RAIRO-Theor. Inf. Appl. 48 (2014) 149–171
DOI: 10.1051/ita/2013044 Available online at:
www.rairo-ita.org

ONE-RULE LENGTH-PRESERVING REWRITE SYSTEMS AND RATIONAL TRANSDUCTIONS

Michel Latteux[1] and Yves Roos[1]

Abstract. We address the problem to know whether the relation induced by a one-rule length-preserving rewrite system is rational. We partially answer to a conjecture of Éric Lilin who conjectured in 1991 that a one-rule length-preserving rewrite system is a rational transduction if and only if the left-hand side u and the right-hand side v of the rule of the system are not quasi-conjugate or are equal, that means if u and v are distinct, there do not exist words x, y and z such that $u = xyz$ and $v = zyx$. We prove the *only if* part of this conjecture and identify two non trivial cases where the *if* part is satisfied.

Mathematics Subject Classification. 68Q45, 68Q42, 68R15.

1. Introduction

Rewrite systems are of primordial interest for computational problems. Mainly, the problems that are investigated for rewrite systems are the accessibility problem, the common descendant problem, the confluence problem, the termination and uniform termination problem. For several years they have been intensively studied and several deep results have been obtained. However some intriguing decidability problems remain open even for very simple rewrite systems[2]. The most known among these problems is certainly the decidability of the termination of one-rule rewrite systems, a question that remains open for more than twenty years.

One-rule length-preserving rewrite systems are among the simplest rewriting systems. Indeed, they are defined by two words u, v over an alphabet A, with

Keywords and phrases. String rewriting - rationality.

[1] Laboratoire d'Informatique Fondamentale de Lille, Université Lille 1, France.
yves.roos@lifl.fr

[2]See [3] for a history of these problems and the attempts to solve them.

Article published by EDP Sciences

$|u| = |v|$, and noted $S = \{u \to v\}$. For a word w, $S(w)$ is the set of words obtainable from w by replacing repeatedly u by v. Thus S induces a relation over A^* and we address here the problem to decide whether this relation is rational. Abusing notations, the question is: *given a one-rule length-preserving rewrite system S, is S a rational transduction?*

One of the simplest one-rule length-preserving rewrite system that is not a rational transduction is the system $S_0 = \{ba \to ab\}$ where a and b are two distinct letters. Indeed $S_0((ba)^*) \cap a^*b^* = \{a^n b^n \mid n \geq 0\}$. On the other hand, the system $S_1 = \{abb \to baa\}$ is a rational transduction. Indeed, it can be proved[3] that, for all word w, $S_1(w)$ can be computed in two passes, the first one reading w from left to right and the second one reading w from right to left, each pass substituting non-deterministically the occurrences of abb with baa. For this, we say that S_1 is *left-right*. Nevertheless, there exist one-rule length-preserving rewrite systems that define a rational transduction but are neither left-right nor right-left.

In 1991, Éric Lilin has proposed in [9] the following conjecture: *a one-rule rewrite system $S = \{u \to v\}$ with $|u| = |v|$ and $u \neq v$ is a rational transduction if and only if the words u and v are not quasi-conjugate.* Two words u and v are quasi-conjugate if there exist some words x, y and z such that $u = xyz$ and $v = zyx$. If this conjecture is true, it would give a nice decidable characterization of one-rule length-preserving rewrite systems that define a rational transduction. In this paper we prove the *only if* part of the conjecture and, conversely, we consider two cases for which the *if* part is satisfied. These cases are based on the kind of overlaps that exist between the left-hand side u and the right-hand side v of the rewriting rule. As for the termination problem (see for instance [5]) and the confluence problem [16] for one-rule rewrite systems, the study of these overlaps plays a central role. More precisely, for a one-rule length-preserving rewrite system $S = \{u \to v\}$, we consider two sets: the set X of overlaps of the left factors of u with the right factors of v and the set Z of overlaps of the right factors of u with the left factors of v. The members of $X \times Z$ are called *pairs of overlaps*. The different cases that we consider for the *if* part of the conjecture depends on the presence of *short* pairs of overlaps, that are pairs of overlaps (x, z) such that $|xz| \leq |u|$ and the presence of *large* pairs of overlaps, that are pairs of overlaps (x, z) such that $|xz| > |u|$.

The paper is organized as follows: first, we define the quasi-conjugacy relation and prove some of its properties. Next we prove that if a one-rule length-preserving rewrite system $S = \{u \to v\}$ with $u \neq v$ is a rational transduction then the words u and v are not quasi-conjugate. Then, in Section 4, we give a sufficient condition which ensures that a one-rule length-preserving rewrite system is left-right or right-left and so is a rational transduction. The two following sections are devoted to some cases for which the *if* part of the conjecture is satisfied: when the system S has no large pair of overlaps and when the system S has no short pair of overlaps. At last, in the conclusion, we identify some problems that deserve to be studied in the context of the rationality of one-rule length-preserving rewrite systems.

[3]As a matter of fact, it is a consequence of a result given in Section 4.

2. Preliminaries and notations

In the following, A will denote a finite alphabet, A^* the free monoid over A and ε the empty word in A^*. A word $w' \in A^*$ is a *factor* of a word $w \in A^*$ if there exist some words x and y in A^* such that $w = xw'y$. We denote by $\mathrm{F}(w)$ the set of the factors of the word w. We denote by $\mathrm{RF}(w)$ (respectively $\mathrm{LF}(w)$) the set of *right factors* (respectively *left factors*) of the word w, that is:

$$\mathrm{RF}(w) = \{w' \in A^* \mid \exists w'' \in A^*, w = w''w'\},$$

$$\mathrm{LF}(w) = \{w' \in A^* \mid \exists w'' \in A^*,\ w = w'w''\}.$$

A word w' is a *subword* of w if there exist some words $x_0, x_1, \ldots x_n$ and $y_0, \ldots y_{n-1}$ in A^* with $n > 0$ such that $w = x_0y_0 \ldots x_{n-1}y_{n-1}x_n$ and $w' = y_0 \ldots y_{n-1}$. For a word $w \in A^*$, $|w|$ denotes the length of the word w. Two words u and v are *conjugate* if there exist words x and y such that $u = xy$ and $v = yx$. It is well known that two words u and v are conjugate if and only if there exists a word z such that $uz = zv$.

A rewrite system over an alphabet A is a subset $S \subseteq A^* \times A^*$. Members of S are denoted $u \underset{S}{\longrightarrow} v$ (or $u \rightarrow v$ if there is no ambiguity). We shall denote S^{-1} the system obtained from the system S by reversing the rules of S, that is $u \rightarrow v \in S$ iff $v \rightarrow u \in S^{-1}$. One-step derivation, denoted $\underset{S}{\longrightarrow}$ ($\rightarrow$ if no ambiguity), is the binary relation over words defined by: $\forall w, w' \in A^*, w \rightarrow w'$ iff there exists $u \rightarrow v \in S$ and $\alpha, \beta \in A^*$ such that $w = \alpha u\beta$ and $w' = \alpha v\beta$. The relation $\xrightarrow{*}$ (resp. $\xrightarrow{+}$), called *derivation relation*, is the reflexive and transitive closure (resp. transitive closure) of the relation $\rightarrow$ and, for any word $w \in A^*$, we shall denote $S(w)$ the set $S(w) = \{w' \in A^* \mid w \underset{S}{\xrightarrow{*}} w'\}$. We extend these notations to languages: for any language $L \subseteq A^*$, $S(L) = \cup_{w\in L}S(w)$. For a derivation $w = w_0 \rightarrow w_1 \cdots \rightarrow w_n = w'$, n is called the length of the derivation that will be denoted by $w \xrightarrow{n} w'$. Abusing notation we shall identify in the following a given rewrite system S with its associated transformation over languages. The main question that is studied in this paper is: *given a one-rule length-preserving rewrite system S, is S a rational transduction?* That means *given a one-rule length-preserving rewrite system S, does there exists a finite transducer T_S such that for all word w, $T_S(w) = S(w)$?*

A particular binary relation over words seems of crucial interest for this question as it was conjectured in [9,15]. This relation can be seen as a generalization of the conjugacy relation, hence we call it quasi-conjugacy.

Definition 2.1. Two words u and v are *quasi-conjugate* if there exist words x, y and z such that $u = xyz$ and $v = zyx$.

Observe that if one of the words x, y, z is empty, then u and v are conjugate. So two conjugate words are also quasi-conjugate words. Note also that if u and v are quasi-conjugate, their factorization $u = xyz$ and $v = zyx$ need not be unique. For instance, for $u = aab$ and $v = baa$ we have the factorizations $u = aa.b$ and $v = b.aa$ but also $u = a.a.b$ and $v = b.a.a$. Similarly, for $u = abaca$ and $v = acaba$, we have $u = a.b.aca$ and $v = aca.b.a$ but also $u = aba.c.a$ and $v = a.c.aba$. A significant remark is that, at the contrary of the conjugacy relation, the quasi-conjugacy relation is not an equivalence relation since it is not transitive. As a matter of fact, one can show that if two words u and v are commutatively equivalent, there exists a *chain* of quasi-conjugate words from u to v *i.e.* words $x_0, \ldots x_n$ such that $u = x_0$, $v = x_n$ and for any $0 \leq i < n$, x_i and x_{i+1} are quasi-conjugate, hence:

Property 2.2. The transitive closure of the quasi-conjugacy relation is the total commutation relation.

Proof. Clearly it suffices to be able to permute two consecutive letters. For any words w, w' and letters a, b we have the following chain of quasi-conjugate words: $wab.w', w'.wab = w'w.a.b, b.a.w'w = baw'.w, w.baw'$. $\square$

Another property that is true for quasi-conjugacy and false for conjugacy is the following simplification lemma:

Lemma 2.3. *For any words u, v, w and w', if uwv and $uw'v$ are quasi-conjugate then w and w' are quasi-conjugate.*

Proof. The proof is an induction over $|u|+|v|$. First we consider the case $|u|+|v| = 1$ and we suppose that $u = a$ where a is a letter and $v = \varepsilon$. Let $aw = xyz$ and $aw' = zyx$. We consider different cases:

1. $x = \varepsilon$. Then $aw = yz$ and $aw' = zy$. If $y = \varepsilon$ or $z = \varepsilon$ it follows that $w = w'$. Else $y = ay'$ and $z = az'$ and we get $w = y'az'$ and $w' = z'ay'$, thus w and w' are quasi-conjugate.
2. $z = \varepsilon$. This case is symmetric with Case 1.
3. $x \neq \varepsilon$ and $z \neq \varepsilon$. Then $x = ax'$ and $z = az'$ and it follows $w = x'yaz'$ and $w' = z'yax'$, thus w and w' are quasi-conjugate.

Similarly, we can prove that w and w' are quasi-conjugate when $v = a$ and $u = \varepsilon$ therefore, by induction, it is also true when u and v are arbitrary words. $\square$

Now, if $u = xyz$ and $v = zyx$ are quasi-conjugate, u and v are obtainable by simplification from the two conjugate words yu and yv. So the quasi-conjugacy relation is the smallest relation containing the conjugacy relation and closed by simplification.

The following remark gives alternative characterizations for quasi-conjugacy:

Property 2.4. The following properties are equivalent:

1. The words u and v are quasi-conjugate.
2. There exists a word y such that yu and yv are conjugate.
3. There exist words y, z such that $yuz = zvy$.

Proof. 2 $\Longrightarrow$ 1 from Lemma 2.3. If $u = xyz$ and v=zyx for some words x, y, z then $yzy.u.y = yzyxyzy = y.v.yzy$ so 1 $\Longrightarrow$ 3. If $yuz = zvy$ for some words y, z then yu and vy are conjugate and it follows that yu and yv are conjugate so 3 $\Longrightarrow$ 2. □

It has been conjectured in [9], and later in [15], the following:

Conjecture 2.5. A one-rule rewrite system $S = \{u \to v\}$ with $|u| = |v|$ and $u \neq v$ is a rational transduction if and only if the words u and v are not quasi-conjugate.

In this paper, we "split" this conjecture in order to "split" its proof. More precisely, given a one-rule rewrite system $S = \{u \to v\}$ with $|u| = |v|$ and $u \neq v$, we want to prove that the following properties are equivalent:

(RAT) S is a rational transduction.
(REG) S preserves regularity.
(NCM) There do not exist words s and t such that $st \xrightarrow{+} ts$.
(NQC) u and v are not quasi-conjugate.

That would imply in particular that Conjecture 2.5 (which gives a decidable characterization of the property (RAT) since the property (NQC) is clearly decidable) is true. The property (NCM) points out for its part that the "rationality" of a one-rule length-preserving rewrite system depends on the impossibility to obtain by derivation the commutation of two words.

The implication $(RAT) \Longrightarrow (REG)$ is well known and the implication $(NCM) \Longrightarrow (NQC)$ is straightforward: indeed, if $S = \{u \to v\}$ with $u = xyz$ and $v = zyx$ for some words x, y and z, it follows $yxyz \to yzyx$. We shall prove in the rest of this paper that $(REG) \Longrightarrow (NCM)$ and that the implication $(NQC) \Longrightarrow (RAT)$ holds for some non trivial cases.

3. $(REG) \Longrightarrow (NCM)$

The proof of this implication uses a (pseudo) distance between words, defined in [14].

For all word w and all word $v \neq \varepsilon$, $\binom{w}{v}$ denotes the number of occurrences of the word v as a subword in w (see [4, 14] where it is called the generalized binomial coefficient). By convention, $\binom{w}{\varepsilon} = 1$ and, thanks to this convention, we get the relation

$$\binom{ww'}{z} = \sum_{z=xy} \left(\binom{w}{x} \times \binom{w'}{y} \right).$$

For instance, if $w = aab$, $w' = ab$, $\binom{ww'}{a} = \binom{w}{a} + \binom{w'}{a}$ and $\binom{ww'}{ab} = \binom{w}{ab} + \binom{w'}{ab} + \binom{w}{a} \times \binom{w'}{b} = 2 + 1 + 2 = 5$.

The notion of k-spectrum of a word appears in [11].

Definition 3.1. The k-spectrum with multiplicity of a word w defined over an alphabet A, denoted as $\mathrm{SP}_k(w) = \sum_{u\in A^k} \binom{w}{u}.u$, is the multiset of subwords of length k of the word w.

Example 3.2. Let $w = aabab$, then $\mathrm{SP}_2(w) = 3aa + 5ab + ba + bb$.

Given two words w and w', we can consider the integer $\Delta(w, w')$ defined as the smallest integer k such that $\mathrm{SP}_k(w) \neq \mathrm{SP}_k(w')$. This gives a kind of measure of similarity or distinguishability between two words. An equivalent definition of $\Delta(w, w')$ can be found in [14] and, following these authors, we will call $\Delta(w, w')$ a "distance" by abuse of language.

Definition 3.3. The *subword distance* between two distinct words w and w', denoted as $\Delta(w, w')$ is the smallest integer k such that $\mathrm{SP}_k(w) \neq \mathrm{SP}_k(w')$.

Example 3.4. The subword distance between the two words *abba* and *baab* is $\Delta(abba, baab) = 3$ since $\mathrm{SP}_2(abba) = aa+2ab+2ba+bb = \mathrm{SP}_2(baab)$, $\mathrm{SP}_3(abba) = abb + 2aba + bba$ and $\mathrm{SP}_3(baab) = baa + 2bab + aab$.

We observe that the subword distance between two distinct words is always strictly positive and that two distinct words are commutatively equivalent if and only if their subword distance is strictly greater than 1. One can also prove[4]:

Property 3.5. Let w and w' be two distinct words defined over an alphabet A and $k = \Delta(w, w')$ then

1. For any word w'' distinct from w and from w', $\Delta(w, w') \geq \min(\Delta(w'', w), \Delta(w'', w'))$.
2. If $|w| = |w'|$ then $\Delta(ww', w'w) = 1 + \Delta(w, w')$.
3. For all words p and q, for all word z such that $|z| \leq k$, $\binom{pwq}{z} - \binom{pw'q}{z} = \binom{w}{z} - \binom{w'}{z}$.
4. For any integer $i \geq 1$, for all word z such that $|z| \leq k$, $\binom{w^i}{z} - \binom{(w')^i}{z} = i(\binom{w}{z} - \binom{w'}{z})$.

Proof.

1. This comes directly from the definition of subword distance between two words: for any $k' < \min(\Delta(w'', w), \Delta(w'', w'))$, $\mathrm{SP}_{k'}(w) = \mathrm{SP}_{k'}(w'') = \mathrm{SP}_{k'}(w')$.
2. Firstly, one has $\Delta(ww', w'w) > k = \Delta(w, w')$. Indeed, for $z \in A^k$ and since $k > 0$,

$$\binom{ww'}{z} = \binom{w}{z} + \binom{w'}{z} + \sum_{\substack{z = xy \\ x, y \neq \varepsilon}} \left(\binom{w}{x} \times \binom{w'}{y} \right) = \binom{w'w}{z}.$$

[4]The property 3.5.3 already appeared as Lemma 3.5 in [10].

In order to prove that $\Delta(ww', w'w) \leq k+1$, we have to find $z \in A^{k+1}$ such that $\binom{ww'}{z} \neq \binom{w'w}{z}$. Let us consider two cases:

- If $\Delta(w, w') = 1$ then there exists a letter a such that $\binom{w}{a} \neq \binom{w'}{a}$. We can suppose that $\binom{w}{a} < \binom{w'}{a}$. Now, since $|w| = |w'|$, there exists a letter b, distinct from the letter a, such that $\binom{w}{b} > \binom{w'}{b}$ and it follows that $\binom{ww'}{ab} < \binom{w'w}{ab}$.
- If $k > 1$, w and w' are commutatively equivalent and for all letter a, for all integer i, $\binom{w}{a^i} = \binom{w'}{a^i}$. Let us consider the greatest integer i such that there exist a word $z \in A^{k-i}$ and a letter $a \in A$ which satisfy $\binom{w}{za^i} \neq \binom{w'}{za^i}$. It follows
$$\begin{aligned}\binom{ww'}{za^{i+1}} - \binom{w'w}{za^{i+1}} &= \sum_{xy=za^{i+1}} \left(\binom{w}{x} \times \binom{w'}{y}\right) - \sum_{xy=za^{i+1}} \left(\binom{w'}{x} \times \binom{w}{y}\right)\\ &= \binom{w}{za^i} \times \binom{w'}{a} - \binom{w'}{za^i} \times \binom{w}{a}\\ &= \left(\binom{w}{za^i} - \binom{w'}{za^i}\right) \times \binom{w}{a} \neq 0\end{aligned}$$
which proves that $\Delta(ww', w'w) = k+1$.

Observe that the condition $|w| = |w'|$ is necessary: for instance $\Delta(baab, ba) = 1$ but $\Delta(baabba, babaab) = 3$.

3. The proof is based on an induction over $|pq|$. Clearly, it is sufficient to consider the case $|pq| = 1$. In this case there exists a letter a such that $p = a$ and $q = \varepsilon$ or $p = \varepsilon$ and $q = a$. Let us consider the first case (the second one is symmetric) and let us prove that for all word z with $|z| \leq \Delta(w, w')$, $\binom{aw}{z} - \binom{aw'}{z} = \binom{w}{z} - \binom{w'}{z}$. This is clear if $z = \varepsilon$ or if z does not begin with a or if $z = a$. Else $z = az'$ with $0 < |z'| < \Delta(w, w')$ and $\binom{aw}{z} - \binom{aw'}{z} = \binom{w}{z'} + \binom{w}{z} - \binom{w'}{z'} - \binom{w'}{z} = \binom{w}{z'} + \binom{w}{z} - \binom{w}{z'} - \binom{w'}{z} = \binom{w}{z} - \binom{w'}{z}$.
4. The property is clearly true for $z = \varepsilon$. Else, observe that, for any word $z \neq \varepsilon$,
$$\binom{w^i}{z} = \binom{w}{z} + \binom{w^{i-1}}{z} + \sum_{\substack{z=xy\\ x,y\neq\varepsilon}} \left(\binom{w}{x} \times \binom{w^{i-1}}{y}\right).$$
We shall prove by induction over $i > 0$ that $\binom{w^i}{z} - \binom{(w')^i}{z} = i \times \left(\binom{w}{z} - \binom{w'}{z}\right)$ for any word $z \neq \varepsilon$ that satisfies $|z| \leq k$.
$$\binom{w^i}{z} - \binom{(w')^i}{z} = \binom{w}{z} + \binom{w^{i-1}}{z} + \sum_{\substack{z=xy\\ x,y\neq\varepsilon}} \left(\binom{w}{x} \times \binom{w^{i-1}}{y}\right)$$
$$-\binom{w'}{z} - \binom{(w')^{i-1}}{z} - \sum_{\substack{z=xy\\ x,y\neq\varepsilon}} \left(\binom{w'}{x} \times \binom{(w')^{i-1}}{y}\right)$$

and, by induction hypothesis,

$$\binom{w^i}{z} - \binom{(w')^i}{z} = i \times \left(\binom{w}{z} - \binom{w'}{z}\right) + \sum_{\substack{z = xy \\ x, y \neq \varepsilon}} \left(\binom{w}{x} \times (i-1) \times \left(\binom{w}{y} - \binom{w'}{y}\right)\right)$$

Since $|z| \leq k$, it follows

$$\sum_{\substack{z = xy \\ x, y \neq \varepsilon}} \left(\binom{w}{x} \times (i-1) \times \left(\binom{w}{y} - \binom{w'}{y}\right)\right) = 0$$

and $\binom{w^i}{z} - \binom{(w')^i}{z} = i \times (\binom{w}{z} - \binom{w'}{z})$.
Finally, for any word z, if $|z| < k$ then $\binom{w^i}{z} - \binom{(w')^i}{z} = 0$ and if we consider a word v such that $|v| = k$ and $\binom{w}{v} \neq \binom{w'}{v}$ then $\binom{w^i}{v} \neq \binom{(w')^i}{v}$ that proves the property.

As a direct consequence of items 3 and 4 of Property 3.5, we get:

Property 3.6. Let w and w' be two distinct words defined over an alphabet A and $k = \Delta(w, w')$ then

1. For all words p and q, $\Delta(pwq, pw'q) = \Delta(w, w')$.
2. For any integer $i \geq 1$, $\Delta(w^i, w'^i) = \Delta(w, w')$.

Thanks to subword distance, we can now prove the following lemma:

Lemma 3.7. *Let $S = \{u \to v\}$ be a one-rule rewrite system with $|u| = |v|$ and $u \neq v$. Then for all words w and w':*

1. $w \xrightarrow{+} w' \Longrightarrow \Delta(w, w') = \Delta(u, v)$;
2. *if* $w \xrightarrow{n} w'$ *and* $w \xrightarrow{m} w'$ *then* $n = m$.

Proof. From Item 3 of Property 3.5, we easily obtain by induction over n that, for all integer $n > 0$, $w \xrightarrow{n} w'$ implies that for all word z with $|z| \leq \Delta(u, v)$, $\binom{w}{z} - \binom{w'}{z} = n(\binom{u}{z} - \binom{v}{z})$. It follows $\Delta(w, w') = \Delta(u, v)$. For 2, if we consider a word z such that $|z| = \Delta(u, v)$ and $\binom{u}{z} \neq \binom{v}{z}$, we get $n(\binom{u}{z} - \binom{v}{z}) = m(\binom{u}{z} - \binom{v}{z})$ so $n = m$. □

We can deduce:

Lemma 3.8. *Let $S = \{u \to v\}$ be a one-rule rewrite system with $|u| = |v|$ and $u \neq v$. If $xy \xrightarrow{+} yx$ and $(xy)^n \xrightarrow{+} y^s x^t$ for some words x and y and some integers n, s, t then $s = t = n$. Hence $S((xy)^*) \cap y^* x^* = \{y^n x^n \mid n \geq 0\}$ is not a regular language.*

Proof. We consider two cases:

1. $|x| = |y|$. Let us suppose that $s > n$, we prove that it leads to a contradiction (the case $s < n$ is symmetric). Since $(xy)^n \xrightarrow{+} y^s x^t$ and $(xy)^n \xrightarrow{+} y^n x^n$, we get $\Delta((xy)^n, y^s x^t) = \Delta((xy)^n, y^n x^n) = \Delta(u, v)$ from Lemma 3.7. Now, from Item 1 of Property 3.5, we get

$$\Delta(y^s x^t, y^n x^n) \geq \min(\Delta((xy)^n, y^s x^t), \Delta((xy)^n, y^n x^n)) = \Delta(u, v).$$

Moreover $\Delta(y^s x^t, y^n x^n) = \Delta(y^{s-n}, x^{n-t})$ from Item 1 of Property 3.6 and, since $s - n = n - t$, $\Delta(y^{s-n}, x^{n-t}) = \Delta(x, y)$. It follows $\Delta(x, y) \geq \Delta(u, v)$. On the other hand, we get $\Delta(xy, yx) = \Delta(u, v)$ from Lemma 3.7, and it follows from Item 2 of Property 3.5 that $\Delta(x, y) = \Delta(u, v) - 1$, a contradiction.
2. $|x| \neq |y|$. Let $i > 0$ and $j > 0$ such that $|x^i| = |y^j|$ and suppose that $i < j$ (the case $i > j$ is symmetric). Since $xy \xrightarrow{+} yx$, we have $x^i y^j \xrightarrow{+} y^j x^i$, $x^i y^j \xrightarrow{+} y^{j-i} x^i y^i$ and $x^i y^i \xrightarrow{+} (xy)^i$. We get $(x^i y^j)^{jn} \xrightarrow{+} y^{jn(j-i)} (xy)^{nij} \xrightarrow{+} y^{jn(j-i)} (y^s x^t)^{ij} \xrightarrow{+} (y^j)^{nj-ni+is} (x^i)^{jt}$. It follows that $nj - ni + is = jn$ and $jt = jn$ which implies $s = t = n$. □

From this lemma, we directly obtain that $(REG) \Longrightarrow (NCM)$:

Proposition 3.9. *Let $S = \{u \to v\}$ be a one-rule rewrite system with $u \neq v$. If S preserves regularity then there do not exist words s and t such that $st \xrightarrow{+} ts$.*

Remark 3.10. This proposition does not hold for arbitrary rewrite systems, even length-preserving: as a counterexample, consider the rewrite system $\{a \to b, b \to a\}$.

As a direct consequence of Proposition 3.9, we get that the only if part of Conjecture 2.5 is true:

Proposition 3.11. *If a one-rule rewrite system $S = \{u \to v\}$ with $|u| = |v|$ and $u \neq v$ is a rational transduction then the words u and v are not quasi-conjugate.*

4. A sufficient condition for the rationality of S

The following notion generalizes to rewrite systems the notion of left-to-right derivation that was introduced by Jacques Sakarovitch in [13] for one rule rewrite systems.

Definition 4.1. Let S be a rewrite system. A derivation $w_0 \to w_1 \ldots \to w_k$ is *strictly left* if $k \leq 1$ or else if it holds that for all $i > 0$ if $w_{i-1} = xuy$, $w_i = xvy = x'u'y'$ and $w_{i+1} = x'v'y'$ with $u \to v \in S$ and $u' \to v' \in S$ then

(i) $|x| \leq |x'|$ and
(ii) $|y'| \leq |y|$.

Remark 4.2.

- if we consider only Condition (i), we find the definition of left-to-right derivation;
- we can define symmetrically the notion of strictly right derivation from a word w to a word w'.

Definition 4.3. Let S be a rewrite system. A derivation $w \xrightarrow{*} w' \xrightarrow{*} w''$ is a *left-right* (resp. right-left) derivation if it is the composition of a strictly left (resp. strictly right) derivation from w to w' with a strictly right (resp. strictly left) derivation from w' to w''.

We also define the notion of strictly left (resp. strictly right, left-right, right-left) rewriting system and the notion of strictly left (resp. strictly right, left-right, right-left) rewriting induced by a rewrite system:

Definition 4.4. A rewrite system S is strictly left (resp. strictly right, left-right, right-left) if for any word w and $w' \in S(w)$ there exists a strictly left (resp. strictly right, left-right, right-left) derivation from w to w'.

Definition 4.5. The strictly left (resp. strictly right, left-right, right-left) rewriting induced by a rewrite system S is the application that associates with any word w the set of words w' such that there exists a strictly left (resp. strictly right, left-right, right-left) derivation from w to w'.

Observe that the fact that a derivation is strictly left is a *global* property of the derivation. More precisely, if $w \xrightarrow{*} w'$ and $w' \xrightarrow{*} w''$ are two strictly left derivations, it does not imply that the derivation $w \xrightarrow{*} w' \xrightarrow{*} w''$ is strictly left. It does not imply either that there exists some strictly left derivation from w to w'' as shown in the following example:

Example 4.6. Let $S_2 = \{b_0c \to b_1c, ab_1 \to ab_2, b_1c \to b_2c, ab_2 \to ab_3\}$. Let us consider the possible derivations from ab_0c to ab_3c. There are only two possibilities: $a\underline{b_0c} \to a\underline{b_1c} \to \underline{ab_2}c \to ab_3c$ and $a\underline{b_0c} \to \underline{ab_1}c \to \underline{ab_2}c \to ab_3c$ that are not strictly left, so S_2 is not strictly left. Observe that, for any derivation $w \xrightarrow{*} w'$ of length 2 in S_2, there exists a strictly left derivation (of length 2) in S_2 from w to w'. This condition does not imply that S_2 is strictly left.

Proposition 4.7. *The strictly left (resp. strictly right, left-right, right-left) rewriting induced by a rewrite system S is a rational transduction.*

Proof. Clearly, it is sufficient to consider the case of a strictly left rewrite system: indeed, the proof is symmetric for a strictly right rewrite system and the class of rational transductions is closed under composition. So, let A be an alphabet and $S \subseteq A^* \times A^*$ be a strictly left rewrite system. We define, following the notations of [1], the transducer $T = (A, A, Q, q_-, Q_+, E)$ where A is the input and output alphabet, Q is the finite set of states, q_- is the initial state, Q_+ is the set of final states and $E \subseteq Q \times A^* \times A^* \times Q$ is the finite set of transitions by:

- $Q = \mathrm{LF}\{u \in A^* \mid \exists v \in A^*, u \to v \in S\}$,

- $q_- = \varepsilon$,
- $Q_+ = \{\varepsilon\}$,
- $E = \{(\varepsilon, a, a, \varepsilon) \mid a \in A\} \cup \{(l_1, l_2, r_1, r_2) \ \mid \ r_2 \in Q \wedge l_1 l_2 \to r_1 r_2 \in S)\}$.

From this definition, we easily obtain by induction over the length of paths in this transducer that any path from the state ε to a state u, labeled by (w, w') in E^* satisfies that there exists a strictly left derivation from w to $w'u$. Conversely, we shall first prove that for all words w, x and all rule $l \to r$ such that there exists a strictly left derivation $w \xrightarrow{*} xl \to xr$ and for all factorization $r = r_1 r_2$ with $r_2 \in Q$, there exists a path from ε to r_2, labeled by (w, xr_1) in E^*. The proof is based on an induction over the length of the derivation $w \xrightarrow{*} xl$. If $w = xl$ then there exists a path from ε to r_2, labeled by (xl, xr_1) in E^* since we have a path from ε to ε labeled by (x, x) and $(\varepsilon, l, r_1, r_2)$ in E. Else there exists a strictly left derivation $w \xrightarrow{*} x'l'y' \to x'r'y' = xl \to xr$ for some words x', l', r', y'. Observe that $w = w'y'$ for some word w', with a strictly left derivation $w' \xrightarrow{*} x'r'$, since the derivation is strictly left. We consider two cases:

1. $|l| \leq |y'|$. In this case $y' = \alpha l$ for some word α. From the induction hypothesis, we have a path from ε to ε labeled by $(w', x'r')$ and a path from ε to r_2 labeled by $(\alpha l, \alpha r_1)$. It follows that there exists a path from ε to r_2 labeled by (w, xr_1) since $w = w'\alpha l$ and $x'r'\alpha r_1 = xr_1$.
2. $|l| > |y'|$. Observe that $|l| \leq |r'| + |y'|$ since the derivation is strictly left. It follows that $x = x'r'_1$, $l = r'_2 y'$ and $r = r'_1 r'_2$ for some words r'_1, r'_2. From the induction hypothesis, we have a path from ε to r'_2 labeled by $(w', x'r'_1)$. Moreover $(r'_2, y', r_1, r_2) \in E$ and it follows that there exists a path from ε to r_2 labeled by (w, xr_1) since $w = w'y'$ and $x'r'_1 r_1 = xr_1$.

Now, if we consider a strictly left derivation $w \xrightarrow{+} w'$ for some arbitrary words w and w', there exist words x, l, r, y, w'' such that $w = w''y$, $w' = xry$ and $w'' \xrightarrow{*} xl \to xr$ that is a strictly left derivation. From the above property, there exists a path from ε to ε labeled by (w'', xr). Moreover there exists a path from ε to ε labeled by (y, y) and it follows that there exists a path from ε to ε labeled by (w, w') since $w = w''y$ and $w' = xry$. That finishes the proof of the proposition. □

In the rest of this section we only consider one-rule length-preserving rewrite systems. For such a system $S = \{u \to v\}$, we denote:

- $X = \mathrm{LF}(u) \cap \mathrm{RF}(v) \cap A^+$,
- $Z = \mathrm{RF}(u) \cap \mathrm{LF}(v) \cap A^+$,
- $V' = vX^{-1} = \{v' \in A^* \ \mid \ \exists x \in X, v = v'x\}$,
  $V'' = Z^{-1}v = \{v'' \in A^* \ \mid \ \exists z \in Z, v = zv''\}$,
  $U' = uZ^{-1}$ and $U'' = X^{-1}u$,
- $Y' = Z^{-1}V' = V''X^{-1}$,
- $Y = X^{-1}U' = U''Z^{-1}$.

Definition 4.8. A *pair of overlaps* for a system $S = \{u \to v\}$ is a pair $(x, z) \in X \times Z$. A pair of overlaps (x, z) is called a *large pair of overlaps* if $|xz| > |u|$, else it is called a *short pair of overlaps*.

Example 4.9. Let $S_1 = \{abb \to baa\}$. Then $X = \{a\}$, $Z = \{b\}$, $V' = \{ba\}$, $V'' = \{aa\}$, $U' = \{ab\}$, $U'' = \{bb\}$, $Y' = \{a\}$ and $Y = \{b\}$. System S_1 has a single (short) pair of overlaps.

Example 4.10. Let $S_3 = \{abbaa \to bbaab\}$. Then $X = \{ab\}$, $Z = \{bbaa\}$, $V' = \{bba\}$, $V'' = \{b\}$, $U' = \{a\}$, $U'' = \{baa\}$ and $Y' = Y = \emptyset$. System S_3 has a single (large) pair of overlaps.

Example 4.11. Let $S_4 = \{babaabbababbbab \to ababbbabaabbaba\}$.
Then $X = \{ba, baba, babaabbaba\}$, $Z = \{ab, ababbbab\}$ and S_4 has five short pairs of overlaps and one large pair of overlaps.
We also have $V' = \{ababbbabaabba, ababbbabaab, ababb\}$,
$V'' = \{abbbabaabbaba, aabbaba\}$,
$U' = \{babaabbababbb, babaabb\}$,
$U'' = \{baabbababbbab, abbababbbab, bbbab\}$ and
$Y' = \{abbbabaabba, abbbabaab, abb, aabba, aab\}$,
$Y = \{baabbababbb, abaabb, abbababbb, abb, bbb\}$.
Observe that $Y \cap Y' = \{abb\} \neq \emptyset$; nevertheless u and v are not quasi-conjugate since they are not commutatively equivalent.

Lemma 4.12. *Let $S = \{u \to v\}$ be a one-rule length-preserving rewrite system over an alphabet A.*

- *S is strictly left iff $Z \subseteq X$,*
- *S is strictly right iff $X \subseteq Z$.*

Proof. Let us suppose that $Z \subseteq X$. Let w and w' be two words with $w \xrightarrow{n} w'$. We shall prove by induction over n that we can build a strictly left derivation of length n from w to w'. This is clearly true for $n \leq 1$. Else, let us consider a derivation $w = \alpha_0 u \beta_0 \to \alpha_0 v \beta_0 \xrightarrow{n-1} w'$ such that if $w = \alpha_0' u \beta_0'$ with $|\alpha_0'| < |\alpha_0|$ then there is no derivation of length $n-1$ from $\alpha_0' v \beta_0'$ to w'. From the induction hypothesis, there exists a strictly left derivation, whose length is $n-1$ thanks to Lemma 3.7: $w_1 \to w_2 \xrightarrow{n-2} w_n = w'$ with

$$\forall 1 < i < n, w_i = \alpha_{i-1} v \beta_{i-1} = \alpha_i u \beta_i \wedge |\alpha_{i-1}| \leq |\alpha_i|.$$

If $|\alpha_0| \leq |\alpha_1|$ then $w_0 \to w_1 \to w_2 \xrightarrow{n-2} w_n = w'$ is strictly left therefore we can suppose $|\alpha_0| > |\alpha_1|$. Moreover, we cannot have $|\alpha_0| \geq |\alpha_1| + |u|$: in this case $w_0 = \alpha_1 u \beta_1' u \beta_0$ and $w_1 = \alpha_1 u \beta_1' v \beta_0$ for some word β_1'. It follows that the word $w_1' = \alpha_1 v \beta_1' u \beta_0$ satisfies $w_1' \xrightarrow{n-1} w'$ that contradicts the hypothesis since $|\alpha_1| < |\alpha_0|$. Thus we can suppose $|\alpha_0| < |\alpha_1| + |u|$. It follows $\alpha_0 = \alpha_1 u'$ with $u = u'z$, $v = zv'$ and $\beta_1 = v'\beta_0$ for some words u', v' and some word $z \in Z \subseteq X$. Since $z \in X$, there exist words u'' and v'' such that $u = zu''$ and $v = v''z$, so $u'u = uu''$, $vu'' = v''u$ and $v''v = vv'$. We finally obtain the following strictly left derivation from w to w': $w = \alpha_0 u \beta_0 = \alpha_1 u' u \beta_0 = \alpha_1 u u'' \beta_0 \to \alpha_1 v u'' \beta_0 = \alpha_1 v'' u \beta_0 \to \alpha_1 v'' v \beta_0 = \alpha_1 v v' \beta_0 = \alpha_1 v \beta_1 = w_2 \xrightarrow{n-2} w_n = w'$.

Conversely, let $z \in Z \setminus X$. Then $u = u'z$ and $v = zv''$. Let us consider the derivation $u'u \to u'v = uv'' \to vv''$. We shall prove that there does not exist a strictly left derivation from $u'u$ to vv''. Thanks to Lemma 3.7, the length of all derivation from $u'u$ to vv'' is 2. Let us consider a derivation

$$u'u = \alpha_0 u\beta_0 \to \alpha_0 v\beta_0 = \alpha_1 u\beta_1 \to \alpha_1 v\beta_1 = vv''$$

with $|\alpha_0| \leq |\alpha_1|$ and $\beta_0 \neq \varepsilon$, $\alpha_1 \neq \varepsilon$. If $u'u = \alpha_0 u\beta_0$, then $u' = \alpha_0\alpha_0'$ for some word α_0' and it follows $\alpha_0' u = u\beta_0$ that implies $u \in \mathrm{LF}(\alpha_0'^*)$. Moreover $\alpha_0\alpha_0' = \alpha_0'\alpha_0$ so $u \in \mathrm{LF}(\alpha_0\alpha_0')$ that implies $u \in \mathrm{LF}(u'u)$. On the other hand, since u and v are distinct, $u = sbs'$ and $v = sas''$ for some words s, s', s'' and some distinct letters a and b. Now, if we suppose $\alpha_0 \neq \varepsilon$, it follows $sb \in \mathrm{LF}(u'u)$, $sb \in \mathrm{LF}(\alpha_0 v\beta_0)$ and $sb \in \mathrm{LF}(\alpha_1 v\beta_1)$, a contradiction since $v = sas''$. Thus $\alpha_0 = \varepsilon$ and symmetrically we get $\beta_1 = \varepsilon$. Now, since $u \in \mathrm{LF}(u'u)$ it follows $z \in \mathrm{LF}(u)$ and, since $z \notin X$, it follows $z \notin \mathrm{RF}(v)$, a contradiction with the equality $v\beta_0 = \alpha_1 u$. Thus there is no strictly left derivation from $u'u$ to vv''.

Similarly, we can prove that S is strictly right iff $X \subseteq Z$. □

Depending on whether the set $S(V''A^*) \cap U''A^*$ is empty or not is a key property in the following of the paper. Here, emptiness of this set gives a sufficient condition for a one-rule length-preserving rewrite system to be a rational transduction. More precisely the condition implies that the system is left-right.

Lemma 4.13. *Let $S = \{u \to v\}$ be a one-rule length-preserving rewrite system. If $S(V''A^*) \cap U''A^* = \emptyset$ then the system S is left-right.*

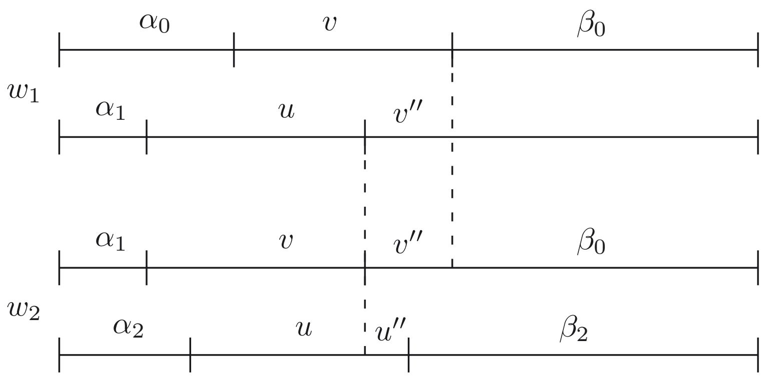
Proof. The proof is based on an induction over the length n of the derivation between two words w and $w' \in S(w)$. More precisely she shall prove that for all derivation from a word w to a word w' of length n, there exists a left-right derivation from w to w'. This is clearly true for $n \leq 1$ since in this case the derivation is strictly left. If $n > 1$, let us consider a derivation $w = \alpha_0 u\beta_0 \to \alpha_0 v\beta_0 \xrightarrow{n-1} w'$ such that if $w = \alpha_0' u\beta_0'$ with $|\alpha_0'| < |\alpha_0|$ then there is no derivation of length $n-1$ from $\alpha_0' v\beta_0'$ to w'. From the induction hypothesis, there exists a left-right derivation (of length $n-1$) from w_1 to $w_n = w'$: there exists an index $1 \leq q \leq n$ such that $w_1 \xrightarrow{q-1} w_q$ is strictly left and $w_q \xrightarrow{n-q} w_n$ is strictly right. It follows that for all $0 < i < n$, $w_i = \alpha_{i-1} v\beta_{i-1} = \alpha_i u\beta_i$ with:

- for all $1 < i \leq q$, $|\alpha_{i-1}| < |\alpha_i|$;
- for all $q < i \leq p$, $|\alpha_{i-1}| > |\alpha_i|$.

If $|\alpha_0| \leq |\alpha_1|$ then $w_0 \xrightarrow{*} w_q$ is strictly left therefore we can suppose $|\alpha_0| > |\alpha_1|$. Moreover, we cannot have $|\alpha_0| \geq |\alpha_1| + |u|$: in this case $w_0 = \alpha_1 u\beta_1' u\beta_0$ and $w_1 = \alpha_1 u\beta_1' v\beta_0$ for some word β_1'. It follows that the word $w_1' = \alpha_1 v\beta_1' u\beta_0$ satisfies $w_0 = \alpha_1 u\beta_1' u\beta_0 \to w_1' \xrightarrow{n-1} w'$ that contradicts the hypothesis since $|\alpha_1| < |\alpha_0|$.

Thus we can suppose $|\alpha_0| < |\alpha_1| + |u|$. We consider two cases:

- $\forall j > 1, |\alpha_j| \geq |\alpha_1|+|u|$. It follows that for all $1 < j < p$, $\alpha_j u\beta_j = \alpha 1 v\alpha'_j u\beta_j$ and we get $w_0 \to \alpha_1 u\alpha'_2 u\beta_2 \xrightarrow{*} \alpha_1 u\alpha'_{p-1} v\beta_{p-1} \to w_p$ that is a left-right derivation.
- $\exists j > 1, |\alpha_j| < |\alpha_1| + |u|$. Let us consider the smallest index $s > 1$ such that $|\alpha_s| < |\alpha_1| + |u|$. Necessarily, $s = 2$ or $s > q$ since the derivation is strictly left from w_1 to w_q. If $s = 2$, then $w_1 = \alpha_0 v\beta_0 = \alpha_1 u\beta_1$. Since $|\alpha_1| < |\alpha_0| < |\alpha_1| + |u|$, it follows that $\beta_1 = v''\beta_0$ with $v'' \in V''$ and $w_2 = \alpha_1 vv''\beta_0 = \alpha_2 u\beta_2$. Since $|\alpha_1| < |\alpha_2| < |\alpha_1| + |u|$ we finally get, as described in the following picture, $v''\beta_0 = u''\beta_2$ with $u'' \in U''$ that contradicts the hypothesis $S(V''A^*) \cap U''A^* = \emptyset$.



If $s > q$ then we obtain a similar contradiction: indeed, $|\alpha_1| < |\alpha_0| < |\alpha_1| + |u|$ implies that $\beta_1 = v''\beta_0$ with $v'' \in V''$ and $w_2 = \alpha_1 vv''\beta_0$. Moreover, for all $2 \leq i < s$, $\alpha_1 v$ is a prefix of w_i since $|\alpha_i| > |\alpha_1| + |u|$. It follows that $w_2 = \alpha_1 vv''\beta_0 \xrightarrow{*} \alpha_1 vz = \alpha_s u\beta_s$ for some word z with $v''\beta_0 \xrightarrow{*} z$. Now, since $|\alpha_s| < |\alpha_1| + |u|$, it follows $z = u''z'$ with $u'' \in U''$ and $v''\beta_0 \xrightarrow{*} u''z'$ that contradicts the hypothesis $S(V''A^*) \cap U''A^* = \emptyset$ and finishes the proof. □

We do not know whether the converse holds or not. Nevertheless, we shall prove in Section 6 that this condition becomes a necessary and sufficient condition when the system S has no short pair of overlaps. We first consider in the next section the dual case, that is when S has no large pair of overlaps.

5. Without large pair of overlaps

We shall prove in this section that Conjecture 2.5 is true for all one-rule length-preserving rewrite system S that has no large pair of overlaps. This result is a consequence of the following key lemma:

Lemma 5.1. *If $S = \{u \to v\}$ is a one-rule length-preserving rewrite system without large pair of overlaps, the three following statements are equivalent:*

(i) *u and v are quasi-conjugate,*

(ii) $S(V''A^*) \cap U''A^* \neq \emptyset$,
(iii) $S(A^*V') \cap A^*U' \neq \emptyset$.

Proof. We prove 5.1 implies 5.1 and 5.1 implies 5.1, the proofs for 5.1 implies 5.1 and 5.1 implies 5.1 are similar.

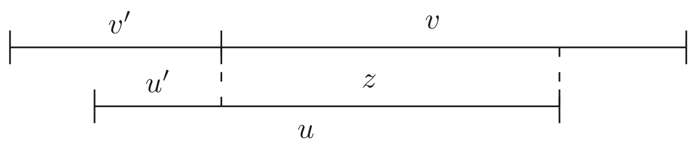
- 5.1 implies 5.1: Assume that $u = xyz$ and $v = zyx$. Then $yx \in V''$, $yz \in U''$ and $yxyz \xrightarrow{*} yzyx \in S(V''A^*) \cap U''A^*$.
- 5.1 implies 5.1: Assume that u and v are not quasi-conjugate and prove that $S(V''A^*) \cap U''A^* = \emptyset$. We need some intermediate results.

Claim 5.2. $V''A^* \cap U''A^* = \emptyset$ and $A^*U' \cap A^*V' = \emptyset$.

We prove $V''A^* \cap U''A^* = \emptyset$, the proof for $A^*U' \cap A^*V' = \emptyset$ is similar. Let $u'' \in U''$ and $v'' \in V''$ and assume that $u''f = v''f'$ for some words f and f'. By definition, $u = xu''$ with $x \in X$ and $v = zv''$ with $z \in Z$. Since S is without large pair of overlaps, $|x| + |z| \leq |u|$ and it follows that there exists words y and y' such that $u = xyz$ and $v = zy'x$ that implies $u'' = yz$ and $v'' = y'x$ therefore $yzf = y'zf'$. Since $|y| = |y'|$, it follows $y = y'$, a contradiction since u and v are not quasi-conjugate.

Claim 5.3. $u \notin \mathrm{F}(V'^*v)$.

In order to prove this claim, we assume that $u \in \mathrm{F}(V'^*v)$ and we consider two cases:
– $u \in \mathrm{F}(v'v)$ with $v' \in V'$.



This would imply $A^*U' \cap A^*V' \neq \emptyset$ that contradicts Claim 5.2.
– $u = rv'z$ with $v' \in V'$ and $z \in Z$ for some word r. It follows rv' in U' and $A^*U' \cap A^*V' \neq \emptyset$ that contradicts again Claim 5.2.

Claim 5.4. $S(V''A^*) = V''A^* \cup Y'V'^*vA^*$.

In order to prove this claim, observe that $S(V''A^*) = V''A^* \cup S(Y'vA^*)$. Moreover, since $Y' = Z^{-1}V'$ we get $F(Y'V'^*v) = F(V'^*v)$ and, from Claim 5.3, $u \notin F(Y'V'^*v)$. In particular, $u \notin F(Y'v)$ and it follows that $S(Y'vA^*)$ is included in $Y'V'^*vA^*$. That proves the equality $S(Y'vA^*) = Y'V'^*vA^*$ since the converse inclusion is clear.

Claim 5.5. $Y \cap Y'V'^* = \emptyset$.

Indeed, if $Y \cap A^*V' \neq \emptyset$, it follows $U' \cap A^*V' \neq \emptyset$ that contradicts Claim 5.2. Therefore $Y \cap Y'V'^+ = \emptyset$. If $y \in Y \cap Y'$, it follows $u = xyz$ and $v = z'yx'$ with $x, x' \in X$ and $z, z' \in Z$. Assume that $|x| \geq |x'|$. We get $|z'| \geq |z|$ and, since $|x| + |z'| \leq |u|$, $u = xy''z'$ and $v = z'y'''x$ for some words y'', y'''. Thus y'' and y''' are two words in $\mathrm{LF}(y)$ having the same length. That implies $y'' = y'''$ and u, v are quasi-conjugate, a contradiction.

We can now finish the proof 5.1 implies 5.1: thanks to Claim 5.2 and Claim 5.4, in order to get $S(V''A^*) \cap U''A^* = \emptyset$, it remains to prove $Y'V'^*vA^* \cap U''A^* = \emptyset$ that is $\mathrm{LF}(Y'V'^*v) \cap U'' = \emptyset$. Assume that $u'' \in \mathrm{LF}(Y'V'^*v) \cap U''$. From Claim 5.2, $u'' \notin \mathrm{LF}(Y')$. Then there exist $v_1, \ldots, v_t \in V'$ such that $u'' = y'v_1 \ldots v_{t-1}z$ with $z \in \mathrm{RF}(u'') \cap \mathrm{LF}(v_t) \subseteq \mathrm{RF}(u) \cap \mathrm{LF}(v)$. Thus $z \in Z$ and $y'v_1 \ldots v_{t-1} \in Y$, contradicting Claim 5.5. □

As a consequence of Proposition 3.9, Proposition 4.7, Lemmas 4.13 and 5.1, we get:

Proposition 5.6. *Conjecture 2.5 is true for all one-rule length-preserving rewrite system S that has no large pair of overlaps. Moreover, when S is a rational transduction, S is left-right and right-left.*

Proof. Thanks to Proposition 3.11, it remains to prove that $(NQC) \Longrightarrow (RAT)$ for a one-rule length-preserving rewrite system $S = \{u \to v\}$ with $u \neq v$ that has no large pair of overlaps. From Lemma 5.1, if u and v are not quasi-conjugate then $S(V''A^*) \cap U''A^* = \emptyset$ and $S(A^*V') \cap A^*U' = \emptyset$. It follows from Lemma 4.13 that S is left-right and right-left that implies from Proposition 4.7 that S is a rational transduction. □

In the case when the set of letters occurring in u differs from the set of letters occurring in v, the system $S = \{u \to v\}$ cannot have a large pair of overlaps and the words u and v cannot be quasi-conjugate. It follows:

Corollary 5.7. *Let $S = \{u \to v\}$ be a one-rule length-preserving rewrite system. If there is a letter that occurs in u but not in v, or that occurs in v but not in u, then S is a rational transduction.*

It has been proved in [7] (and stated again in [16]) that a one-rule length-preserving rewrite system $S = \{u \to v\}$ is confluent if and only if

$$(\mathrm{LF}(u) \cap \mathrm{RF}(u)) \setminus \{u\} \subseteq \mathrm{LF}(v) \cap \mathrm{RF}(v)$$

One can deduce:

Corollary 5.8. *Conjecture 2.5 is true for all confluent one-rule length-preserving rewrite system S. Moreover, when S is a rational transduction, S is left-right and right-left.*

Proof. Let $S = \{u \to v\}$ be a confluent one-rule length-preserving rewrite system with $u \neq v$. It is sufficient to prove that, if S has a large pair of overlaps then u and v are quasi-conjugate. Let (x, z) be a large pair of overlaps of S such that there does not exist another large pair of overlaps (x', z') with $|xz| > |x'z'|$. It follows $u = xu'' = u'z = u'mu''$ and $v = v'x = zv'' = v'm'v''$ for some words u', u'', v', v'', m, m' with $|m| = |m'|$. That implies $m' \in \mathrm{LF}(u) \cap \mathrm{RF}(u)$ and $m \in \mathrm{LF}(v) \cap \mathrm{RF}(v)$. Since S is confluent, we obtain $m = m' \in X \cap Z$. Moreover, since there does not exist another large pair of overlaps (x', z') with $|xz| > |x'z'|$, $|xm| \leq |u|$ and $|mz| \leq |u|$ that implies $x = msm$ and $z = mtm$ for some words s and t. It follows that $u = msmtm$ and $v = mtmsm$ are quasi-conjugate. $\square$

Remark 5.9. We also directly obtain from Corollary 5.8 that Conjecture 2.5 is true for all one-rule length-preserving rewrite system S such that S^{-1} is confluent.

The condition S *has no large pair of overlaps* in Lemma 5.1 is a mandatory condition. Indeed, when a system $S = \{u \to v\}$ has large pairs of overlaps, it is possible to get $S(V''A^*) \cap U''A^* \neq \emptyset$ with u and v that are not quasi-conjugate as shown in the following example.

Example 5.10. Let $S_3 = \{abbaa \to bbaab\}$. Although $abbaa$ and $bbaab$ are not quasi-conjugate, we have $X = \{ab\}$, $Z = \{bbaa\}$, $V'' = \{b\}$, $U'' = \{baa\}$ so $V''A^* \cap U''A^*$ is not empty. Moreover let us consider the derivation $a\underline{abbaa}aa \to \underline{abbaa}baa \to bba\underline{abbaa} \to bb\underline{abbaa}b \to bbbbaabb$. It is the only possible derivation from $abbaaaa$ to $bbbbaabb$ and it is neither left-right nor right-left. Nevertheless we shall see in the next section that the system $S_3 = \{abbaa \to bbaab\}$ is still a rational transduction.

6. Without short pair of overlaps

In the previous section, we have proved that for a system S that has no large pair of overlaps, (NQC) implies (RAT). Up to now, we are unable to prove this implication in the general case. However, we shall show the rationality of S under a stronger hypothesis than (NQC): if S has no short pair of overlaps then S is a rational transduction.

A first remark concerning this situation is that the language $S(V''A^*) \cap U''A^*$ occurring in Section 4 becomes easily calculable. Indeed, $\mathrm{RF}(V'') \cap \mathrm{LF}(u) = \{\varepsilon\}$ and $S(V''A^*) = V''A^*$. Similarly, we have $S(A^*V') = A^*V'$. It follows:

Lemma 6.1. *If $S = \{u \to v\}$ is a one-rule length-preserving rewrite system that has no short pair of overlaps then:*

- $S(V''A^*) \cap U''A^* = V''A^* \cap U''A^*$.
- $S(A^*V') \cap A^*U' = A^*V' \cap A^*U'$.

Another important property in this case, proved below, is that if S has at least one pair of overlaps but no short pair of overlaps, then S has a unique (large) pair of overlaps. Moreover u and v have particular factorizations:

Lemma 6.2. *Let $S = \{u \to v\}$ be a one-rule length-preserving rewrite system that has no short pair of overlaps. If $X \neq \emptyset$ and $Z \neq \emptyset$ then:*

1. *$X = \{x\}$ and $Z = \{z\}$;*
2. *$u = m'lmrm'$ and $v = mrm'lm$ with $x = m'lm$ and $z = mrm'$ for some words m, m', l, r.*

Proof.

1. Let x be the shortest element in X and z be the shortest element in Z. Suppose that there exists $x' \neq x$ in X. It follows $x \in \mathrm{LF}(x') \cap \mathrm{RF}(x')$. If $|xx| > |x'|$, then there exists $x'' \in \mathrm{LF}(x) \cap \mathrm{RF}(x) \subseteq X$ with $|x''| < |x|$, therefore $x' = xfx$ for some word f. Since $|xz| > |u|$, there exist $x_2 \in \mathrm{RF}(x)$ such that $x_2 fx \in \mathrm{LF}(z)$ and $x_1 \in \mathrm{LF}(x)$ such that $xfx_1 \in \mathrm{RF}(z)$. It follows $x_2 fx_1 \in \mathrm{LF}(z) \cap \mathrm{RF}(z) \subseteq Z$, a contradiction since $x_2 fx_1$ is shorter than z. The proof is similar if we consider that Z is not a singleton.
2. $u = u'mu''$ and $v = v'm'v''$ with $x = u'm = m'v''$ and $z = mu'' = v'm'$. If $|m'z| > |u|$, there exists a word $z' \neq z \in \mathrm{LF}(z) \cap \mathrm{RF}(m') \subseteq Z$, a contradiction. It follows $|m'z| \leq |u|$ and, similarly, $|mx| \leq |u|$. Therefore there exist words l and r such that $u' = m'l$ and $u'' = rm'$. $\square$

Example 6.3. Recall that for system $S_3 = \{abbaa \to bbaab\}$, we have $X = \{ab\}$, $Z = \{bbaa\}$ and it has no short pair of overlaps. For this system, $m = b$, $m' = a$, $r = ba$ and $l = \varepsilon$.

The simplest example of such a one-rule length-preserving rewrite system $S = \{u \to v\}$ that has only one large pair of overlaps is the case $l = r = \varepsilon$ and m, m' reduced to two distinct letters, that gives $u = bab$ and $v = aba$. The particular forms of u and v are now used to establish that the condition of the Lemma 4.13 becomes here a necessary and sufficient condition:

Lemma 6.4. *If $S = \{u \to v\}$ is a one-rule length-preserving rewrite system that has no short pair of overlaps then:*

1. *S is left-right iff $S(V''A^*) \cap U''A^* = \emptyset$;*
2. *S is right-left iff $S(A^*V') \cap A^*VU' = \emptyset$.*

Proof.

1. Assume that $S(V''A^*) \cap U''A^* \neq \emptyset$. That implies $X \neq \emptyset$ and $Z \neq \emptyset$ and it follows from Lemma 6.2 that S has a single large pair of overlaps and $u = m'lmrm'$, $v = mrm'lm$, $x = m'lm$, $z = mrm'$, $u' = m'l$, $u'' = rm'$, $v' = mr$

and $v'' = lm$ for some words m, m', l, r. We consider two cases:

(a) $v'' = u''f$ for some word f. We shall prove that the derivation

$$u'\underline{u} \to u'\underline{v} = \underline{u'z}u''f \to \underline{v}u''f = v'\underline{xu''}f \to v'\underline{v}f$$

is the unique derivation going from $u'u$ to $v'vf$. For that we show that there is a unique occurrence of u in $u'u$, $u'v$ and vv''. It is clear for $u'v$ and vv'' since $X = \{x\}$ and $Z = \{z\}$. Assume now that $u'u = \alpha u \beta$ for some words α and β such that $\beta \neq \varepsilon$. Since $u' \in \mathrm{LF}(u)$, we have $u' = \alpha\gamma = \gamma\alpha$ for some $\gamma \in \mathrm{LF}(u)$. Let ρ be the root of u', we get $u', \gamma \in \rho^*$ and $u \in \mathrm{LF}(\gamma^*) \subseteq \mathrm{LF}(\rho^*)$. It follows $u \in \mathrm{LF}(u'u)$, a contradiction since $u \in m'lmA^*$ and $u'u \in m'lm'A^*$ with $|m| = |m'|$ and $m \neq m'$. Finally, this unique derivation going from $u'u$ to $v'vf$ is not left-right since $|u'| > 0$ and $|v'| > 0$.

(b) $u'' = v''f$ for some word f. The proof is symmetric to the first case, considering S^{-1} and the derivation

$$v'\underline{v} \xrightarrow[S^{-1}]{} v'\underline{xu''} = \underline{v}u'' \xrightarrow[S^{-1}]{} \underline{u'z}u'' = u'\underline{v}f \xrightarrow[S^{-1}]{} u'\underline{u}f$$

2. The proof is similar to 1. □

We shall now prove the main result of this section: a one-rule length-preserving rewrite system that has no short pair of overlaps is always a rational transduction. For this, we use a result of Bala Ravikumar in [12], extended by Alfons Geser et al. in [6], that gives a sufficient condition for a length-preserving rewrite system to be a rational transduction. This condition is based on the notion of *change-bounded* length-preserving rewrite system that was introduced in the same article and that we recall here:

Let $S = \{u_1 \to v_1, \ldots, u_t \to v_t\}$ be a length-preserving rewrite system and consider a derivation

$$d = \alpha_0 u_{i_0} \beta_0 \to \alpha_0 v_{i_0} \beta_0 = \alpha_1 u_{i_1} \beta_1 \ldots \to \alpha_{n-1} v_{i_{n-1}} \beta_{n-1} = \alpha_n u_{i_n} \beta_n \to \alpha_n v_{i_n} \beta_n$$

For all integer j, we denote $\mathcal{N}(d,j)$ the subset of indexes $\mathcal{N}(d,j) = \{i \in [0,n] \mid |\alpha_i| = j\}$. The system S is called change-bounded (by K) if there exists an integer K such that for all derivation d and all integer j, the cardinality of $\mathcal{N}(d,j)$ is bounded by K. Intuitively, this means that in all derivation in S, the number of times that a change (an application of a rule) is made at a same position during the derivation is bounded by K. Bala Ravikumar has proved:

Proposition 6.5 [12]**.** *A change-bounded length-preserving rewrite system is a rational transduction.*

We prove now that a one-rule length-preserving rewrite system that has no short pair of overlaps is always change-bounded by 1:

Lemma 6.6. *Let $S = \{u \to v\}$ be a one-rule length-preserving rewrite system that has no short pair of overlaps and a derivation $\alpha v \beta \xrightarrow{*} \alpha' w \beta'$ with $|\alpha| = |\alpha'|$ and $|\beta| = |\beta'|$, then $w \neq u$.*

Proof. Clearly, if S has no pair of overlaps, we are done. So, one can assume that S has a unique large pair of overlaps and from Lemma 6.2, we have $u = xu'' = u'z$ and $v = v'x = zv''$ with $x = m'lm, z = mrm', u'' = rm', u' = m'l, v' = mr$ and $v'' = lm$ for some words m, m', l, r. Let us consider two cases:

1. $u''A^* \cap v''A^* = \emptyset = A^*u' \cap A^*v'$. We shall reason by induction on the length n of the derivation $\alpha v\beta \xrightarrow{n} \alpha' w\beta'$. If $n = 0$ then $w = v \neq u$, otherwise one can consider the first step of the derivation: $\alpha v\beta = \alpha'' u\beta'' \rightarrow \alpha'' v\beta'' \xrightarrow{n-1} \alpha' w\beta'$. There are three cases:
   (a) $|\alpha'' u| \leq |\alpha|$ or $|\alpha''| \geq |\alpha v|$. Then $\alpha'' v\beta'' = \alpha''' v\beta''$ with $|\alpha'''| = |\alpha| = |\alpha'|$ and $|\beta'''| = |\beta| = |\beta'|$ and it follows $w \neq u$ by the induction hypothesis.
   (b) $|\alpha| < |\alpha'' u| < |\alpha v|$. Then $\alpha = \gamma u'$ and $\gamma u' v\beta \rightarrow \gamma v v''\beta \xrightarrow{n-1} \alpha' w\beta'$. Since $S(\gamma v v'') \subseteq A^* v(v'')^+$ and that $\mathrm{LF}(u) \cap \mathrm{RF}(v(v'')^+) = \{\varepsilon\}$, it follows $w \in \mathrm{RF}(v(v'')^+) \subseteq A^* m$ that implies $w \neq u$.
   (c) $|\alpha| < |\alpha''| < |\alpha v|$. Similarly it comes that $w \in \mathrm{LF}(v'^+ v) \subseteq mA^*$ that implies $w \neq u$.
2. $u''A^* \cap v''A^* \neq \emptyset$ (the case $A^*u' \cap A^*v' \neq \emptyset$ is similar). One can assume that $|u''| \geq |v''|$, otherwise we consider S^{-1}. It follows $v'' \in \mathrm{LF}(u'')$. Let i be the largest integer such that $(v'')^i \in LF(u'')$. Since $\mathrm{LF}(m') \cap \mathrm{RF}(m) = \{\varepsilon\}$, it follows $r = (v'')^i r'$ with $v'' \notin \mathrm{LF}(r'm')$ so $u'' = (v'')^i r'm' \notin \mathrm{LF}((v'')^*)$. Let us consider now the language $K = \mathrm{LF}(v') \cap m(v'')^i A^*$. We need the following properties:

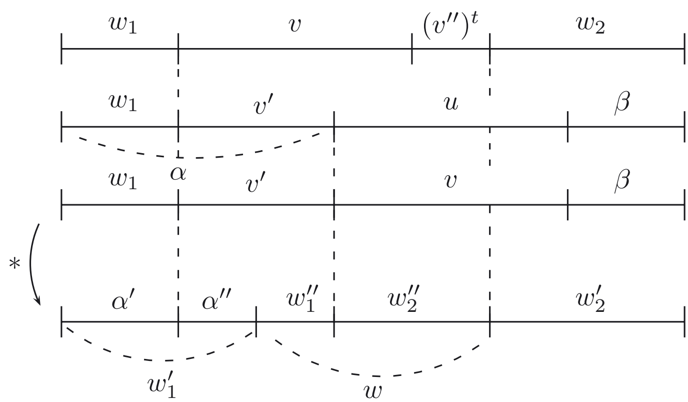
   Claim 6.7. For all word w, $S(wK^+vA^*) = S(w)K^+vA^*$.

   First, we prove that $S(K^+vA^*) = K^+vA^*$. For this, take $w_1 \in K^+$, $w_2 \in A^*$ and assume that $w_1 v w_2 = w_1' u w_2'$. We show that $w_1' v w_2' \in K^+vA^*$. Since $K \subseteq \mathrm{LF}(v')$ and $\mathrm{LF}(v') \cap \mathrm{RF}(u) = \{\varepsilon\}$, $u \notin \mathrm{F}(w_1)$ and $|w_1' u| > |w_1|$. If $|w_1'| \geq |w_1 v|$ then $w_1' v w_2' \in K^+vA^*$ else first assume $|w_1 v| > |w_1' u| > |w_1|$. It follows that $w_1 = w_1' u' \in K^* m(v'')^i \gamma$ for some γ. Since $|u'| = |v''|$ and $\mathrm{LF}(u') \cap \mathrm{RF}(v'') = \emptyset$ it follows $u' \in \mathrm{RF}(\gamma)$ so $w_1' \in K^* m(v'')^i A^* \subseteq K^+$ and $w_1' v w_2' \in K^+vA^*$. Assume, now, that $|w_1| < |w_1'| < |w_1 v|$. Since $v' \in K$, $w_1' v w_2' \in w_1 v' v A^* \subseteq K^+vA^*$. So, $S(K^+vA^*) = K^+vA^*$. Let us now consider $S(wK^+vA^*)$ for some word w. Let $w' \in wK^+vA^*$ and assume that $w' = \alpha u\beta$ with $|\alpha| < |w| < |\alpha u|$. Since $\mathrm{LF}(v') \cap \mathrm{RF}(u) = \{\varepsilon\}$, this would imply $\mathrm{RF}(u') \cap K^+ \neq \emptyset$, a contradiction since $|u'| < |mv''|$. It follows $S(wK^+vA^*) = S(w)S(K^+vA^*) = S(w)K^+vA^*$.

   Claim 6.8. $\forall t \geq 0$, if $w_1 v(v'')^t w_2 \xrightarrow{*} w_1' w w_2'$ with $|w| = |v|$ and $|w_2'| = |w_2|$ then $w \neq u$.

   We prove this claim by induction on the length n of the derivation from $w_1 v(v'')^t w_2$ to $w_1' w w_2'$. If $n = 0$, $w \neq u$ since $u \notin \mathrm{RF}(v(v'')^*)$. If $n > 0$, let us consider the first step of the derivation: $w_1 v(v'')^t w_2 = \alpha u\beta \rightarrow \alpha v\beta \xrightarrow{n-1} w_1' w w_2'$. Since $\mathrm{RF}(v'') \cap \mathrm{LF}(u) = \{\varepsilon\}$ we only have to consider four cases:
   (a) $|\alpha u| \leq |w_1|$ or $|u\beta| \leq |w_2|$. In this case $\alpha v\beta = w_1'' v(v'')^t w_2''$ with $|w_2''| = |w_2| = |w_2'|$ and it follows $w \neq u$ by induction hypothesis.

(b) $|w_1| < |\alpha u| < |w_1 v|$. Then $\alpha v\beta = \alpha v(v'')^{t+1}w_2$ and it follows $w \neq u$ by induction hypothesis.
(c) $|w_1| < |\alpha| < |w_1 v|$ and $|\beta| \geq |w_2|$. This would imply $u'' \in \mathrm{LF}((v'')^*)$, a contradiction.
(d) $|w_1| < |\alpha| < |w_1 v|$ and $|\beta| < |w_2|$. Then $\alpha u\beta = w_1 v' z(v'')^t w_2$ with $(v'')^t \in \mathrm{LF}(u'')$ so $t \leq i$. Moreover, $S(\alpha v\beta) \subseteq S(w_1)K^+vA^*$ from the previous claim since $v' \in K$. Let us consider $w_1' w w_2' \in S(\alpha v\beta)$ with $|w_2'| = |w_2|$ and $|w| = |v|$, then $w_1' = \alpha'\alpha''$ with $\alpha' \in S(w_1)$ and $w = w_1'' w_2''$ with $\alpha'' w_1'' \in K$. It follows $\alpha'' w_1'' = m(v'')^i\delta = m(lm)^i\delta$ for some δ and, since $t \leq i$, we get $\alpha'' w_1'' = (ml)^t(ml)^i m\delta$. Moreover $|\alpha''| = |(v'')^t|$ since $|\alpha'' w| = |v(v'')^t|$. It follows $\alpha'' = (ml)^t$ and $w_1'' = (ml)^{i-t}m$. This implies $w \in mA^*$ and $w \neq u$.



Finally, taking $t = 0$ in Claim 6.8 finishes the proof of the lemma. □

As a consequence of Lemma 6.6 and Proposition 6.5, we obtain:

Proposition 6.9. *A one-rule length-preserving rewrite system that has no short pair of overlaps is a rational transduction.*

We observe that a direct consequence of this proposition is that Conjecture 2.5 is satisfied in the case of a one-rule length-preserving rewrite system $S = \{u \to v\}$ that has no short pair of overlaps since, in this case the words u and v are not quasi-conjugate and S is a rational transduction so, combining the results of Sections 5 and 6, we get:

Proposition 6.10. *Conjecture 2.5 is true for all one-rule length-preserving rewrite system S that has a single pair of overlaps.*

7. Conclusion

In this paper, we have mainly studied the following conjecture: *a one-rule length-preserving rewrite system $S = \{u \to v\}$, with $u \neq v$, is a rational transduction if* *and only if u and v are not quasi-conjugate*. In Section 3 we have proved the *only if* part: if $S = \{u \to v\}$ is a rational transduction then u and v are not quasi-conjugate. Conversely, in Sections 5 and 6, we have proved the *if* part in two cases: when the system S has no large pair of overlaps and when it has no short pair of overlaps. The main remaining open question is to know whether the conjecture is true for one-rule length-preserving rewrite systems that have both short and large pair of overlaps. Nevertheless some other questions, listed below, deserve to be studied.

The Thue congruence $\overset{*}{\longleftrightarrow}$ generated by a rewrite system S is the reflexive transitive symmetric closure of $\to$. It corresponds to the derivation relation of the Thue system $S^{\leftrightarrow} = S \cup S^{-1}$ generated by S. We may wonder whether or when the rationality of S implies the rationality of $S^{\leftrightarrow}$.

It is the case when S (or S^{-1}) is a confluent one-rule length-preserving rewrite system since, when S is confluent, $S^{\leftrightarrow} = S^{-1} \circ S$ and, as a consequence, if S is a rational transduction, then $S^{\leftrightarrow}$ is a rational transduction; but it is not true for arbitrary one-rule length-preserving rewrite systems: for instance the rewrite system $S_5 = \{aa \to bb\}$ is clearly a rational transduction since it has no pair of overlaps, but, rather surprisingly, it is not the case for $S_5^{\leftrightarrow} = \{aa \leftrightarrow bb\}$. Indeed, $abba \leftrightarrow aaaa \leftrightarrow bbaa \leftrightarrow bbbb \leftrightarrow baab$ so, for all integer n, $(abba)^n \overset{*}{\longleftrightarrow} (ba)^n(ab)^n$. Moreover, we can prove that for all integer $0 < i \le n$, $(ba)^{n-i}(ab)^{n+i} \overset{*}{\longleftrightarrow} (ba)^n(ab)^n$ implies $i = 0$. This last property is a consequence of a more general result: let u and v be two words and a, b be two distinct letters then the system $\{bu \leftrightarrow av\}$ is *left simplifiable*, that means: for all words w, w', w'', if $ww' \overset{*}{\longleftrightarrow} ww''$ then $w' \overset{*}{\longleftrightarrow} w''$ (this can easily be proved for $w = a$ by induction on the length of the derivation $aw' \overset{*}{\longleftrightarrow} aw''$). Symmetrically, system S_5 is also right simplifiable. So, if $(ba)^{n-i}(ab)^{n+i} \overset{*}{\longleftrightarrow} (ba)^n(ab)^n$, it follows $(ab)^i \overset{*}{\longleftrightarrow} (ba)^i$ that implies $i = 0$. Finally, $S_5^{\leftrightarrow}((abba)^*) \cap (ba)^*(ab)^* = \{(ba)^n(ab)^n \mid n \ge 0\}$ and $S_5^{\leftrightarrow}$ is not a rational transduction.

Conversely, we think that the rationality of $S^{\leftrightarrow}$ implies the rationality of S when S is one-rule length-preserving rewrite system while it is not the case when S has several rules: let $S_6 = \{ba \to ab, b \to a\}$ then $S_6((ba)^*) \cap a^*b^* = \{a^nb^p \mid n \ge p \ge 0\}$, so S_6 is not rational while $S_6^{\leftrightarrow}$ is equivalent to the rational substitution s defined by $s(a) = s(b) = \{a, b\}$.

Conjecture 7.1. For all one-rule length-preserving rewrite system S, if $S^{\leftrightarrow}$ is a rational transduction then S is a rational transduction.

Another question about one-rule rewrite systems and regularity is the following: *if a rewrite system S does not preserve regularity, does it transform all regular language into a context-free language ?* Indeed, the simplest example of one-rule length-preserving rewrite system that is not rational is the system $S_6 = \{ba \to ab\}$ and it has been proved in [2], in the context of a particular class of rewrite systems called semi-commutations, that S_6 transform all regular language into a context-free language.

One could think that all one-rule rewrite system satisfies this property but it is not the case: consider the system $S_7 = \{ba \to a^2b^2\}$ then $S_7(b^2a^2)$ is not a context-free language([8]). Nevertheless, we think it is true when S is length-preserving:

Conjecture 7.2. A one-rule length-preserving rewrite system always transforms a regular language into a context-free language.

Acknowledgements. The first author would like to thank Alain Terlutte and Jacques Sakarovitch for fruitful discussions during the early steps of this work.

References

[1] J. Berstel, *Transductions and Context-Free Languages*. Teubner Verlag (1979).

[2] M. Clerbout and Y. Roos, Semi-commutations and algebraic languages, in *STACS 90*, in vol. 415, edited by Christian Choffrut and Thomas Lengauer. *Lect. Notes Comput. Sci.* Springer Berlin/Heidelberg (1990) 82–94.

[3] N. Dershowitz, Open. closed. open, in vol. 3467 of *Lect. Notes Comput. Sci. RTA*, edited by J. Giesl. Springer (2005) 376–393.

[4] S. Eilenberg and B. Tilson, Automata, languages and machines, vol. B, *Pure Appl. Math.* Academic Press, New York, San Francisco, London (1976).

[5] A. Geser, Termination of string rewriting rules that have one pair of overlaps, in vol. 2706, *Lect. Notes Comput. Sci. RTA*, edited by R. Nieuwenhuis. Springer (2003) 410–423.

[6] A. Geser, D. Hofbauer, and J. Waldmann, Match-bounded string rewriting systems. *Appl. Algebra Eng. Commun. Comput.* **15** (2004) 149–171.

[7] W. Kurth, *Termination und Konfluenz von Semi–Thue–Systemen mit nur einer Regel*. Ph.D. thesis, Technische Universitt Clausthal (1990).

[8] M. Latteux and Y. Roos, The image of a word by a one-rule semi-Thue system is not always context-free (2011). `http://www.lifl.fr/~yroos/al/one-rule-context-free.pdf`.

[9] É. Lilin, Une généralisation des semi-commutations. Technical Report IT-210, Laboratoire d'Informatique Fondamentale de Lille, Université de Lille 1, France (1991). In french.

[10] Y. Métivier, Calcul de longueurs de chaînes de réécriture dans le monoïde libre. *Theor. Comput. Sci.* **35** (1985) 71–87.

[11] R. Milner, The spectra of words, in vol. 3838 of *Processes, Terms and Cycles: Steps on the Road to Infinity*, edited by A. Middeldorp, V. van Oostrom, F. van Raamsdonk and R. de Vrijer. *Lect. Notes Comput. Sci.* Springer Berlin/Heidelberg (2005) 1–5.

[12] B. Ravikumar, Peg-solitaire, string rewriting systems and finite automata. *Theor. Comput. Sci.* **321** (2004) 383–394.

[13] J. Sakarovitch, *Elements of Automata Theory*. Cambridge University Press, New York, USA (2009).

[14] J. Sakarovitch and I. Simon, Subwords. In *Combinatorics on words*. Cambridge Mathematical Library. Cambridge University Press (1997) 105–142.

[15] A. Terlutte and D. Simplot, Iteration of rational transductions. *RAIRO: ITA* **34** (2000) 99–130.

[16] C. Wrathall, Confluence of one-rule thue systems, Word Equations and Related Topics, in vol. 572 of *Lect. Notes Comput. Sci.*, edited by K. Schulz. Springer Berlin/Heidelberg (1992) 237–246.

Communicated by Y. Métivier.
Received October 24, 2012. Accepted December 3, 2013.