alpha_1\alpha_2 \leq_p \alpha_2\alpha_2\alpha_1$ for $m = 0$. In both of these cases, α_1 and α_2 commute. Therefore, we assume that x is primitive. If $r > n$ we immediately get that α_1 and α_2 commute. Hence, we can write

$$xu^{-1}\beta' = q^m\alpha_2 q^{n-r},$$

where $0 \leq m + n - r \leq 1$ follows from a simple length argument.

3. I. A) Let $xu^{-1}\beta' = \alpha_2$. There are two subcases.

3. I. A. a) Let first $xu^{-(i+1)}\alpha_1 = \alpha_2$. This case cannot happen: we would have $2|\alpha_1| = \left|u^{i+1}\right|$, i.e. $|\alpha_1| < \left|u^i\right|$ — a contradiction.

3. I. A. b) Let $xu^{-(i+1)}x\alpha_1 = \alpha_2$. Then x and u commute by Lemma 23.

3. I. B) Let $xu^{-1}\beta' = q\alpha_2 = \alpha_2 x$.

3. I. B. a) In the first case, we have $xu^{-(i+1)}\alpha_1 = \alpha_2 x$ which is equal to $\alpha_1\alpha_2 u^{-(i+1)}\alpha_1 = \alpha_2\alpha_1\alpha_2$. Thus x and u commute by Lemma 22.

3. I. B. b) Let $xu^{-(i+1)}x\alpha_1 = \alpha_2 x$. That means $2|\alpha_1| = |u^{i+1}|$. This implies ($i+1$ is odd) that $u = u_1u_2$ with $|u_1| = |u_2|$ and $\alpha_1 = u^d u_1$, $d = i/2$. Since α_1 is a suffix of x, and also u is a suffix of x, we deduce $u_1 = u_2$ and $\alpha_1 = u_1^{i+1}$. Then $u_1^{i+1}\alpha_2 u_1^{-2(i+1)} u_1^{i+1}\alpha_2 u_1^{i+1} = \alpha_2 u_1^{i+1}\alpha_2$ is a nontrivial relation showing that u_1 and α_2 commute. It follows that u and x commute as well.

3. I. C) Let $xu^{-1}\beta' = \alpha_2 q$.

3. I. C. a) The case $xu^{-(i+1)}\alpha_1 = \alpha_2 q$ leads to $\alpha_1\alpha_2 = \alpha_2\alpha_2 u^{i+1}$. Hence, x and u commute by Lemma 21.

3. I. C. b) In the case $xu^{-(i+1)}x\alpha_1 = \alpha_2 q$, we get $xu^{-(i+1)}x = \alpha_2\alpha_2$. We also have $2|\alpha_1| = |u^{i+1}|$. Now we can proceed similarly to case 3. I. B. b). Once again, we obtain $u = u_1u_2$ with $|u_1| = |u_2|$ and $\alpha_1 = u^d u_1$, $d = i/2$. Hence $u_2 \leq_p \alpha_2$. The equality $xu^{-(i+1)}x = \alpha_2\alpha_2$ implies that $u_1 \leq_p \alpha_2$. Thus $u_1 = u_2$ and $\alpha_1 = u_1^{i+1}$. Then $u_1^{i+1}\alpha_2 u_1^{-2(i+1)} u_1^{i+1}\alpha_2 = \alpha_2\alpha_2$ is a nontrivial relation showing that u_1 and α_2 commute. It follows that u and x commute as well.

3. II) Let now

$$xu^{-(i+1)}\alpha q^{-2} \in \alpha_2' q^*,$$

i.e. $xu^{-1}\beta' q^r = \alpha_2' q^m$ for some $m \geq 0$, where $\alpha_1\alpha_2' = \alpha_2\alpha_1$. Then $\alpha_1 xu^{-1}\beta' q^r = \alpha_1\alpha_2' q^m = \alpha_2\alpha_1 q^m$. If x is not primitive, we obtain $\alpha_2\alpha_1 \leq_p \alpha_1\alpha_1\alpha_2$ and α_1 and

α_2 commute by Lemma 5. Hence, assume that x is primitive. By simple length arguments, we can show $0 \leq m - r \leq 1$.

3. II. A) Let $xu^{-1}\beta' = \alpha_2'$.

3. II. A. a) The case $xu^{-(i+1)}\alpha_1 = \alpha_2'$ cannot occur; the reasoning is the same as in case 3. I. A. a).

3. II. A. b) In the case $xu^{-(i+1)}\alpha_1 q = \alpha_2'$, we get $xu^{-(i+1)}\alpha_1\alpha_1 = \varepsilon$, that is $u^{i+1} = \alpha_1\alpha_1 x$. Hence x and u commute by Lemma 24.

3. II. B) Let $xu^{-1}\beta' = \alpha_2' q$.

3. II. B. a) The case $xu^{-(i+1)}\alpha_1 = \alpha_2' q$ implies $xu^{-(i+1)} = \alpha_2'\alpha_2$, i.e. $\alpha_1\alpha_2 = \alpha_2'\alpha_2 u^{i+1}$. Then $\alpha_1\alpha_1\alpha_2 = \alpha_1\alpha_2'\alpha_2 u^{i+1} = \alpha_2\alpha_1\alpha_2 u^{i+1}$. Since $\alpha_2\alpha_1 \leq_p \alpha_1\alpha_1\alpha_2$, α_1 and α_2 commute by Lemma 5.

3. II. B. b) In the case $xu^{-(i+1)}\alpha_1 q = \alpha_2' q$, we have $\alpha_1 x u^{-(i+1)}\alpha_1 = \alpha_1\alpha_2' = \alpha_2\alpha_1$. That leads to $\alpha_1\alpha_1\alpha_2 = \alpha_2 u^{i+1}$. Since $\alpha_1 \leq_p u^i$, we get $\alpha_2\alpha_1 \leq_p \alpha_1\alpha_1\alpha_2$ and α_1 and α_2 commute by Lemma 5.

4. *Case* $i = 4$ *and* $k = 1$. Here we have $|\alpha| = \frac{3}{2}|x| + \frac{5}{2}|u|$ and $|yy'| = \frac{5}{2}(|x| - |u|)$.

4. I) If x is not primitive, we get $|\alpha| > 3|p_x|$ and $|yy'| > 3|p_x|$. Denote by z the word $\alpha_1 x u^{-(i+1)} p_x$. Then $|y'y| - |p_x z p_x| = |\alpha| - |p_x p_x p_x \alpha_1| \geq 0$, which means that the word $(yy')^2 = (\alpha_2 p_x^l)^2$ contains $p_x z p_x$. We already know that $l \geq 1$ and $|\alpha_1| - (i+k)|u| \equiv |\alpha_2| \pmod{|p_x|}$. Therefore, z satisfies the conditions of Lemma 20.

4. I. A) Let first

$$\alpha_1 x u^{-(i+1)} p_x \in p_x^* \alpha_2 p_x^*.$$

Here we get that $\alpha_1\alpha_1\alpha_2$ is a prefix of $(\alpha_1\alpha_2)^*\alpha_2(\alpha_1\alpha_2)^*$, because x is not primitive. If $\alpha_1\alpha_1\alpha_2$ is a prefix of $\alpha_1\alpha_2\alpha_1$, then α_1 and α_2 obviously commute. If $\alpha_1\alpha_1\alpha_2$ is a prefix of $\alpha_1\alpha_2\alpha_2\alpha_1$, then α_1 and α_2 commute by Lemma 5. And finally, if $\alpha_1\alpha_1\alpha_2$ is a prefix of $\alpha_2(\alpha_1\alpha_2)^\omega$, then also $\alpha_1\alpha_1\alpha_1\alpha_2$ is a prefix of $(\alpha_1\alpha_2)^\omega$ and α_1 and α_2 commute by Lemma 5.

4. I. B) Let now

$$\alpha_1 x u^{-(i+1)} p_x \in p_x^* \alpha_2'.$$

In this situation, a simple length argument yields $\alpha_1 p_x^m u^{-(i+1)} p_x = p_x^n \alpha_2'$ for some $m > 1$ and $n > 0$. Since $p_x = \alpha_1\alpha_2 = \alpha_2'\alpha_1$, we get $\alpha_1\alpha_2' = \alpha_2'\alpha_1$, i.e. α_1 and α_2' commute. Then $\alpha_2 = \alpha_2'$ and α_1 and α_2 commute as well.

4. II) Assume that x is primitive. Then, there are three possibilities.

4. II. A) Let $|u^5| = |x|$. This means $x = u^5$ and x and u commute.

4. II. B) Let $|u^5| > |x|$. This leads to $\alpha = xx\alpha_1$, $yy' = \alpha_2 x = \alpha_2\alpha_1\alpha_2$ and $y'y = xu^{-5}xx\alpha_1$. We also have $|\alpha_1| < \frac{|u|}{2}$ and, since $|x| > 4|u|$, we get $|\alpha_2| > 7\frac{|u|}{2}$. Thus, we can write $y'y = \alpha_1(\alpha_2 u^{-1})(u^{-4}x)\alpha_1\alpha_2\alpha_1$. We get that $\alpha_1\alpha_2\alpha_1$ is a factor of $(\alpha_2\alpha_1\alpha_2)^\omega$. Then, Lemmas 3 and 4 allow only these three options:

4. II. B. a) Let $\alpha_1(\alpha_2 u^{-1})(u^{-4}x)\alpha_1 = \alpha_2$. Therefore, x and u commute by Lemma 23.

4. II. B. b) Let $\alpha_1\alpha_1(\alpha_2 u^{-1})(u^{-4}x) = \alpha_2$. We get $u^5 = \alpha_1\alpha_1 x$ and x and u commute by Lemma 24.

4. II. B. c) And finally, let $\alpha_1\alpha_1(\alpha_2 u^{-1})(u^{-4}x)\alpha_1 = p_{\alpha_2}^m \alpha_1 p_{\alpha_2}^n$, where $m, n \geq 1$ and $p_{\alpha_2}^{m+n} = \alpha_2$. In this case, we implicitly suppose that α_2 is not primitive. Then either $\alpha_1^2 p_{\alpha_2} \leq_p p_{\alpha_2}^m \alpha_1$ or $p_{\alpha_2}^m \alpha_1 \leq_p \alpha_1^2 p_{\alpha_2}$. In either case, α_1 and α_2 commute by Lemma 5.

4. II. C) Let $\left|u^5\right| < |x|$. We get $\alpha = x\alpha_1$, $yy' = \alpha_2 xx$ and $y'y = xu^{-5}x\alpha_1$. Here we have $|\alpha_1| = |\alpha_2| + \left|u^5\right|$. Since $|\alpha_1| > 5|u|$, we can write $y'y = (\alpha_1\alpha_2 u^{-1}) \times (u^{-4}\alpha_1)\alpha_2\alpha_1$, i.e. $\alpha_2\alpha_1$ is a factor of $y'y$. Since $y'y$ and yy' are conjugate, this $\alpha_2\alpha_1$ must occur somewhere within $(\alpha_2\alpha_1\alpha_2\alpha_1\alpha_2)^2$. Lemma 4 allows these three options:

4. II. C. a) Let $\alpha_1\alpha_2 u^{-5}\alpha_1 = \alpha_2\alpha_1\alpha_2$. Therefore, x and u commute by Lemma 22.

4. II. C. b) Let $\alpha_1\alpha_2 u^{-5}\alpha_1 = \alpha_2\alpha_2\alpha_1$. This case leads to $\alpha_1\alpha_2 = \alpha_2\alpha_2 u^5$, and x and u commute by Lemma 21.

4. II. C. c) Let $\alpha_2\alpha_1$ be "over the edge" $\alpha_2\alpha_2$. Formally, this situation corresponds to $(\alpha_2\alpha_1\alpha_2\alpha_1\alpha_2)^2 = w_1\alpha_1\alpha_2 u^{-5}\alpha_1\alpha_2\alpha_1 w_2$ where $w_1 \neq \varepsilon$ and $|w_2| > |\alpha_2\alpha_1\alpha_2|$. Since u^4 is a prefix of α_1, we may write $\alpha_1 = u^4\overline{\alpha}_1$. Then we can divide this situation into four subcases.

4. II. C. c. i) Let $|w_2| \leq |\alpha_2\alpha_2\alpha_1\alpha_2|$. In this case, α_2 is "over the edge" or the edge is between α_2 and α_1. Hence, this α_2 must occur within $\alpha_2\alpha_2$. Lemma 3 implies that $\alpha_1\alpha_1\alpha_2 u^{-5}\alpha_1 = p_{\alpha_2}^m \alpha_1\alpha_2\alpha_1 p_{\alpha_2}^n$ for some m and n such that $p_{\alpha_2}^{m+n} = \alpha_2$ and $m > 0$. Since $|\alpha_2\alpha_1| < \left|\alpha_1\alpha_1\alpha_2 u^{-1}\right|$, we get $p_{\alpha_2}^m \alpha_1 \leq_p \alpha_1\alpha_1$ and α_1 and α_2 commute by Lemma 5.

4. II. C. c. ii) Let $|\alpha_2\alpha_2\alpha_1\alpha_2| < |w_2| < \left|\alpha_2 u^4\alpha_2\alpha_1\alpha_2\right|$. Since p_u is a suffix of $\overline{\alpha}_1\alpha_2$, Lemma 3 implies that $\alpha_2\alpha_1\alpha_2\alpha_1\alpha_2 = p_u^m\overline{\alpha}_1\alpha_1\alpha_2 u^{-1}\overline{\alpha}_1\alpha_2 p_u^n$ for some m and n such that $p_u^{m+n} = u^4$ and $n > 0$. The word $p_u^n\overline{\alpha}_1\alpha_2$ is a suffix of the left-hand side of the equation while $\overline{\alpha}_1\alpha_2 p_u^n$ is a suffix of the right-hand side. Hence, p_u commutes with $\overline{\alpha}_1\alpha_2$ and thus also with $p_x = p_u^{m+n}\overline{\alpha}_1\alpha_2$.

4. II. C. c. iii) Let $|w_2| \geq \left|\alpha_2 u^4\alpha_2\alpha_1\alpha_2\right|$. In this case, the primitive word $\overline{\alpha}_1\alpha_2 u^4$ from $y'y$ (see Lemma 8) must occur in $\alpha_2\alpha_1\alpha_2\alpha_1\alpha_2$, which in turn is a factor of $(\overline{\alpha}_1\alpha_2 u^4)^\omega$. Lemma 4 forces the equality $\overline{\alpha}_1 u^4\overline{\alpha}_1\alpha_2 u^{-1} = \overline{\alpha}_1\alpha_2\alpha_2 u^4$, i.e. $\alpha_1\alpha_2 = \alpha_2\alpha_2 u^5$. Therefore, x and u commute by Lemma 21. $\square$

6. Binary equality words with two b's — part 2

Let us now explain how our results can be used to deal with binary equality words with two b's. The following lemma is an elementary case of a general theory developed in [4] and [5].

Lemma 26. *Suppose that w is not a binary equality word. Let f be a binary morphism*

$$f\colon \quad \begin{aligned} a &\mapsto a \\ b &\mapsto a^i b a^j, \end{aligned}$$

with $i, j \geq 0$. Then $f(w)$ is not a binary equality word.

Proof: Suppose that $g \circ f(w) = h \circ f(w)$ for binary morphisms g and h. Since w is not a binary equality word, we have that either $g \circ f = h \circ f$, or both $g \circ f$ and $h \circ f$ are periodic.

If $g \circ f = h \circ f$, then $g(a) = h(a)$, and $g(a^i b a^j) = h(a^i b a^j)$. Therefore $g = h$.

Let both $g \circ f$ and $h \circ f$ be periodic. Then $g(a), g(a^i b a^j) \in t^*$ for some t. Therefore also g is periodic. Similarly, h is periodic.

This completes the proof. □

Note that the previous proof has in fact verified the two conditions of [4, Theorem 6].

Example 27. Take the word $w' = ababaaaa$ and suppose that there exist two distinct nonperiodic morphisms $g, h\colon \{a,b\}^* \to \Sigma^*$ such that $g(w') = h(w')$. Now we can take the word $w = abbaaa$ and a morphism $f\colon \{a,b\}^* \to \{a,b\}^*$ defined by

$$f\colon \quad \begin{aligned} a &\mapsto a \\ b &\mapsto ba. \end{aligned}$$

It is easy to see that $f(w) = w'$ and $g \circ f(w) = h \circ f(w)$. Since the word w is not a binary equality word by Theorem 25, neither w' is.

If w is an equality word, then J. D. Day at al., see [4, Theorem 6], do not tell us anything about $f(w)$. However, sometimes the solution of w yields a solution of $f(w)$. We give a nontrivial example.

Example 28. Consider the word $ababa$. We can take $w = bba$ and $f(b) = ab$. Take the most simple solution for bba, namely

$$g'\colon \quad \begin{aligned} a &\mapsto a \\ b &\mapsto cc \end{aligned} \qquad\qquad h'\colon \quad \begin{aligned} a &\mapsto cca \\ b &\mapsto c. \end{aligned}$$

This solution is not helpful because $g'(a)$ is not a prefix of $g'(b)$, nor $h'(a)$ a prefix of $h'(b)$. However, considering instead $\theta \circ g''$ and $\theta \circ h''$, where $\theta(a) = a$, $\theta(c) = ab$, and

$$g''\colon \quad \begin{aligned} a &\mapsto a \\ b &\mapsto c^4 \end{aligned} \qquad\qquad h''\colon \quad \begin{aligned} a &\mapsto cca \\ b &\mapsto c^3, \end{aligned}$$

we get

$$g'''\colon \quad \begin{aligned} a &\mapsto a \\ b &\mapsto (ab)^4 \end{aligned} \qquad\qquad h'''\colon \quad \begin{aligned} a &\mapsto (ab)^2 a \\ b &\mapsto (ab)^3. \end{aligned}$$

Now $g'''(a) \leq_p g'''(b)$ and $h'''(a) \leq_p h'''(b)$ and we obtain a solution

$$g\colon \quad a \mapsto a \qquad\qquad h\colon \quad a \mapsto (ab)^2a$$
$$b \mapsto b(ab)^3 \qquad\qquad b \mapsto b$$

for the word $ababa$.

We have just seen that $a^l baba^l$ is a binary equality word for $l = 1$. The following lemma shows that it is true for greater l's as well.

Lemma 29. *The word $a^l baba^l$, where $l \geq 2$, is an equality word of two nonperiodic morphisms.*

Proof: Take the morphisms

$$\begin{aligned} g\colon \quad & a \mapsto a^{2l-1}ba^{2l-1}ba^{2l-1} \\ & b \mapsto ba^{2l-1}(g(a))^{l-2}a^{2l-1}b \\ h\colon \quad & a \mapsto a \\ & b \mapsto a^{l-1}ba^{2l-1}g(b)a^{2l-1}ba^{2l-1}g(b)a^{2l-1}ba^{l-1}. \end{aligned}$$

It is straightforward to verify that $g(a^l baba^l) = h(a^l baba^l)$ and both g and h are obviously nonperiodic. □

The morphisms in the proof of the previous lemma can be derived from the solutions of the equation $x^{l-1}y^2x^l = u^{l-1}v^2u^l$. We can take the following solution (see pages 52–53 in [8])

$$\begin{aligned} x &= (a^{l-1}b)^2a^{2l-1}, & u &= a, \\ y &= a^{l-1}bx^{l-1}a^{l-1}b, & v &= ba^{l-1}ba^{2l-1}x^{l-2}a^{l-1}bx^{l-1}(a^{l-1}b)^2a^{l-1}, \end{aligned}$$

and apply the morphism

$$f\colon \quad a \mapsto a, \qquad b \mapsto a^l b.$$

Then we obtain $g(a) = f(x)$, $g(b) = (g(a))^{-1}f(y)$, $h(a) = f(u)$ and $h(b) = (h(a))^{-1}f(v)$.

We can now present a result that yields a complete classification of binary equality words with two b's.

Theorem 30. *Let $g, h\colon \{a,b\}^* \to \Sigma^*$ be two different nonperiodic morphisms. Let $w \in \Sigma^*$ and $|w|_b = 2$. Then w is a binary equality word if and only if w is one of the following words:*

$$a^nb^2,\ \ b^2a^n,\ \ ba^nb,\ \ a^{n+1}b^2a^n,\ \ a^nb^2a^{n+1},\ \ a^nbaba^n,\ \ a^nbba^n,\ \ a^nba^{n+m}ba^m,$$

where $m, n \geq 0$.

Proof: We first list morphisms witnessing that all listed words are binary equality words.

For

$$g\colon\quad a \mapsto a^m \qquad\qquad h\colon\quad a \mapsto (a^{mn}b)^m$$
$$\quad\;\; b \mapsto (ba^{mn})^n \qquad\qquad\quad b \mapsto a^n$$

we have $a^n b^m \in \mathrm{Eq}(g,h)$ (see [3, Example 5.1]).

For $n = 2\,l + 1$ and

$$g\colon\quad a \mapsto a \qquad\qquad h\colon\quad a \mapsto ba^{2\,l+1}b$$
$$\quad\;\; b \mapsto b(ba^{2\,l+1}b)^{\,l}b \qquad\qquad\quad b \mapsto b$$

we have $ba^n b \in \mathrm{Eq}(g,h)$ (cf. [3, Theorem 6.2]).

For

$$g\colon\quad a \mapsto a^{2n+1}(ba^n)^2 \qquad\qquad h\colon\quad a \mapsto a$$
$$\quad\;\; b \mapsto ba^n \qquad\qquad\quad b \mapsto (a^n b)^2(a^{3n+1}ba^n b)^n$$

we have $a^{n+1}ba^n \in \mathrm{Eq}(g,h)$ (see Conclusion in [9]).

For $a^n baba^n$, see Lemma 29.

Example 13 yields morphisms with $\mathrm{Eq}(g,h) = \{a^n b, ba^n\}^+$. Therefore also $a^n bba^n, ba^{2n}b \in \mathrm{Eq}(g,h)$.

For

$$g\colon\quad a \mapsto a^2 \qquad\qquad h\colon\quad a \mapsto a$$
$$\quad\;\; b \mapsto b \qquad\qquad\quad b \mapsto a^n ba^m$$

we have $a^n ba^m \in \mathrm{Eq}(g,h)$. Then also $a^n ba^{n+m}ba^m \in \mathrm{Eq}(g,h)$.

The remaining words are mirror images of words already covered.

We now show that no other binary equality words with two b's exist. Let $w = a^i ba^j ba^k$. If $i = k = 0$, then w is one of the allowed words. By symmetry, we can further assume $i \geq k$ and $i > 0$.

I. Let $j \geq i + k + 1$. Then $aba^{j-(i+k)+1}b$ is periodicity forcing by Lemma 14. Lemma 26 implies that w is periodicity forcing using the morphism $f(b) = a^{i-1}ba^k$.

II. If $j = i + k$, then w is one of the allowed words.

III. Let $j < i + k$.

a) Let $k < j$. Then $a^{i+k-j+1}bab$ is periodicity forcing by Lemma 15. Then also w is periodicity forcing by Lemma 26 using $f(b) = a^{j-k-1}ba^k$.

b) Let $k = j$. Then w is a binary equality word for $j = k = 0$ or for $i = j = k = 1$. Otherwise, $i \geq 2$ and $k \geq 1$. Then $a^i baba$ is periodicity forcing by Lemma 16, and also w is periodicity forcing by Lemma 26 using $f(b) = ba^{k-1}$.

c) Let $k > j$. The word w is a binary equality word if $i = k$ and $j \leq 1$, or $i = k + 1$ and $j = 0$. Otherwise, $a^i bba^{k-j}$ is periodicity forcing by Theorem 25. Then also w is periodicity forcing by Lemma 26 with $f(b) = ba^j$. □

7. Conclusion

In this paper, we have covered an important part of unsolved cases in the classification of binary equality words. The difficulty of the proof, namely of Theorem 25, may be surprising. It is interesting to stress, that while the dual PCP is decidable even in the general case, there is no efficient decision procedure even in the binary case. This reflects a complicated question of algorithmic solving of general word equations, of which our equations are a special case (symmetric equations in four unknowns).

References

[1] Barbin-Le R. E., Le Rest M., *Sur la combinatoire des codes à deux mots*, Theoret. Comput. Sci. **41** (1985), no. 1, 61–80 (French. English summary).

[2] Baumslag G., *Topics in Combinatorial Group Theory*, Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 1993.

[3] Culik K. II, Karhumäki J., *On the equality sets for homomorphisms on free monoids with two generators*, RAIRO Inform. Théor. **14** (1980), no. 4, 349–369.

[4] Day J. D., Reidenbach D., Schneider J. C., *On the dual post correspondence problem*, Internat. J. Found. Comput. Sci. **25** (2014), no. 8, 1033–1048.

[5] Day J. D., Reidenbach D., Schneider J. C., *Periodicity forcing words*, Theoret. Comput. Sci. **601** (2015), 2–14.

[6] Ehrenfeucht A., Karhumäki J., Rozenberg G., *The (generalized) Post correspondence problem with lists consisting of two words is decidable*, Theoret. Comput. Sci. **21** (1982), no. 2, 119–144.

[7] Ehrenfeucht A., Karhumäki J., Rozenberg G., *On binary equality sets and a solution to the test set conjecture in the binary case*, J. Algebra **85** (1983), no. 1, 76–85.

[8] Hadravová J., *Structure of Equality Sets*, PhD. Thesis, Charles University in Prague, Praha, 2015.

[9] Hadravová J., Holub Š., *Equation $x^iy^jx^k = u^iv^ju^k$ in words*, Language and Automata Theory and Applications, Lecture Notes in Comput. Sci., Springer, Cham, 2015, pp. 414–423.

[10] Halava V., Harju T., Hirvensalo M., *Binary (generalized) Post correspondence problem*, Theoret. Comput. Sci. **276** (2002), no. 1–2, 183–204.

[11] Halava V., Holub Š., *Binary (Generalized) Post Correspondence Problem is in P*, TUCS Technical Report, 785, Turku, 2006.

[12] Holub Š., *A unique structure of two-generated binary equality sets*, Developments in Language Theory (Ito M., ed.), 6th International Conf., Kyoto, 2002, Lecture Notes in Comput. Sci., 2450, Springer, Berlin, 2003, pp. 245–257.

[13] Holub Š., *Binary equality sets are generated by two words*, J. Algebra **259** (2003), no. 1, 1–42.

[14] Holub Š., *Binary equality languages for periodic morphisms*, Algebraic Systems, Formal Languages and Conventional and Unconventional Computation Theory, RIMS Kokyuroku, 1366, Kyoto University, 2004, pp. 1880–2818.

[15] Karhumäki J., Maňuch J., Plandowski W., *On defect effect of bi-infinite words*, Mathematical Foundations of Computer Science, 1998 (Brno), Lecture Notes in Comput. Sci., 1450, Springer, Berlin, 1998, pp. 674–682.

[16] Lothaire M., *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and Its Applications, 90, Cambridge University Press, Cambridge, 2002.

[17] Lyndon R. C., Schützenberger, M. P., *The equation $a^M = b^Nc^P$ in a free group*, Michigan Math. J. **9** (1962), no. 4, 289–298.

[18] Maňuch J., *Defect Theorems and Infinite Words*, TUCS Dissertations, 41, Turku, 2002.

[19] Post E. L., *A variant of a recursively unsolvable problem*, Bull. Amer. Math. Soc. **52** (1946), no. 4, 264–268.

[20] Rozenberg G., Salomaa A., eds., *Handbook of Formal Languages, Vol. 1: Word, Language, Grammar*, Springer, New York, 1997.

[21] Spehner J.-C., *Quelques problèmes d'extension, de conjugaison et de presentation des sous-monoïdes d'un monoïde libre*, Thèse, Université Paris VII, Paris, 1976 (French).

Š. Holub, J. Sýkora:
Charles University, Faculty of Mathematics and Physics,
Department of Algebra, Sokolovská 83, 186 75 Praha 8, Czech Republic

E-mail: holub@karlin.mff.cuni.cz
sykora.jir@seznam.cz

(*Received* January 19, 2018, *revised* February 4, 2018)