Periodicity and Unbordered Words: A Proof of the Extended Duval Conjecture *

Tero Harju
University of Turku

Dirk Nowotka
University of Stuttgart

April 12, 2007

Abstract

The relationship between the length of a word and the maximum length of its unbordered factors is investigated in this paper. Consider a finite word w of length n. We call a word *bordered* if it has a proper prefix which is also a suffix of that word. Let $\mu(w)$ denote the maximum length of all unbordered factors of w, and let $\partial(w)$ denote the period of w. Clearly, $\mu(w) \leq \partial(w)$.

We establish that $\mu(w) = \partial(w)$, if w has an unbordered prefix of length $\mu(w)$ and $n \geq 2\mu(w) - 1$. This bound is tight and solves the stronger version of an old conjecture by Duval (1983). It follows from this result that, in general, $n \geq 3\mu(w) - 3$ implies $\mu(w) = \partial(w)$ which gives an improved bound for the question raised by Ehrenfeucht and Silberger in 1979.

1 Introduction

Periodicity and borderedness are two properties of words which are investigated in this paper. These two fundamental notions play a rôle (explicitly or implicitly) in many areas. Just a few of those areas are string searching algorithms (Knuth, Morris, and Pratt 1977; Boyer and Moore 1977; Crochemore and Perrin 1991), data compression (Ziv and Lempel 1977; Crochemore, Mignosi, Restivo, and Salemi 1999), and codes (Berstel and Perrin 1985). These are classical examples, but also computational biology, e.g., sequence assembly (Margaritis and Skiena 1995) or superstrings (Breslauer, Jiang, and Jiang 1997), and serial data communications systems (Bylanski and Ingram 1980) are areas among others where periodicity and borderedness of words (sequences) are important concepts. It is well known that these two properties of words are not independent of each other. However, it is somewhat surprising that no clear relation has been established so far, despite the fact that this basic question has been around for more than 25 years.

Let us consider a finite word (a sequence of letters) w. We denote the length of w by $|w|$ and call a subsequence of consecutive letters of w a *factor* of w. The period of w, denoted by $\partial(w)$, is the smallest positive integer p such that the i-th letter equals the $(i+p)$-th letter for all $1 \leq i \leq |w| - p$. Let $\mu(w)$ denote the maximum length of all unbordered factors of w. A word is bordered if it has a proper prefix that is also a suffix, where we call a prefix proper if it is neither empty nor the entire word. For the investigation of the relationship between $|w|$ and the maximality of $\mu(w)$, that is, $\mu(w) = \partial(w)$, we consider the special case where the longest unbordered prefix of a word is of maximum length, that is, no unbordered factor is longer than that prefix. Let w be an unbordered word. Then a word wu is called a *Duval extension* (of w) if every unbordered factor of wu has length at most $|w|$, that is, $\mu(wu) = |w|$. We call wu a *trivial* Duval extension if $\partial(wu) = |w|$, or in other words, if u is a prefix of w^k for some $k \geq 1$. For example, let $w = abaabb$ and $u = aaba$. Then $wu = abaabbaaba$ is a nontrivial Duval extension of w since (i) w is unbordered, (ii) all factors of wu longer than w are bordered, that is, $|w| = \mu(wu) = 6$, and (iii) the period of wu is 7, and hence, $\partial(wu) > |w|$. Note that this example satisfies $|u| = |w| - 2$.

*A preliminary version of this paper appeared in the proceedings of the STACS conference 2004 (LNCS 2996:294–304, Springer-Verlag, Berlin, 2004).

In 1979 a line of research was initiated (Ehrenfeucht and Silberger 1979; Assous and Pouzet 1979; Duval 1982) exploring the relationship between the length of a word w and $\mu(w)$. In 1982 these efforts culminated in the following result by Duval: If $|w| \geq 4\mu(w) - 6$ then $\partial(w) = \mu(w)$. However, it was conjectured (Assous and Pouzet 1979) that $|w| \geq 3\mu(w)$ implies $\partial(w) = \mu(w)$ which follows from Duval's conjecture (Duval 1982).

Conjecture 1.1. *Let wu be a nontrivial Duval extension of w. Then $|u| < |w|$.*

After that, no progress was recorded, to the best of our knowledge, for 20 years. However, the topic remained popular, see for example Chapter 8 in (Lothaire 2002). The most recent results are by Mignosi and Zamboni (Mignosi and Zamboni 2002) and the authors of this article (Duval, Harju, and Nowotka). However, not Duval's conjecture but rather its opposite is investigated in those papers, that is: which words admit only trivial Duval extensions? It is shown in (Mignosi and Zamboni 2002) that unbordered, finite factors of Sturmian words allow only trivial Duval extensions; in other words if an unbordered, finite factor of a Sturmian word of length $\mu(w)$ is a prefix of w, then $\partial(w) = \mu(w)$. Sturmian words are binary infinite words of minimal subword complexity, that is, a Sturmian word contains exactly $n + 1$ different factors of length n for every $n \geq 1$; see (Morse and Hedlund 1940) or Chapter 2 in (Lothaire 2002). This result was later improved (Duval, Harju, and Nowotka) by showing that Lyndon words (Lyndon 1954) allow only trivial Duval extensions and the fact that every unbordered, finite factor of a Sturmian word is a Lyndon word but not vice versa. A Lyndon word is a primitive word that is minimal among all its conjugates with respect to some lexicographic order.

The main result in this paper is a proof of the extended version of Conjecture 1.1.

Theorem 1.2. *Let wu be a nontrivial Duval extension of w. Then $|u| < |w| - 1$.*

The example mentioned above already indicates that this bound on the length of a nontrivial Duval extension is tight. An example for arbitrary lengths of w is given later in Section 4. Recently, a new proof of Theorem 1.2 was given by Holub in (Holub 2005). Theorem 1.2 implies the truth of Duval's conjecture, as well as the following corollary (for any word w).

Corollary 1.3. *If $|w| \geq 3\mu(w) - 3$, then $\partial(w) = \mu(w)$.*

This corollary (see Section 4) confirms the conjecture by Assous and Pouzet in (Assous and Pouzet 1979) about a question asked by Ehrenfeucht and Silberger in (Ehrenfeucht and Silberger 1979).

Our main result, Theorem 1.2, is presented in Section 4 and its corollary in Section 5. Sections 4 and 5 use the notation introduced in Section 2 and preliminary results from Section 3. We conclude with Section 6.

2 Notation

In this section we introduce the notation of this paper. We refer to (Lothaire 1983; Lothaire 2002) for more basic and general definitions.

We consider a finite alphabet A of letters. Let A^* denote the monoid of all finite words over A including the empty word denoted by ε. We denote the i-th letter of a word w with $w_{(i)}$.[1] Let $w = w_{(1)}w_{(2)} \cdots w_{(n)}$. The word $w_{(n)} \cdots w_{(2)}w_{(1)}$ is called the *reversal* of w denoted by $\widetilde{w}$. We denote the length n of w by $|w|$. If w is not empty, then let $w^\bullet = w_{(1)}w_{(2)} \cdots w_{(n-1)}$. We define $\varepsilon^\bullet = \varepsilon$. An integer $1 \leq p \leq n$ is a *period* of w if $w_{(i)} = w_{(i+p)}$ for all $1 \leq i \leq n - p$. The smallest period of w is called the *minimum period* (or simply, the period) of w, denoted by $\partial(w)$. A word w is called *primitive* if $w = u^k$ implies $k = 1$, that is, $\partial(w)$ does not divide $|w|$. A *conjugate* of w is a word $w' = uv$ such that $vu = w$. Note that every conjugate of w occurs in $ww^\bullet$. A nonempty word u is called a *border* of a word w, if $w = uv = v'u$ for some words v and v'. We call w *bordered*, if it has a border that is shorter than w, otherwise w is called *unbordered*. Note that every

[1] In general, subscripts without brackets are used for variables in A^*, for example $w_i \in A^*$, and subscripts with brackets for variables in A, for example $w_{(i)} \in A$.

unbordered word is primitive and every bordered word w has a minimum border u such that $w = uvu$, where u is unbordered. Let $\mu(w)$ denote the maximum length of unbordered factors of w. We have that

$$\mu(w) \leq \partial(w) \ .$$

Indeed, let $u = u_{(1)}u_{(2)} \cdots u_{(\mu(w))}$ be an unbordered factor of w. If $\mu(w) > \partial(w)$ then $u_{(i)} = u_{(i+\partial(w))}$ for all $1 \leq i \leq \mu(w) - \partial(w)$ and $u_{(1)}u_{(2)} \cdots u_{(\mu(w)-\partial(w))}$ is a border of u; a contradiction.

Suppose $w = uv$, then u is called a *prefix* of w, denoted by $u \leq_{\mathrm{p}} w$, and v is called a *suffix* of w, denoted by $v \leq_{\mathrm{s}} w$. If u and v are both not the empty word, then u is called *proper prefix* of w, denoted by $u <_{\mathrm{p}} w$, and v is called *proper suffix* of w, denoted by $v <_{\mathrm{s}} w$. Let u and v be two nonempty words. We say that u *overlaps v from the left* (resp. *from the right*) if there is a word w such that $|w| < |u| + |v|$, and $u <_{\mathrm{p}} w$ and $v <_{\mathrm{s}} w$, (resp. $v <_{\mathrm{p}} w$ and $u <_{\mathrm{s}} w$). We say that u *overlaps* with v, if u overlaps v from the left or right. We say that u *intersects* with v, if u and v overlap or one is a factor of the other.

Example 2.1. *Let $A = \{a, b\}$ and $u, v, w \in A^*$ such that $u = abaa$ and $v = baaba$ and $w = abaaba$. Then $|w| = 6$, and 3, 5, and 6 are periods of w, and $\partial(w) = 3$. We have that a is the shortest border of u and w, whereas ba is the shortest border of v. We have $\mu(w) = 3$. We also have that u and v overlap since $u \leq_{\mathrm{p}} w$ and $v \leq_{\mathrm{s}} w$ and $|w| < |u| + |v|$.*

We continue with some more notation. Let w and u be words where w is unbordered. We call wu a *Duval extension* of w if every factor of wu longer than $|w|$ is bordered, that is, $\mu(wu) = |w|$. A Duval extension wu of w is called *trivial*, if $\partial(wu) = \mu(wu) = |w|$. A nontrivial Duval extension wu of w is called *minimal* if $u = u'a$ and $w = u'bw'$ where $a, b \in A$ and $a \neq b$, that is, wu is a nontrivial Duval extension and $wu^\bullet$ is a trivial Duval extension.

Example 2.2. *Let $w = abaabbabaababb$ and $u = aaba$. Then*

$$w.u = abaabbabaababb.aaba$$

(for the sake of readability, we use a dot to mark where w ends) is a nontrivial Duval extension of w of length $|wu| = 18$, where $\mu(wu) = |w| = 14$ and $\partial(wu) = 15$. However, wu is not a minimal Duval extension, whereas

$$w.u' = abaabbabaababb.aa$$

is minimal, with $u' = aa \leq_{\mathrm{p}} u$. Note that wu is not the longest nontrivial Duval extension of w since

$$w.v = abaabbabaababb.abaaba$$

is longer, with $v = abaaba$ and $|wv| = 20$ and $\partial(wv) = 17$. One can check that wv is a nontrivial Duval extension of w of maximum length, and at the same time wv is also a minimal Duval extension of w.

Let an integer p with $1 \leq p < |w|$ be called *point* in w. Intuitively, a point p denotes the place between $w_{(p)}$ and $w_{(p+1)}$ in w. A nonempty word u is called a *repetition word* at point p if $w = xy$ with $|x| = p$ and there exist words x' and y' such that $u \leq_{\mathrm{s}} x'x$ and $u \leq_{\mathrm{p}} yy'$. For a point p in w, let

$$\partial(w, p) = \min\big\{|u| \bigm| u \text{ is a repetition word at } p\big\}$$

denote the *local period* at point p in w. Note that the repetition word of length $\partial(w, p)$ at point p is necessarily unbordered and $\partial(w, p) \leq \partial(w)$. A factorization $w = uv$, with $u, v \neq \varepsilon$ and $|u| = p$, is called *critical*, if $\partial(w, p) = \partial(w)$, and if this holds, then p is called *critical point*.

Example 2.3. *The word*

$$w = ab.aa.b$$

has the period $\partial(w) = 3$ and two critical points, 2 and 4, marked by dots. The shortest repetition words at the critical points are aab and baa, respectively. Note that the shortest repetition words at the remaining points 1 and 3 are ba and a, respectively.

Let us consider alphabets of any finite size larger than one for the rest of this paper.

3 Preliminary Results

We state some auxiliary and well-known results about repetitions and borders in this section. These results will be used to prove Theorem 1.2 and Corollary 1.3 in Section 4. The first lemma recalls a well-known fact.

Lemma 3.1. *Let w be a primitive word over a k-letter alphabet. Then there exist at least k unbordered conjugates of w.*

Indeed, for every letter a in an alphabet A a lexicographic order $\lhd_a$ can be chosen such that a is minimal in A. It is not hard to show that the smallest conjugate w' of w with respect to $\lhd_a$ is unbordered. Note that $a \leq_{\mathrm{p}} w'$, and hence, every smallest conjugate with respect to a chosen order is different for a different letter.

Lemma 3.2. *Let $zf = gzh$ where $f, g \neq \varepsilon$. Let az' be the maximum unbordered prefix of az where a is a letter. If az does not occur in zf, then agz' is unbordered.*

Proof. Assume agz' is bordered, and let y be its shortest border. In particular, y is unbordered. If $|z'| \geq |y|$ then y is a border of az' which is a contradiction. If $|az'| = |y|$ or $|az| < |y|$ then az occurs in zf which is again a contradiction. If $|az'| < |y| \leq |az|$ then az' is not maximum since y is unbordered; a contradiction. □

The proof of the following lemma is easy and therefore omitted.

Lemma 3.3. *Let w be an unbordered word and $u \leq_{\mathrm{p}} w$ and $v \leq_{\mathrm{s}} w$. Then uw and wv are unbordered.*

The critical factorization theorem (CFT) is one of the main results about periodicity of words. A weak version of it was first conjectured by Schützenberger (Schützenberger 1979) and proved by Césari and Vincent (Césari and Vincent 1978). It was developed into its current form by Duval (Duval 1979). We refer to (Harju and Nowotka 2002a) for a short proof of the CFT.

Theorem 3.4 (CFT). *Every word w, with $\partial(w) \geq 2$, has at least one critical factorization $w = uv$, with $u, v \neq \varepsilon$ and $|u| < \partial(w)$, i.e., $\partial(w, |u|) = \partial(w)$.*

We have the following two lemmas about properties of critical factorizations.

Lemma 3.5. *Let $w = uv$ be unbordered and $|u|$ be a critical point of w. Then u and v do not intersect.*

Proof. Note that $\partial(w, |u|) = \partial(w) = |w|$ since w is unbordered. Let $|u| \leq |v|$ without loss of generality. Assume that u and v do intersect. First, if $u = u's$ and $v = sv'$ for a nonempty s, then $\partial(w, |u|) \leq |s| < |w|$. On the other hand, if $u = su'$ and $v = v's$, then s is a border of w. Finally, if $v = sut$, then $\partial(w, |u|) \leq |su| < |w|$. These contradictions prove the claim. □

The next result follows from Lemma 3.5.

Lemma 3.6. *Let $w = u_0u_1$ be unbordered and $|u_0|$ be a critical point of w. Then u_0xu_1 (resp. u_1xu_0) is either unbordered or has a minimum border g such that $|g| \geq |u_0| + |u_1|$ for any word x.*

Proof. Indeed, since $|u_0|$ is critical for w (for which $\partial(w) = |w|$), the words u_0 and u_1 are not factors of each other, and no suffix of u_0 can be a prefix of u_1. Therefore if g is a border of u_0xu_1, then it must be of the form u_0yu_1 for some y. □

The next theorem states a basic fact about minimal Duval extensions; see (Harju and Nowotka 2004) for a proof of it.

Theorem 3.7. *Let wu be a minimal Duval extension of the unbordered word w. Then au occurs in w where a is the last letter of w.*

The following Lemmas 3.8, 3.9 and 3.10 and Corollary 3.11 are given in (Duval 1982). Let $a_0, a_1 \in A$, with $a_0 \neq a_1$, and $t_0 \in A^*$. Let the sequences (a_i), (s_i), (s'_i), (s''_i), and (t_i), for $i \geq 1$, be defined by

- $a_i = a_{i \pmod 2}$, that is, $a_i = a_0$ (resp. $a_i = a_1$), if i is even (resp. odd),
- s_i is chosen so that $a_i s_i$ is the shortest border of $a_i t_{i-1}$,
- s'_i is chosen so that $a_{i+1} s'_i$ is the longest unbordered prefix of $a_{i+1} s_i$,
- s''_i is chosen so that $s'_i s''_i = s_i$,
- t_i is chosen so that $t_i s''_i = t_{i-1}$.

For any parameters of the above definition, the following holds.

Lemma 3.8. *For any a_0, a_1, and t_0 there exists an $m \geq 1$ such that*

$$|s_1| < \cdots < |s_m| = |t_{m-1}| \leq \cdots \leq |t_0|$$

and $s_m = t_{m-1}$ and $|t_0| \leq |s_m| + |s_{m-1}|$.

Lemma 3.9. *Let $z \leq_{\mathrm{p}} t_0$ such that neither of $a_0 z$ and $a_1 z$ occurs in t_0. Let $a_0 z_0$ and $a_1 z_1$ be the longest unbordered prefixes of $a_0 z$ and $a_1 z$, respectively, and let m a number given as in Lemma 3.8. Then*

1. *if $m = 1$ then $a_0 t_0$ is unbordered,*
2. *if $m > 1$ is odd, then $a_1 s_m$ is unbordered and $|t_0| \leq |s_m| + |z_0|$,*
3. *if $m > 1$ is even, then $a_0 s_m$ is unbordered and $|t_0| \leq |s_m| + |z_1|$.*

Lemma 3.10. *Let v be an unbordered factor of the unbordered word w of length $\mu(w)$. If v occurs twice in w, then $\mu(w) = \partial(w)$.*

Corollary 3.11. *Let wu be a Duval extension of the unbordered word w. If w occurs twice in wu, then wu is a trivial Duval extension.*

4 Main Result

The extended Duval conjecture is proven in this section.

Theorem 1.2. *Let wu be a nontrivial Duval extension of the unbordered word w. Then $|u| < |w| - 1$.*

Proof. Recall that every factor of wu longer than $|w|$ is bordered since wu is a Duval extension of w. Let z be the longest suffix of w that occurs twice in zu, the second occurrence possibly overlapping with the first z. We have $z \neq w$ since wu is otherwise trivial by Corollary 3.11.

If $z = \varepsilon$, we are done. Indeed, in that case the last letter a of w does not occur in u. Let $w = u'bw''$ and $u = u'cu''$ such that $b, c \in A$ and $b \neq c$. Now $wu'c$ is a minimal Duval extension of w, and by Theorem 3.7, w has the form $w = w'_0 a u' c w'_1$, where a is the last letter of w. Consider the factor $x = au'cw'_1 u$. If it is unbordered then $|u| + 1 < |x| \leq |w|$ and so $|u| < |w| - 1$. Otherwise, the shortest border g of x satisfies $|au| \leq |g|$, since, in this case, a does not occur in u. Since now g occurs in w and does not contain the last letter of w, we have $|u| < |w| - 1$ as claimed.

Assume $z \neq w$ and $z \neq \varepsilon$ in the following. Let $w = w'az$ and bz occur in zu. Note that bz does not overlap az from the right, since such an overlap gives $azz' = z''bz$ where $|z'| \leq |z|$ and wz' is unbordered by Lemma 3.3. We have

$$w = w'az \qquad \text{and} \qquad u = u'bzr$$

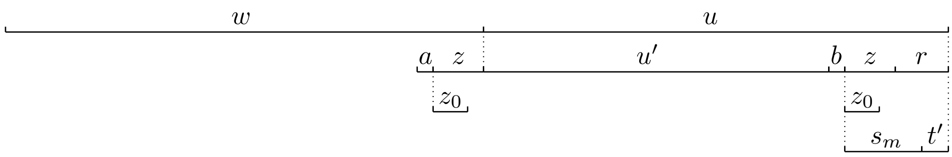
where z occurs in zr only once, that is, we assume bz to match the rightmost occurrence of z in u. Naturally, $a \neq b$ by maximality of z. Also, $w' \neq \varepsilon$, for otherwise $w = az$ and the prefix $azu'bz$ of wu is bordered, say with the shortest border g, but then either w is bordered (if $|g| \leq |z|$) or az occurs in zu (if $|g| > |z|$); a contradiction in both cases.

Let az_0 and bz_1 denote the longest unbordered prefix of az and bz, respectively. Let $a_0 = a$ and $a_1 = b$ and $t_0 = zr$ and the integer m be defined as in Lemma 3.9. We have then a word s_m, with its properties defined by Lemma 3.9, such that

$$t_0 = s_m t' \ .$$

Consider $x' = azu'bz_0$. We have $az \leq_{\mathrm{p}} a_0 zu$ and $x' \leq_{\mathrm{p}} a_0 zu$, and $bz_0 \leq_{\mathrm{s}} x'$. Also, az occurs only as a prefix in x'. It follows from Lemma 3.2 that x' is unbordered (where $z' = z_0$ and $f = u'bzr$ and $g = zu'b$ and $h = r$ in Lemma 3.2), and hence,

$$|x'| = |azu'bz_0| \leq |w| \ . \tag{1}$$



In the following we separately consider the two cases of even and odd parity of m.

Claim 4.1. *If m is even then $|u| < |w| - 1$.*

Now $m \geq 2$ and as_m $(= a_m s_m)$ is unbordered since m is even, and $|t_0| \leq |s_m| + |z_1|$ by Lemma 3.9.

Case: Let $|t_0| = |s_m| + |z_1|$ with $z_1 = z$. Then $|z| \leq |s_{m-1}|$ by Lemma 3.8, and moreover, $a_{m-1}s_{m-1}$ is the shortest border of $a_{m-1}t_{m-2} = bt_{m-2} \leq_{\mathrm{p}} bt_0 = bzr$. Because bs_{m-1} occurs twice in bt_{m-2} and zr marks the rightmost occurence of z in u, we have that z is not a proper prefix of s_{m-1}, and therefore, $|s_{m-1}| \leq |z|$. Hence, $|s_{m-1}| = |z|$. We have an immediate contradiction if $m = 2$ since then $|s_1| < |z|$ which contradicts $|z| \leq |s_{m-1}|$. Assume $m > 2$. But now, bz occurs in t_0 since bs_{m-1} is a border of bt_{m-2} and $t_i \leq_{\mathrm{p}} t_0$, for all $0 \leq i < m$, which is a contradiction.

Case: Let $|t_0| < |s_m| + |z_1|$ or $|z_1| < |z|$. Then $|t'| < |z|$.

Subcase: Let $|s_m| \leq |z_0|$. According to (1), $|azu'bz_0| \leq |w|$, and so

$$\begin{aligned}
|u| &= |azu| - |z| - 1 \\
&= |azu'bz_0| - |z_0| + |t_0| - |z| - 1 \\
&< |azu'bz_0| - |z_0| + |s_m| + |z_1| - |z| - 1 \\
&\leq |w| + |z_1| - |z| - 1 \\
&\leq |w| - 1
\end{aligned}$$

if $|t_0| < |s_m| + |z_1|$, or

$$\begin{aligned}
|u| &= |azu| - |z| - 1 \\
&= |azu'bz_0| - |z_0| + |t_0| - |z| - 1 \\
&\leq |azu'bz_0| - |z_0| + |s_m| + |z_1| - |z| - 1 \\
&\leq |w| + |z_1| - |z| - 1 \\
&< |w| - 1
\end{aligned}$$

if $|z_1| < |z|$. We have $|u| < |w| - 1$ in both cases.

Subcase: Let $|s_m| > |z_0|$. We have that as_m is unbordered, and since az_0 is the longest unbordered prefix of az, necessarily az is a proper prefix of as_m, and hence, $|z| < |s_m|$. Now, $azu'bs_m$ is unbordered, for otherwise its shortest border is longer than az, since no prefix of az is a suffix of as_m, and az occurs in u;

a contradiction. We have $|azu'bs_m| \leq |w|$ and similarly to the previous subcase, we obtain

$$\begin{aligned}|u| &= |azu| - |z| - 1\\ &= |azu'bs_m| - |s_m| + |t_0| - |z| - 1\\ &< |azu'bs_m| - |s_m| + |s_m| + |z_1| - |z| - 1\\ &\leq |w| + |z_1| - |z| - 1\\ &\leq |w| - 1\end{aligned}$$

if $|t_0| < |s_m| + |z_1|$, or

$$\begin{aligned}|u| &= |azu| - |z| - 1\\ &= |azu'bs_m| - |s_m| + |t_0| - |z| - 1\\ &\leq |azu'bs_m| - |s_m| + |s_m| + |z_1| - |z| - 1\\ &\leq |w| + |z_1| - |z| - 1\\ &< |w| - 1\end{aligned}$$

if $|z_1| < |z|$. We have $|u| < |w| - 1$ in both cases.

This proves Claim 4.1.

Claim 4.2. *If m is odd then $|u| < |w| - 1$.*

The word bs_m $(= a_m s_m)$ is unbordered, since m is odd. We have $|t_0| \leq |s_m| + |z_0|$; see Lemma 3.9. Note that $t_0 = s_m$ and $t' = \varepsilon$ by Lemma 3.9, if $m = 1$. Surely $s_m \neq \varepsilon$. In particular, $|t'| \leq |z_0|$.

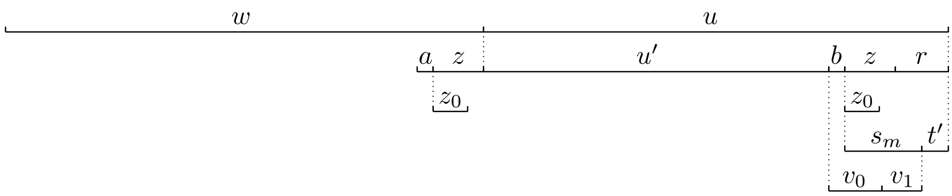
If $|s_m| < |z|$, then $|u| < |w| - 1$, since

$$|u| = |azu'bz_0| - |bz_0| + |bt_0| - |az|$$

and $|azu'bz_0| \leq |w|$, by (1), and $|t_0| \leq |s_m| + |z_0|$.

Assume thus that $|s_m| \geq |z|$, and hence, also $z \leq_{\mathrm{p}} s_m$. Since $s_m \neq \varepsilon$, we have $|bs_m| \geq 2$, and therefore, by the CFT (Theorem 3.4), there exists a critical point p in bs_m such that $bs_m = v_0 v_1$, where $|v_0| = p$. In particular,

$$bz \leq_{\mathrm{p}} v_0 v_1 \ . \tag{2}$$



Claim 4.3. *The factor $v_0 v_1$ occurs in w.*

Let, u_0' and u_1 be such that

$$u = u_0' v_0 v_1 u_1$$

where $v_0 v_1$ does not occur in u_0'. Note that $v_0 v_1$ does not overlap with itself since it is unbordered, and v_0 and v_1 do not intersect by Lemma 3.5. Consider the prefix $wu_0'bz$ of wu which is bordered by definition and has a shortest border g with $|g| > |z|$ (for otherwise g is also a border of w). We have $bz \leq_{\mathrm{s}} g$, and also $g \leq_{\mathrm{p}} w$ since g is unbordered and therefore $|g| \leq |w|$ by definition. Let

$$w = w_0 bz w_1 \tag{3}$$

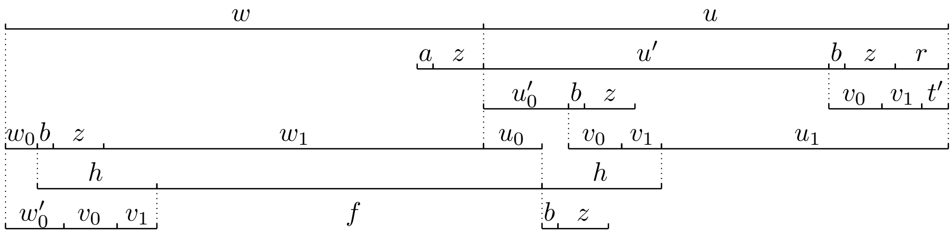
such that bz occurs in $w_0 bz$ only once, that is, we consider the leftmost occurrence of bz in w. Note that

$$|w_0 bz| \leq |g| \leq |u_0' bz| \tag{4}$$

where the first inequality comes from (3) and the second inequality from the fact that $|u_0'bz| < |g|$ implies that w is bordered. Let

$$f = bzw_1u_0'v_0v_1 \ .$$

If f is unbordered, then $|f| \leq |w|$, and hence, $|u_0'v_0v_1| \leq |w_0|$. Now, we have $|u_0'| < |w_0|$, which contradicts (4). Therefore, f is bordered. Let h be its shortest border.



Surely, $|bz| < |h|$, otherwise v_0v_1 is bordered by (2). So, $bz \leq_{\mathrm{p}} h$. Moreover, $|v_0v_1| \leq |h|$ otherwise bz occurs in s_m contradicting our assumption that bzr marks the rightmost occurrence of bz in u. So, $v_0v_1 \leq_{\mathrm{s}} h$, and v_0v_1 occurs in w since $w_0h \leq_{\mathrm{p}} w$ by (4). This proves Claim 4.3.

In the following we will consider a factor $\bar{v}_0\bar{v}_1$ in u of maximum length such that

1. $\bar{v}_0\bar{v}_1$ is unbordered,
2. $|\bar{v}_0|$ is a critical position in $\bar{v}_0\bar{v}_1$,
3. $\bar{v}_0\bar{v}_1$ occurs in w,
4. $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$,
5. v_0v_1 does not occur in $\bar{v}_1$,
6. either $\bar{v}_0 = v_0$ or $v_0v_1 \leq_{\mathrm{p}} \bar{v}_0$,
7. if $\bar{v}_0$ occurs in $\bar{w}_2$ then $az <_{\mathrm{s}} \bar{w}_2$ and $|\bar{t}'| \leq |z|$,

where $\bar{v}_0\bar{v}_1\bar{w}_2 \leq_{\mathrm{s}} w$ and $\bar{v}_0\bar{v}_1\bar{t}' \leq_{\mathrm{s}} u$ and $\bar{v}_0\bar{v}_1$ does neither occur in $\bar{w}_2$ nor in $\bar{t}'$. Note that v_0 and v_1 satisfy all conditions for $\bar{v}_0$ and $\bar{v}_1$, where $\bar{w}_2 = w_2$ and $\bar{t}' = t'$. In particular condition (7) follows from the fact that $v_0 \leq_{\mathrm{p}} bz$ and v_0 and v_1 do not intersect and $|t'| \leq |z_0|$.

Let

$$w = \bar{w}_0'\bar{v}_0\bar{v}_1\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_2'\bar{v}_0\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2$$

for some word $\bar{w}_2$ that does not contain $\bar{v}_0\bar{v}_1$, and

$$u = \bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_2'\bar{v}_0\bar{v}_1\bar{u}_1'\bar{v}_0\bar{v}_1\bar{t}'$$

such that $\bar{v}_0\bar{v}_1$ does not occur in $\bar{w}_k'$, for all $0 \leq k \leq i$, or $\bar{u}_\ell'$, for all $0 \leq \ell \leq j$. Note that these factorizations of w and u are unique, and, moreover, $\bar{w}_2 \neq \varepsilon$. Indeed, if $\bar{w}_2 = \varepsilon$ then $\bar{v}_0\bar{v}_1 \leq_{\mathrm{s}} w$ and $az \leq_{\mathrm{s}} \bar{v}_0\bar{v}_1$, since $|\bar{v}_0\bar{v}_1| \geq |v_0v_1| \geq |az|$, and az would occur in u; a contradiction.

The rest of the proof has the following outline: Claim 4.4 shows that either $\bar{w}_k' = \bar{u}_k'$ for all $1 < k \leq \min\{i,j\}$ or $|u| < |w| - 1$, and the Claims 4.6 (page 13), 4.7 (page 13), and 4.8 (page 14) show that the cases $i < j$, $i > j$, and $i = j$, respectively, imply $|u| < |w| - 1$.

Claim 4.4. *If $|u| \geq |w| - 1$ then $\bar{w}_k' = \bar{u}_k'$ for all $1 < k \leq \min\{i,j\}$.*

The proof goes by induction on k.

Case: First let $k = 1$. We show that $\bar{w}_1' = \bar{u}_1'$. Consider

$$f_1 = \bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_1'\bar{v}_0 \ .$$

If f_1 is unbordered, then $|u| < |w| - 1$ since $|f_1| \leq |w|$ and

$$|u| = |f_1| - |\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2| + |\bar{v}_1\bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume that f_1 is bordered, and let h_1 be its shortest border. We have that $h_1 = \bar{v}_1 g_1 \bar{v}_0$ for some g_1 (possibly empty), since $\bar{v}_0$ and $\bar{v}_1$ do not intersect. We show that $h_1 \leq_{\mathrm{p}} \bar{v}_1\bar{w}_1'\bar{v}_0$. Indeed, otherwise we have one of the following cases.

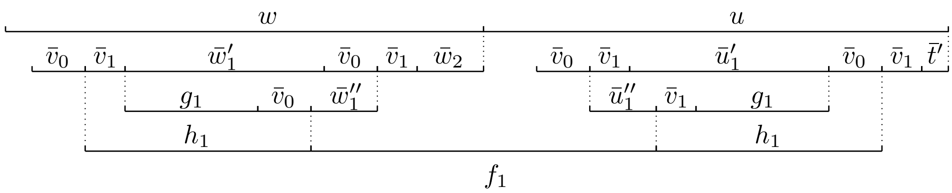
1. If $\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2 \leq_{\mathrm{p}} h_1$ then az occurs in u; contradicting the maximality of z.

2. If $|\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{v}_0| \leq h_1 < |\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2|$ then $\bar{v}_0$ occurs in $\bar{w}_2$. Let $\bar{v}_0\bar{w}_3 \leq_{\mathrm{s}} \bar{w}_2$ for some $\bar{w}_3$. Note that $\bar{v}_0$ is not a prefix of az since it begins with the letter b different from a. If $\bar{v}_0$ occurs in z then it overlaps with $\bar{v}_1$ since $bz \leq_{\mathrm{p}} \bar{v}_0\bar{v}_1$; a contradiction. If $\bar{v}_0$ does not occur in z, that is, $|az| \leq |\bar{v}_0\bar{w}_3|$, then $\bar{v}_0\bar{w}_3u'\bar{v}_0\bar{v}_1$ is unbordered (since otherwise its border is at least as long as $\bar{v}_0\bar{v}_1$, because $\bar{v}_0$ and $\bar{v}_1$ do not intersect and therefore it is longer than $|az|$; a contradiction). But now $|\bar{t}'| < |\bar{v}_0\bar{w}_3| - 1$ since $|\bar{t}'| < |az|$ and $|az| < |\bar{v}_0\bar{w}_3|$, and $|u| < |w| - 1$ follows.

3. If $|\bar{v}_1\bar{w}_1'\bar{v}_0| < h_1 < |\bar{v}_1\bar{w}_1'\bar{v}_0\bar{v}_1\bar{v}_0|$ then $\bar{v}_0$ and $\bar{v}_1$ intersect; a contradiction.

Moreover, $h_1 \leq_{\mathrm{s}} \bar{v}_1\bar{u}_1'\bar{v}_0$ since otherwise $\bar{v}_0\bar{v}_1$ occurs in h_1 (for $\bar{v}_1 \leq_{\mathrm{p}} h_1$ and $\bar{v}_0$ and $\bar{v}_1$ do not intersect) and $\bar{v}_0\bar{v}_1$ occurs in $\bar{v}_1\bar{w}_1'\bar{v}_0$; a contradiction. Let $\bar{w}_1''$ and $\bar{u}_1''$ be such that

$$\bar{w}_1'\bar{v}_0 = g_1\bar{v}_0\bar{w}_1'' \qquad \text{and} \qquad \bar{v}_1\bar{u}_1' = \bar{u}_1''\bar{v}_1 g_1 \,. \tag{5}$$



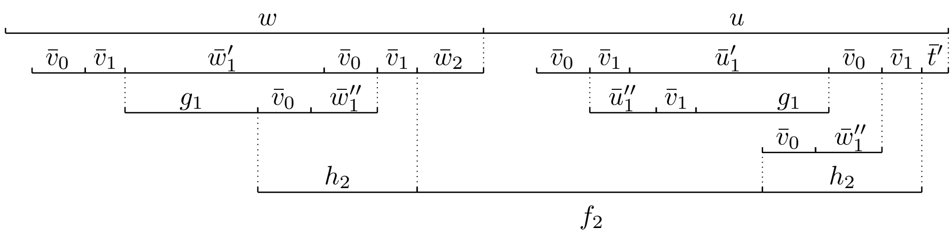
Consider,

$$f_2 = \bar{v}_0\bar{w}_1''\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_1'\bar{v}_0\bar{v}_1 \ .$$

If f_2 is unbordered, then $|u| < |w| - 1$ since $|f_2| \leq |w|$ and

$$|u| = |f_2| - |\bar{v}_0\bar{w}_1''\bar{v}_1\bar{w}_2| + |\bar{t}'|$$

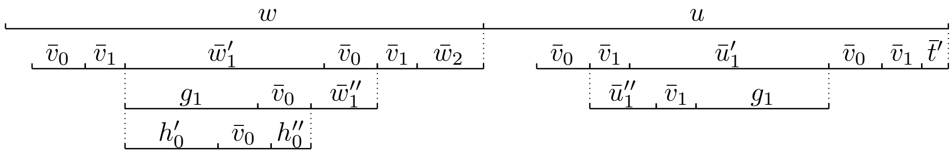
and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume that f_2 is bordered, and let h_2 be its shortest border. Since $\bar{v}_0$ and $\bar{v}_1$ do not intersect, $\bar{v}_0\bar{v}_1 \leq_{\mathrm{s}} h_2$. Also $h_2 \leq_{\mathrm{p}} \bar{v}_0\bar{w}_1''\bar{v}_1$ since $\bar{v}_0\bar{v}_1$ does not occur in $\bar{w}_2$ (and $\bar{v}_0$ and $\bar{v}_1$ do not intersect) and az does not occur in h_2 (and so h_2 does not stretch beyond w). We have $\bar{v}_0\bar{w}_1''\bar{v}_1 \leq_{\mathrm{p}} h_2$ since $\bar{v}_0\bar{v}_1$ occurs in $\bar{v}_0\bar{w}_1''\bar{v}_1$ only as a suffix. Hence, $h_2 = \bar{v}_0\bar{w}_1''\bar{v}_1$. Note that $|h_2| \leq |\bar{u}_1'\bar{v}_0\bar{v}_1|$ since otherwise $|h_2| \geq |\bar{v}_0\bar{v}_1\bar{u}_1'\bar{v}_0\bar{v}_1|$ (because $\bar{v}_0$ and $\bar{v}_1$ do not intersect) and $\bar{v}_0\bar{v}_1$ occurs twice in h_2, but $\bar{v}_0\bar{v}_1$ occurs only once in h_2 since it occurs only once in $\bar{w}_1'\bar{v}_0\bar{v}_1$. We have $\bar{w}_1'\bar{v}_0\bar{v}_1 = g_1 h_2$ and $h_2 \leq_{\mathrm{s}} \bar{u}_1'\bar{v}_0\bar{v}_1$.



Let

$$h_1 = \bar{v}_1 g_1 \bar{v}_0 = \bar{v}_1 h_0' \bar{v}_0 h_0'' \tag{6}$$

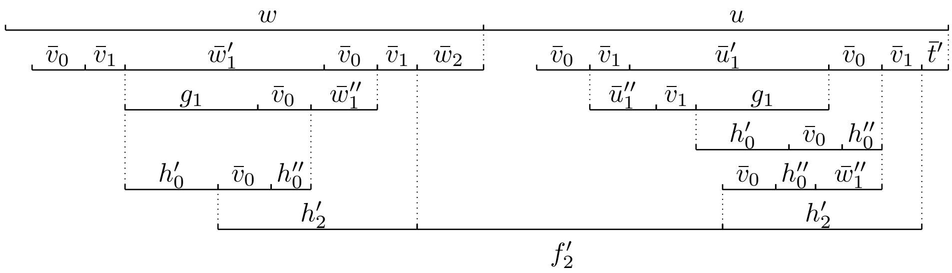
where $\bar{v}_0$ occurs only once in $h_0'\bar{v}_0$.



Let

$$f_2' = \bar{v}_0 h_0'' \bar{w}_1'' \bar{v}_1 \bar{w}_2 \bar{u}_0' \bar{v}_0 \bar{v}_1 \bar{u}_j' \cdots \bar{v}_0 \bar{v}_1 \bar{u}_1' \bar{v}_0 \bar{v}_1$$

with the shortest border h_2' (which exists if $|u| \geq |w| - 1$; as in the case of f_2) and $\bar{v}_0\bar{v}_1 \leq_{\mathrm{s}} h_2'$. We have $h_2' \leq_{\mathrm{p}} \bar{v}_0 h_0'' \bar{w}_1'' \bar{v}_1$ since $\bar{v}_0\bar{v}_1$ does not occur in $\bar{w}_2$ and az does not occur in h_2' (and so h_2' does not stretch beyond w). We have $\bar{v}_0 h_0'' \bar{w}_1'' \bar{v}_1 \leq_{\mathrm{p}} h_2'$ since $\bar{v}_0\bar{v}_1$ does not occur in $\bar{w}_1'$. Hence, $h_2' = \bar{v}_0 h_0'' \bar{w}_1'' \bar{v}_1$ and $\bar{w}_1' \bar{v}_0 \bar{v}_1 = h_0' \bar{v}_0 h_0'' \bar{w}_1'' \bar{v}_1$.



We have $\bar{v}_0 h_0'' \bar{w}_1'' \leq_{\mathrm{s}} g_1 \bar{v}_0 = h_0' \bar{v}_0 h_0''$, and $\bar{w}_1'' = \varepsilon$ follows from (6). This implies $\bar{w}_1' = g_1 \leq_{\mathrm{s}} \bar{u}_1'$; see (5).

Next, we show that actually $\bar{u}_1' = g_1$. Let

$$\bar{v}_1 g_1 = h_1'' \bar{v}_1 h_1' \tag{7}$$

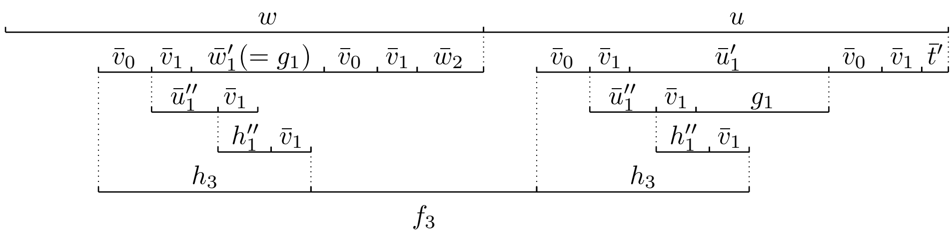
such that $\bar{v}_1$ occurs only once in $\bar{v}_1 h_1'$. Consider,

$$f_3 = \bar{v}_0 \bar{v}_1 \bar{w}_1' \bar{v}_0 \bar{v}_1 \bar{w}_2 \bar{u}_0' \bar{v}_0 \bar{v}_1 \bar{u}_j' \cdots \bar{v}_0 \bar{v}_1 \bar{u}_2' \bar{v}_0 \bar{u}_1'' h_1'' \bar{v}_1 \ .$$

If f_3 is unbordered, then $|u| < |w| - 1$ since $|f_3| \leq |w|$ and

$$|u| = |f_3| - |\bar{v}_0 \bar{v}_1 \bar{w}_1' \bar{v}_0 \bar{v}_1 \bar{w}_2| + |h_1' \bar{v}_0 \bar{v}_1 \bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0 \bar{v}_1|$ and $|\bar{w}_1'| = |g_1| \geq |h_1'|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_3 is bordered. Then f_3 has a shortest border h_3 such that $\bar{v}_0\bar{v}_1 \leq_{\mathrm{p}} h_3$ since $\bar{v}_0$ and $\bar{v}_1$ do not intersect. If $|h_3| > |\bar{v}_0 \bar{u}_1'' h_1'' \bar{v}_1|$ then $|h_3| \geq |\bar{v}_0 \bar{v}_1 g_1 \bar{v}_0 \bar{v}_1|$. Note that $h_3 \neq \bar{v}_0 \bar{v}_1 g_1 \bar{v}_0 \bar{v}_1$ since a shortest border is not bordered. Assume $|h_3| > |\bar{v}_0 \bar{v}_1 g_1 \bar{v}_0 \bar{v}_1|$ and $\bar{u}_1'' \neq \varepsilon$. But now h_3 contradicts the maximality of $\bar{v}_0\bar{v}_1$ since h_3 is unbordered (condition 1) and occurs both in w and u (condition 3) and $|h_3| > |g_1 \bar{v}_0 \bar{v}_1 \bar{t}'|$ (condition 4) and $h_3 = \bar{\bar{v}}_0 \bar{\bar{v}}_1$ where $\bar{\bar{v}}_0 = \bar{v}_0 \bar{v}_1 g_1 \bar{v}_0$ is a critical factorization of h_3, because otherwise $\bar{u}_1'' h_1'' \bar{v}_1$ occurs in $\bar{v}_1 g_1$ (since $\bar{v}_0$ and $\bar{v}_1$ do not intersect) contradicting (7) (conditions 2, 5, and 6), and $\bar{\bar{v}}_0$ does not occur in $\bar{w}_2$ since $\bar{v}_0\bar{v}_1 \leq_{\mathrm{p}} \bar{\bar{v}}_0$ (condition 7). We have $|h_3| = |\bar{v}_0 \bar{u}_1'' h_1'' \bar{v}_1|$ which implies $h_3 = \bar{v}_0 \bar{u}_1'' h_1'' \bar{v}_1$.



We have $\bar{u}_1''h_1''\bar{v}_1 \leq_{\mathrm{p}} \bar{v}_1 g_1 = h_1''\bar{v}_1 h_1'$, and $\bar{u}_1'' = \varepsilon$ follows from (7). We conclude that

$$\bar{w}_1' = g_1 = \bar{u}_1' \ .$$

Case: Let $1 < k \leq \min\{i, j\}$ and $\bar{w}_\ell' = \bar{u}_\ell'$, for all $1 \leq \ell < k$. Let us denote both $\bar{w}_\ell'$ and $\bar{u}_\ell'$ by v_ℓ', for all $1 \leq \ell < k$. We show that $\bar{w}_k' = \bar{u}_k'$. Consider

$$f_4 = \bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_k'\bar{v}_0 \ .$$

If f_4 is unbordered, then $|u| < |w| - 1$ since $|f_4| \leq |w|$ and

$$|u| = |f_4| - |\bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2| + |\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots \hat{v}_1 v_1'\bar{v}_0\bar{v}_1\bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_4 is bordered. Then f_4 has a shortest border h_4 such that $|\bar{v}_0\bar{v}_1| \leq |h_4|$. Let $h_4 = \bar{v}_1 g_4 \bar{v}_0$.

Subcase: Let $|\bar{v}_1\bar{w}_k'\bar{v}_0| < |h_4|$. Then there exists an $\ell < k$ such that

$$h_4 = \bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_{\ell+1}'\bar{v}_0\bar{v}_1 v_\ell''\bar{v}_0$$

where $v_\ell'' \leq_{\mathrm{p}} v_\ell'$. That implies $\bar{u}_k' = v_\ell''$, since $\bar{v}_0\bar{v}_1$ does neither occur in v_ℓ'' nor in $\bar{u}_k'$. Now, consider

$$f_5 = \bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_\ell''\bar{v}_0 \ .$$

If f_5 is unbordered, then $|u| < |w| - 1$ since $|f_4| < |f_5|$, see above. Assume, f_5 is bordered. Then f_5 has a shortest border h_5 such that $|h_4| < |h_5|$, for otherwise h_4 is not the shortest border of f_4, since either $h_4 \leq_{\mathrm{p}} h_5$ or $h_5 \leq_{\mathrm{p}} h_4$, and the latter implies that h_4 is bordered, and hence, not minimal. There exists an $\ell' < \ell$ such that

$$h_5 = \bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_{\ell'+1}'\bar{v}_0\bar{v}_1 v_{\ell'}''\bar{v}_0$$

where $v_{\ell'}'' \leq_{\mathrm{p}} v_{\ell'}'$. We have $|f_4| < |f_5| < |f_6|$ where

$$f_6 = \bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j' \cdots \bar{v}_0\bar{v}_1\bar{u}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_{\ell'}''\bar{v}_0 \ ,$$

which is either unbordered and $|u| < |w| - 1$ since $|f_4| < |f_5|$, or it is bordered with a shortest border h_6, and we have $|h_4| < |h_5| < |h_6|$ and a factor f_7, such that $|f_4| < |f_5| < |f_6| < |f_7|$, and so on, until eventually an unbordered factor is reached proving that $|u| < |w| - 1$.

Subcase: Let $h_4 \leq_{\mathrm{p}} \bar{v}_1\bar{w}_k'\bar{v}_0$. We also have that $h_4 \leq_{\mathrm{s}} \bar{v}_1\bar{u}_k'\bar{v}_0$ since $\bar{v}_0\bar{v}_1$ does not occur in $\bar{w}_k'$. Let $\bar{w}_k'\bar{v}_0 = g_4\bar{v}_0\bar{w}_k'' = g_4'\bar{v}_0\bar{g}_4'\bar{w}_k''$ and $\bar{v}_1\bar{u}_k' = \bar{u}_k''\bar{v}_1 g_4 = \bar{u}_k''\bar{g}_4''\bar{v}_1 g_4''$ such that $\bar{v}_0$ and $\bar{v}_1$ occur only once in $g_4'\bar{v}_0$ and $\bar{v}_1 g_4''$, respectively. We show that $\bar{w}_k'' = \bar{u}_k'' = \varepsilon$ next. Consider

$$f_8 = \bar{v}_0\bar{g}_4'\bar{w}_k''\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j'\bar{v}_0\bar{v}_1 \cdots \bar{u}_k'\bar{v}_0\bar{v}_1 \ .$$

If f_8 is unbordered, then $|u| < |w| - 1$ since $|f_8| \leq |w|$ and

$$|u| = |f_8| - |\bar{v}_0\bar{g}_4'\bar{w}_k''\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2| + |v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_8 is bordered with a shortest border h_8 then $\bar{v}_0\bar{v}_1 \leq_{\mathrm{s}} h_8$. If $|h_8| > |\bar{v}_0\bar{g}_4'\bar{w}_k''\bar{v}_1|$ then the same argument as in the case $|\bar{v}_1\bar{w}_k'\bar{v}_0| < |h_4|$ above shows that $|u| < |w| - 1$, that is, we have an increasing chain of factors longer than f_8 with a corresponding increasing chain of shortest borders longer than h_8 until $|u| < |w| - 1$ is shown. If $|h_8| < |\bar{v}_0\bar{g}_4'\bar{w}_k''\bar{v}_1|$ then $\bar{v}_0\bar{v}_1$ occurs in $\bar{w}_k'$; a contradiction. The remaining case is $h_8 = \bar{v}_0\bar{g}_4'\bar{w}_k''\bar{v}_1$ with $h_8 \leq_{\mathrm{s}} \bar{u}_k'\bar{v}_0\bar{v}_1$. It follows $\bar{v}_0\bar{g}_4'\bar{w}_k'' \leq_{\mathrm{s}} g_4\bar{v}_0$ and $g_4'\bar{v}_0\bar{g}_4'\bar{w}_k'' = g_4'\bar{v}_0\bar{g}_4'\bar{w}_k''\bar{w}_k''$ by the choice of g_4', and we have $\bar{w}_k'' = \varepsilon$. Consider

$$f_9 = \bar{v}_0\bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j'\bar{v}_0\bar{v}_1 \cdots \bar{u}_{k+1}'\bar{v}_0\bar{u}_k''\bar{g}_4''\bar{v}_1 \ .$$

If f_9 is unbordered, then $|u| < |w| - 1$ since $|f_9| \leq |w|$ and

$$|u| = |f_9| - |\bar{v}_0\bar{v}_1\bar{w}_k'\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{w}_2| + |g_4''\bar{v}_0\bar{v}_1 v_{k-1}'\bar{v}_0\bar{v}_1 \cdots v_1'\bar{v}_0\bar{v}_1\bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_9 is bordered with a shortest border h_9 then $\bar{v}_0\bar{v}_1 \leq_\mathrm{p} h_9$. If $|h_9| > |\bar{v}_0\bar{u}''_k\bar{g}''_4\bar{v}_1|$ then the same argument as in the case $|\bar{v}_1\bar{w}'_k\bar{v}_0| < |h_4|$ above shows that $|u| < |w| - 1$, that is, we have an increasing chain of factors longer than f_9 with a corresponding increasing chain of shortest borders longer than h_9 until $|u| < |w| - 1$ is shown. If $|h_9| < |\bar{v}_0\bar{u}''_k\bar{g}''_4\bar{v}_1|$ then $\bar{v}_0\bar{v}_1$ occurs in $\bar{u}'_k$; a contradiction. The remaining case is $h_8 = \bar{v}_0\bar{u}''_k\bar{g}''_4\bar{v}_1$ with $h_9 \leq_\mathrm{p} \bar{v}_0\bar{v}_1\bar{w}'_k$. It follows $\bar{u}''_k\bar{g}''_4\bar{v}_1 \leq_\mathrm{p} \bar{v}_1 g_4$ and $\bar{u}''_k\bar{g}''_4\bar{v}_1 g''_4 = \bar{u}''_k\bar{u}''_k\bar{g}''_4\bar{v}_1 g''_4$ by the choice of g''_4, and we have $\bar{u}''_k = \varepsilon$. We have $h_4 = \bar{v}_0 g_4\bar{v}_1 = \bar{v}_0\bar{w}'_k\bar{v}_1 = \bar{v}_0\bar{u}'_k\bar{v}_1$ and $\bar{w}'_k = \bar{u}'_k$. This proves Claim 4.4.

Let

$$\bar{v} = \bar{v}_0\bar{v}_1\bar{w}'_\iota \cdots \bar{v}_0\bar{v}_1\bar{w}'_2\bar{v}_0\bar{v}_1\bar{w}'_1 = \bar{v}_0\bar{v}_1\bar{u}'_\iota \cdots \bar{v}_0\bar{v}_1\bar{u}'_2\bar{v}_0\bar{v}_1\bar{u}'_1$$

where $\iota = \min\{i, j\}$.

Next, we will show that $|\bar{w}'_0| \leq |\bar{u}'_0|$ (Claim 4.5) which will be used later in proving the Claims 4.6 and 4.8.

Claim 4.5. *It holds that* $|\bar{w}'_0| \leq |\bar{u}'_0|$.

Recall that w_0 and u_0 are such that $w_0bz \leq_\mathrm{p} w$ and $u_0bz \leq_\mathrm{p} u$ and bz occurs only once in w_0bz and u_0bz, respectively. Note that $w_0 \leq_\mathrm{p} \bar{w}'_0$ and $u_0 \leq_\mathrm{p} \bar{u}'_0$. Consider

$$f_0 = wu_0bz$$

which is bordered with a shortest border $\bar{h}_0$. We have $|z| < |\bar{h}_0| \leq |u_0bz|$ otherwise w is bordered. Hence, $bz \leq_\mathrm{s} \bar{h}_0$ and $|\bar{h}_0| \geq |w_0bz|$. Actually, $|\bar{h}_0| = |w_0bz|$ since bz occurs only once in u_0bz, and we have $w_0 \leq_\mathrm{s} u_0$.

Let $\hat{w}$ and $\hat{u}$ be such that $w_0\hat{w}v_0v_1 \leq_\mathrm{p} w$ and $u_0\hat{u}v_0v_1 \leq_\mathrm{p} u$ and v_0v_1 occurs only once in $w_0\hat{w}v_0v_1$ and $u_0\hat{u}v_0v_1$, respectively. Note that $w_0\hat{w} = \bar{w}'_0$ and $u_0\hat{u} = \bar{u}'_0$, if $\bar{v}_0\bar{v}_1 = v_0v_1$. Let $\hat{w}v_0v_1 = bz\hat{w}'$ and $\hat{u}v_0v_1 = bz\hat{u}'$, and let $w = w_0\hat{w}v_0v_1v'$. Consider

$$f'_0 = bz\hat{w}'v'u_0\hat{u}v_0v_1$$

which is bordered with a shortest border $\bar{h}'_0$ (since $|w_0| \leq |u_0|$). We have that $|\bar{h}'_0| > |bz|$ since $bz \leq_\mathrm{p} v_0v_1$ and v_0v_1 is unbordered. Moreover, $|\bar{h}'_0| \geq |v_0v_1|$ since $v_0 \leq_\mathrm{p} bz$ and v_0 and v_1 do not intersect. If $|\bar{h}'_0| > |u_0\hat{u}v_0v_1|$ then also $|\bar{h}'_0| > |azu_0\hat{u}v_0v_1|$ since there is no suffix z' of z such that $z' \leq_\mathrm{p} u$ otherwise wz' is unbordered or w is bordered; a contradiction in any case. If $|\bar{h}'_0| > |azu_0\hat{u}v_0v_1|$ then az occurs twice (nonoverlapping) in w. Let $w = \hat{w}_1az\hat{w}_2az$. But then, $az\hat{w}_2azu'bz_0$ is unbordered (see Lemma 3.2 where $z' = z_0$ and $f = \hat{w}_2azu'bzr$ and $g = zu'b$ and $h = r$) and $|u| < |w| - 1$ since $|zr| < |v_0v_1t'| \leq |v_0v_1| + |z|$ and

$$\begin{aligned}
|u| &\leq |\hat{w}_1| + |zr| \\
&< |\hat{w}_1| + |v_0v_1| + |z| \\
&< |\hat{w}_1| + |v_0v_1| + |az| \\
&\leq |w| \ .
\end{aligned}$$

Assume that $|\bar{h}'_0| \leq |u_0\hat{u}v_0v_1|$. We have $|\bar{h}'_0| \leq |\hat{u}v_0v_1|$ since $bz \leq_\mathrm{p} \bar{h}'_0$. Moreover, $|\bar{h}'_0| \geq |\hat{w}v_0v_1|$, since $v_0v_1 \leq_\mathrm{s} \bar{h}'_0$, and $|\bar{h}'_0| \leq |\hat{w}v_0v_1|$, since v_0v_1 occurs only once in $\hat{u}v_0v_1$. Hence, $\bar{h}'_0 = \hat{w}v_0v_1 = bz\hat{w}' \leq_\mathrm{s} \hat{u}v_0v_1$. Note that we are done if $v_0v_1 = \bar{v}_0\bar{v}_1$.

Assume $v_0v_1 \neq \bar{v}_0\bar{v}_1$. Recall that $w = w_0\hat{w}v_0v_1v'$, and let $\hat{w}_0$ and $\hat{u}_0$ be such that $w_0\hat{w}\hat{w}_0 = \bar{w}'_0$ and $u_0\hat{u}\hat{u}_0 = \bar{u}'_0$. Consider

$$f''_0 = v_0v_1v'\bar{u}'_0\bar{v}_0\bar{v}_1$$

which is bordered with a shortest border $\bar{h}''_0$ (since $|w_0\hat{w}| \leq |u_0\hat{u}|$). We assume that $|\bar{h}''_0| \leq |\bar{w}'_0\bar{v}_0\bar{v}_1|$ otherwise $|\bar{h}''_0| \geq |az\bar{w}'_0\bar{v}_0\bar{v}_1|$ (since $bz \leq_\mathrm{p} \bar{h}''_0$) and az has at least two (nonoverlapping) occurrences in w and $|u| < |w| - 1$ follows by the argument in the previous paragraph. We also have that $|\bar{h}''_0| > |v_0v_1|$ since $v_0v_1 \leq_\mathrm{p} \bar{v}_0$ and $\bar{v}_0\bar{v}_1$ is unbordered. If $|\bar{h}''_0| \geq |\bar{v}_0\bar{v}_1|$ then $\bar{h}''_0 = \hat{w}_0\bar{v}_0\bar{v}_1 \leq_\mathrm{s} \hat{u}_0\bar{v}_0\bar{v}_1$ since v_0v_1 does occur only once in $u_0\hat{u}v_0v_1$, and we are done. Assume $|v_0v_1| < |\bar{h}''_0| < |\bar{v}_0\bar{v}_1|$. Then $|\bar{v}_1| < |\bar{h}''_0|$ because v_0v_1 does not occur in $\bar{v}_1$. We have $\bar{h}''_0 \leq_\mathrm{p} \hat{w}_0$ since $\bar{v}_1 \leq_\mathrm{s} \bar{h}''_0$ and $\bar{v}_0$ and $\bar{v}_1$ do not intersect. Let $\bar{h}''_0 = \hat{h}''_0\bar{v}_1$ and $\hat{w}_0\bar{v}_0 = \bar{h}''_0\hat{w}_1\bar{v}_0\hat{w}_2$ where $\bar{v}_0$ occurs only once in $\hat{w}_1\bar{v}_0$. Note that $\bar{v}_0$ also occurs only once in $\bar{h}''_0\hat{w}_1\bar{v}_0$ since $|\bar{v}_0| > |v_0v_1|$ and $\bar{v}_0$ and $\bar{v}_1$

do not intersect. Let $\hat{u}_0\bar{v}_0 = \hat{u}_1\bar{v}_0\hat{u}_2$ such that $\bar{v}_0$ occurs only once in $\hat{u}_1\bar{v}_0$. Note that $\bar{v}_0$ also occurs only once in $u_0\hat{u}\hat{u}_1\bar{v}_0$. Consider

$$f_1'' = \bar{v}_1\hat{w}_1\bar{v}_0\hat{w}_2\bar{v}_1\bar{w}_i'\bar{v}_0\bar{v}_1\cdots\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2u_0\hat{u}\hat{u}_1\bar{v}_0$$

which is bordered by a shortest border $\bar{h}_1''$ (since $|w_0\hat{w}\hat{h}_0''| < |u_0\hat{u}\hat{u}_1\bar{v}_0|$). We have that $|\bar{v}_0\bar{v}_1| \leq |\bar{h}_1''| \leq |\hat{u}_1\bar{v}_0|$ since $\bar{v}_0$ and $\bar{v}_1$ do not intersect. It follows that $\hat{w}_1 = \hat{u}_1$ since $\bar{v}_0$ does occur only once in $\hat{w}_1\bar{v}_0$. Consider then

$$f_2'' = \bar{v}_0\hat{w}_2\bar{v}_1\bar{w}_i'\bar{v}_0\bar{v}_1\cdots\bar{w}_1'\bar{v}_0\bar{v}_1\bar{w}_2u_0\hat{u}\hat{u}_1\bar{v}_0\hat{u}_2\bar{v}_1$$

which is bordered by a shortest border $\bar{h}_2''$ (since $|w_0\hat{w}\hat{w}_0| < |u_0\hat{u}\hat{u}_1\bar{v}_0\hat{u}_2\bar{v}_1|$). We have that $|\bar{h}_2''| \geq |\bar{v}_0\bar{v}_1|$ since $\bar{v}_0$ and $\bar{v}_1$ do not intersect. Moreover, $\hat{w}_2 = \hat{u}_2$ since $\bar{v}_0\bar{v}_1$ occurs only once in $\bar{v}_0\hat{w}_2\bar{v}_1$ and $\bar{v}_0$ occurs only once in $u_0\hat{u}\hat{u}_1\bar{v}_0$, and we are done. This proves Claim 4.5.

Claim 4.6. *If $i < j$ then $|u| < |w| - 1$.*

We have that

$$|\bar{w}_0'| < |\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j'\cdots\bar{v}_0\bar{v}_1\bar{u}_{i+1}'| \tag{8}$$

since $|\bar{w}_0'| \leq |\bar{u}_0'|$ by Claim 4.5. Let

$$f_{11} = \bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j'\cdots\bar{v}_0\bar{v}_1\bar{u}_{i+1}'\bar{v}\bar{v}_0 \ .$$

Note that $w = \bar{w}_0'\hat{v}\bar{v}_0\bar{v}_1\bar{w}_2$. Then $|w| < |f_{11}|$ by (8), and hence, f_{11} is bordered. Let $h_{11} = \bar{v}_1g_{11}\bar{v}_0$ be the shortest border of f_{11}. Recall, that $\bar{w}_2 \neq \varepsilon$ and either $az \leq_{\mathrm{s}} \bar{v}_1\bar{w}_2$ or $\bar{v}_1\bar{w}_2 \leq_{\mathrm{s}} az$. If $|\bar{v}_1\bar{w}_2| < |az|$ then $\bar{v}_1$ necessarily occurs in z, and hence, it intersects with $\bar{v}_0$ (since $bz \leq_{\mathrm{p}} \bar{v}_0\bar{v}_1$ and $\bar{w}_2 \neq \varepsilon$); a contradiction. We have $az \leq_{\mathrm{s}} \bar{v}_1\bar{w}_2$. Surely, $|h_{11}| < |\bar{v}_1\bar{w}_2|$ (and so $h_{11} \leq_{\mathrm{p}} \bar{v}_1\bar{w}_2$) for otherwise az occurs in u which contradicts our assumption that z is of maximum length. Let $\bar{w}_2 = g_{11}\bar{v}_0\bar{w}_5$. Then $|\bar{t}'| < |az| < |\bar{v}_0\bar{w}_5|$ by condition (7) (page 8). Consider,

$$f_{12} = \bar{v}_0\bar{w}_5\bar{u}_0'\bar{v}_0\bar{v}_1\bar{u}_j'\cdots\bar{v}_0\bar{v}_1\bar{u}_{i+1}'\bar{v}\bar{v}_0\bar{v}_1 \ .$$

If f_{12} is unbordered, then $|u| < |w| - 1$ since $|f_{12}| \leq |w|$ and

$$|u| = |f_{12}| - |\bar{v}_0\bar{w}_5| + |\bar{t}'|$$

and $|\bar{t}'| < |az| < |\bar{v}_0\bar{w}_5|$. Assume f_{12} is bordered. Then f_{12} has a shortest border $h_{12} = g_{12}\bar{v}_0\bar{v}_1$ for some g_{12} (since $\bar{v}_0$ and $\bar{v}_1$ do not intersect) and $h_{12} \leq_{\mathrm{p}} \bar{v}_0\bar{w}_5$ (otherwise az occurs in h_{12} and also in u; a contradiction). But now, $\bar{v}_0\bar{v}_1$ occurs in $\bar{w}_2$; a contradiction. This proves Claim 4.6.

Claim 4.7. *If $i > j$ then $|u| < |w| - 1$.*

We have

$$w = \bar{w}_0'\bar{v}_0\bar{v}_1\bar{w}_i'\cdots\bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2 \qquad \text{and} \qquad u = \bar{u}_0'\bar{v}\bar{v}_0\bar{v}_1\bar{t}' \ .$$

Consider

$$f_{13} = \bar{v}_1\bar{w}_i'\cdots\bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}_0 \ .$$

If f_{13} is unbordered, then $|u| < |w| - 1$ since $|f_{13}| \leq |w|$ and

$$|u| = |f_{13}| - |\bar{v}_1\bar{w}_i'\cdots\bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2| - |\bar{v}_0| + |\bar{v}\bar{v}_0\bar{v}_1\bar{t}'| \leq |\bar{w}_0'| + |\bar{v}\bar{v}_0\bar{v}_1\bar{t}'|$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_{13} is bordered. Then f_{13} has a shortest border $h_{13} = \bar{v}_1g_{13}\bar{v}_0$.

If $\bar{v}_0\bar{v}_1$ occurs in h_{13}, then also az has to occur in h_{13} since $bz \leq_{\mathrm{p}} \bar{v}_0\bar{v}_1$ and $\bar{v}_0\bar{v}_1$ does not occur in $az\bar{u}_0'\bar{v}_0$. It follows that az occurs twice non-overlapping in w (since h_{13} does not overlap itself because it is the shortest border of f_{13}). Let $w = \hat{w}_1az\hat{w}_2az$. But then, $az\hat{w}_2azu'bz_0$ is unbordered (see Lemma 3.2 where $z' = z_0$ and $f = \hat{w}_2azu'bzr$ and $g = zu'b$ and $h = r$) and $|u| < |w| - 1$ since $|zr| < |v_0v_1t'| \leq |v_0v_1| + |z|$ and

$$\begin{aligned} |u| &\leq |\hat{w}_1| + |zr| \\ &< |\hat{w}_1| + |v_0v_1| + |z| \\ &< |\hat{w}_1| + |v_0v_1| + |az| \\ &\leq |w| \ . \end{aligned}$$

Assume that $\bar{v}_0\bar{v}_1$ does not occur in h_{13}.

If $|g_{13}| \geq |\bar{u}_0'|$ then $|u| < |w| - 1$ since

$$\begin{aligned}
|u| &\leq |g_{13}\bar{v}\bar{v}_0\bar{v}_1\bar{t}'| \\
&\leq |\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{t}'| \\
&< |\bar{v}_0\bar{v}_1| + |\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1| \\
&< |\bar{w}_0'| + |\bar{v}_0\bar{v}_1| + |\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1| + |\bar{w}_2| \\
&= |w|
\end{aligned}$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume $|g_{13}| < |\bar{u}_0'|$.

Let $u = \hat{u}g_{13}\bar{v}\bar{v}_0\bar{v}_1\bar{t}'$. Note that $\bar{v}_1 \leq_{\mathrm{s}} w\hat{u}$. Consider

$$f_{14} = \bar{v}_0\bar{v}_1\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2\hat{u}\ .$$

If f_{14} is unbordered, then $|u| < |w| - 1$ since $|f_{14}| \leq |w|$ and

$$\begin{aligned}
|u| &= |f_{14}| - |\bar{v}_0\bar{v}_1\bar{w}_i' \cdots \bar{v}_0\bar{v}_1\bar{w}_{j+1}'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2| + |g_{13}\bar{v}\bar{v}_0\bar{v}_1\bar{t}'| \\
&\leq |\bar{w}_0'| + |g_{13}\bar{v}\bar{v}_0\bar{v}_1\bar{t}'| \\
&< |\bar{w}_0'| + |\bar{v}_0\bar{v}_1| + |g_{13}\bar{v}\bar{v}_0\bar{v}_1| \\
&< |\bar{w}_0'| + |\bar{v}_0\bar{v}_1| + |g_{13}\bar{v}\bar{v}_0\bar{v}_1| + |\hat{w}_2| \\
&\leq |w|
\end{aligned}$$

and $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $\bar{w}_2 \neq \varepsilon$. Assume f_{14} is bordered. Then f_{14} has a shortest border $h_{14} = \bar{v}_0g_{14}\bar{v}_1$ and $\bar{v}_0\bar{v}_1 \leq_{\mathrm{p}} h_{14}$ since $\bar{v}_0$ and $\bar{v}_1$ do not intersect. However, $\bar{v}_0\bar{v}_1$ can only occur in h_{14} if also az occurs in h_{14} (again, since $\bar{v}_0\bar{v}_1$ does not occur in $az\bar{u}_0'\bar{v}_0$), and $|u| < |w| - 1$ follows by the same arguments as in the case of h_{13} above. This proves Claim 4.7.

Claim 4.8. *If $i = j$ then $|u| < |w| - 1$.*

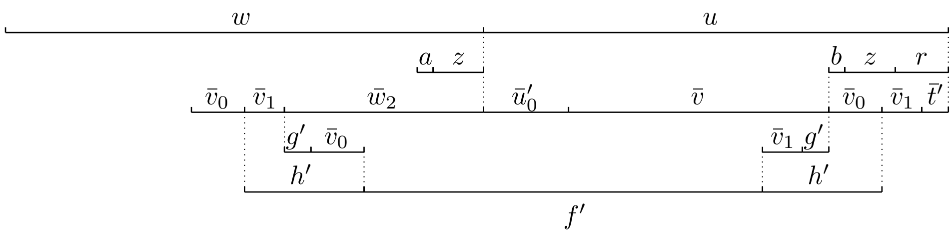
We have

$$w = \bar{w}_0'\bar{v}\bar{v}_0\bar{v}_1\bar{w}_2 \qquad \text{and} \qquad u = \bar{u}_0'\bar{v}\bar{v}_0\bar{v}_1\bar{t}'\ .$$

Consider

$$f' = \bar{v}_1\bar{w}_2\bar{u}_0'\bar{v}\bar{v}_0\ .$$

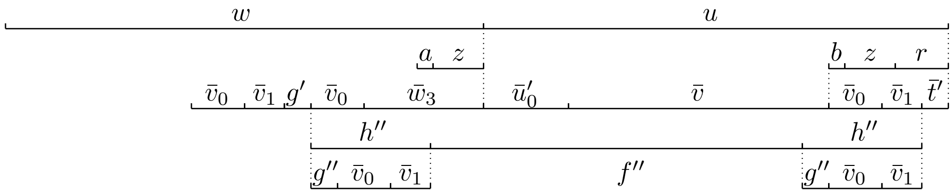
Case: Assume that f' is bordered. Then f' has a shortest border $h' = \bar{v}_1g'\bar{v}_0$.



Recall, that $\bar{w}_2 \neq \varepsilon$ and either $az \leq_{\mathrm{s}} \bar{v}_1\bar{w}_2$ or $\bar{v}_1\bar{w}_2 \leq_{\mathrm{s}} az$. If $|\bar{v}_1\bar{w}_2| < |az|$ then $\bar{v}_1$ occurs in z, and hence, intersects with $\bar{v}_0$ since $bz \leq_{\mathrm{p}} \bar{v}_0\bar{v}_1$; a contradiction. We have $az \leq_{\mathrm{s}} \bar{v}_1\bar{w}_2$. It follows that $|h'| < |\bar{v}_1\bar{w}_2|$ otherwise az occurs in u contradicting the maximality of z. Let $\bar{w}_2 = g'\bar{v}_0\bar{w}_3$ and consider

$$f'' = \bar{v}_0\bar{w}_3\bar{u}_0'\bar{v}\bar{v}_0\bar{v}_1\ .$$

Subcase: Assume that f'' is unbordered. Then it follows that $|u| < |w| - 1$ since we have $|\bar{t}'| < |az|$ by condition (7) (page 8) and $|az| < |\bar{v}_0\bar{w}_3|$.



Subcase: Assume then that f'' is bordered. Then it has a shortest border $h'' = g''\bar{v}_0\bar{v}_1$ with $|az| < |h''|$, for otherwise az occurs in u. But then $\bar{v}_0\bar{v}_1$ occurs in $\bar{w}_2$; a contradiction.

Case: Assume that f' is unbordered. Then $|f'| \leq |w|$, and hence, $|\bar{w}'_0| \geq |\bar{u}'_0|$. But, we also have $|\bar{w}'_0| \leq |\bar{u}'_0|$; see Claim 4.5. That implies $|\bar{w}'_0| = |\bar{u}'_0|$. Moreover, the factors $\bar{w}_0$ and $bz\hat{w}'$ have both nonoverlapping occurrences in $\bar{u}'_0\bar{v}_0\bar{v}_1$ by Claim 4.5. Therefore, $\bar{w}'_0 = \bar{u}'_0$. Let, $w = xa\bar{w}_6$ and $u = xb\bar{t}''$, where $\bar{w}'_0\bar{v}\bar{v}_0\bar{v}_1 \leq_{\mathrm{p}} x$ and a and b are different letters and $\bar{w}_6 \leq_{\mathrm{s}} \bar{w}_2$ and $\bar{t}'' \leq_{\mathrm{s}} \bar{t}'$. We have that xb occurs in w by Theorem 3.7. Since xb is not a prefix of w and $\bar{v}_0\bar{v}_1$ does not overlap itself and az does not occur in u, we have $|xb| + |\bar{v}_0\bar{v}_1| < |w|$. From $|\bar{t}'| < |\bar{v}_0\bar{v}_1|$ and $|\bar{t}''| \leq |\bar{t}'|$ follows $|u| < |w| - 1$.

This proves the Claims 4.8 and 4.2 and finishes the proof of Theorem 1.2. ☐

5 Corollary

Note that the bound $|u| < |w| - 1$ on the length of a nontrivial Duval extension wu of w is tight, as the following example shows.

Example 5.1. *Let $w = a^n ba^{n+m} bb$ and $u = a^{n+m} ba^n$ with $n, m \geq 1$. Then*

$$w.u = a^n ba^{n+m} bb.a^{n+m} ba^n$$

is a nontrivial Duval extension of w and $|u| = |w| - 2$.

In general, Duval (Duval 1982) proved that we have $\partial(w) = \mu(w)$, for any word w, if $|w| \geq 4\mu(w) - 6$. Duval also noted that already $|w| \geq 3\mu(w)$ implies $\partial(w) = \mu(w)$, provided his conjecture holds. Corollary 1.3 follows from Theorem 1.2.

Corollary 1.3. *If $|w| \geq 3\mu(w) - 3$ then $\partial(w) = \mu(w)$.*

Proof. Assume $\partial(w) \neq \mu(w)$ and $|w| \geq 3\mu(w) - 3$. Let $w = xvy$ such that v is the leftmost unbordered factor of w of maximum length, that is, $|v| = \mu(w)$ and $\mu(xv^\bullet) < \mu(xv)$. Then $\widetilde{xv}$ and vy are Duval extensions of $\widetilde{v}$ and v, respectively. We have by our assumption that $|x|$ or $|y|$ is larger than $\mu(w) - 2$.

If $|x| \geq \mu(w) - 1$ then $\widetilde{xv}$ is a trivial Duval extension by Theorem 1.2 and all conjugates of v occur in xv. Since v is primitive and only alphabets with at least two letters are considered there occurs an unbordered conjugate u of v in $xv^\bullet$ by Lemma 3.1 contradicting our assumption that v is the leftmost unbordered factor of w of maximum length.

If $|y| \geq \mu(w) - 1$ then vy is a trivial Duval extension by Theorem 1.2 and all conjugates of v occur in vy. Moreover, all conjugates of v occur in the suffix y' of vy of length $2|v| - 1$. Let u be an unbordered conjugate of v, with $u \neq v$ (which exists since we consider only words with at least two different letters), occurring in y', that is $w = sut$ with $|t| \leq |v| - 2$. Consider the Duval extension $\widetilde{su}$. If $\widetilde{su}$ is trivial than $\partial(w) = \mu(w)$ contradicting our assumption. So, $\widetilde{su}$ is a nontrivial Duval extension, and hence, $|s| < |v| - 1$ by Theorem 1.2. Now, $|w| < 3|v| - 3$ which is again a contradiction. ☐

However, this bound is unlikely to be tight. The best example for a large bound that we could find is taken from (Assous and Pouzet 1979).

Example 5.2. *Let*

$$w = a^n ba^{n+1} ba^n ba^{n+2} ba^n ba^{n+1} ba^n \ .$$

We have $|w| = 7n + 10$ and $\mu(w) = 3n + 6$ and $\partial(w) = 4n + 7$.

Example 5.2 shows that the precise bound for the length of a word that implies $\partial(w) = \mu(w)$ is larger than $(7/3)\mu(w) - 4$ and not larger than $3\mu(w) - 3$ (by Corollary 1.3). The characterization of the precise bound of the length of a word as a function of its longest unbordered factor is still an open problem.

6 Conclusions

In this paper we have given an affirmative answer to a long standing conjecture (Duval 1982) by proving that a Duval extension wu of w longer than $2|w| - 2$ is trivial. This bound is tight and also gives a new bound on the relation between the length of an arbitrary word w and its longest unbordered factors $\mu(w)$, namely that $|w| \geq 3\mu(w) - 3$ implies $\partial(w) = \mu(w)$ as conjectured (more weakly) in (Assous and Pouzet 1979). However, the best known example of a word w satisfying $\partial(w) > \mu(w)$ gives $|w| = (7/3)\mu(w) - 4$. We believe that the actual bound of $|w|$ is indeed close to $(7/3)\mu(w)$ rather than $3\mu(w)$. We pose the following conjecture.

Conjecture 6.1. *If* $|w| \geq \frac{7}{3}\mu(w) - 3$ *then* $\partial(w) = \mu(w)$.

Certainly, more information about the structure of nontrivial Duval extensions, like the one described in (Duval, Harju, and Nowotka), would be useful for solving Conjecture 6.1.

Acknowledgement

The authors would like to thank the anonymous referees for the time and effort that they put into the review of this manuscript and their detailed comments and suggestions which greatly helped to improve this paper.

References

Assous, R. and M. Pouzet (1979). Une caractérisation des mots périodiques. *Discrete Math. 25*(1), 1–5.

Berstel, J. and D. Perrin (1985). *Theory of codes*, Volume 117 of *Pure and Applied Mathematics*. Orlando, FL: Academic Press Inc.

Boyer, R. S. and J. S. Moore (1977, October). A fast string searching algorithm. *Commun. ACM 20*(10), 762–772.

Breslauer, D., T. Jiang, and Z. Jiang (1997). Rotations of periodic strings and short superstrings. *J. Algorithms 24*(2), 340–353.

Bylanski, P. and D. G. W. Ingram (1980). *Digital transmission systems*. IEE.

Césari, Y. and M. Vincent (1978). Une caractérisation des mots périodiques. *C. R. Acad. Sci. Paris Sér. A 286*, 1175–1177.

Crochemore, M., F. Mignosi, A. Restivo, and S. Salemi (1999). Text compression using antidictionaries. In *26th Internationale Colloquium on Automata, Languages and Programming (ICALP), Prague*, Volume 1644 of *Lecture Notes in Comput. Sci.*, pp. 261–270. Berlin: Springer.

Crochemore, M. and D. Perrin (1991). Two-way string-matching. *J. ACM 38*(3), 651–675.

Duval, J.-P. (1979). Périodes et répétitions des mots de monoïde libre. *Theoret. Comput. Sci. 9*(1), 17–26.

Duval, J.-P. (1982). Relationship between the period of a finite word and the length of its unbordered segments. *Discrete Math. 40*(1), 31–44.

Duval, J.-P., T. Harju, and D. Nowotka. Unbordered factors and Lyndon words. *Discrete Math.*. to appear, see also (Harju and Nowotka 2002b).

Ehrenfeucht, A. and D. M. Silberger (1979). Periodicity and unbordered segments of words. *Discrete Math. 26*(2), 101–109.

Harju, T. and D. Nowotka (2002a). Density of critical factorizations. *Theor. Inform. Appl. 36*(3), 315–327.

Harju, T. and D. Nowotka (2002b). Unbordered factors and Lyndon words. technical report 479, Turku Centre of Computer Science (TUCS), Turku, Finland.

Harju, T. and D. Nowotka (2004). Minimal Duval extensions. *Internat. J. Found. Comput. Sci. 15*(2), 349–354.

Holub, S. (2005). A proof of the extended Duval's conjecture. *Theoret. Comput. Sci. 339*(1), 61–67.

Knuth, D. E., J. H. Morris, and V. R. Pratt (1977). Fast pattern matching in strings. *SIAM J. Comput. 6*(2), 323–350.

Lothaire, M. (1983). *Combinatorics on Words*, Volume 17 of *Encyclopedia of Mathematics*. Reading, MA: Addison-Wesley.

Lothaire, M. (2002). *Algebraic Combinatorics on Words*, Volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge, United Kingdom: Cambridge University Press.

Lyndon, R. C. (1954). On Burnside's problem. *Trans. Amer. Math. Soc. 77*, 202–215.

Margaritis, D. and S. Skiena (1995). Reconstructing strings from substrings in rounds. In *36th Annual Symposium on Foundations of Computer Science (FOCS)*, Milwaukee, WI, pp. 613–620. IEEE Computer Society.

Mignosi, F. and L. Q. Zamboni (2002). A note on a conjecture of Duval and Sturmian words. *Theor. Inform. Appl. 36*(1), 1–3.

Morse, M. and G. A. Hedlund (1940). Symbolic dynamics II: Sturmian trajectories. *Amer. J. Math. 61*, 1–42.

Schützenberger, M.-P. (1979). A property of finitely generated submonoids of free monoids. In *Algebraic theory of semigroups (Proc. Sixth Algebraic Conf., Szeged, 1976)*, Volume 20 of *Colloq. Math. Soc. János Bolyai*, pp. 545–576. Amsterdam: North-Holland.

Ziv, J. and A. Lempel (1977). A universal algorithm for sequential data compression. *IEEE Trans. Information Theory 23*(3), 337–343.