Theoretical Computer Science 58 (1988) 81-101
North-Holland



# COUNTING WITH RATIONAL FUNCTIONS

C. CHOFFRUT
*Faculté des Sciences, Université de Rouen, 76130 Mont-Saint-Aignan, France*

M.P. SCHUTZENBERGER
*U.E.R. de Mathématiques et d'Informatique, Université de Paris VII 75251 Paris Cedex 05, France*

**Abstract.** Rational functions of a free monoid $A^*$ into the free cyclic monoid $t^*$ generated by a unique element $t$ can be viewed as assigning an integer to every word $u \in A^*$. We investigated those functions which count occurrences of some fixed (and special) subsets $X \subseteq A^*$ in all words of $A^*$ and show that they are exactly those which satisfy a Lipschitz condition relatively to some metric on the free monoid.

**Résumé.** Les fonctions rationnelles d'un monoïde libre $A^*$ dans le monoïde libre cyclique $t^*$ engendré par l'unique élément $t$ peuvent être considérées comme des applications qui à tout mot $u \in A^*$ associent un entier. Nous étudions plus précisément celles qui comptent les occurrences d'ensembles fixes (et particulier) $X \subseteq A^*$ dans chaque mot de $A^*$. Nous montrons que ce sont exactement celles qui vérifient une condition de Lipschitz pour une certaine métrique du monoïde libre.

## 1. Introduction

Rational functions of a free monoid $A^*$ into another $B^*$ are obtained by providing a finite, not necessarily deterministic automaton $\mathfrak{A}$ with an output function, thus associating a word in $B^*$ with every transition of $\mathfrak{A}$. Morphisms of $A^*$ into $B^*$ are special cases of such functions.

Numerous areas of computer science are directly dealing with rational functions: codes (encoding and decoding of messages), lexical analysers (assigning a token to some portion of a program), sorting (the Soundex encoding of surnames, cf. [8, p. 391], defines a special case of rational function [7]), text editing (systems like Multics or Unix provide a large range of commands substituting to all occurrences of some rational expression a given word) etc...

The importance of rational functions is also theoretical since they play, with respect to free monoids, a role similar to that of rational fractions with respect to complex numbers. In both cases these functions are directly defined from the structure on which they act (concatenation in the former case, addition and multiplication in the latter).

We are here concerned with a problem which has long been considered in automata theory, to wit what automata can possibly count. Refining the notion of threshold and modulo counting, various classes of rational languages were defined. In the

present paper we consider rational functions $\alpha$ of the input monoid $A^*$ into the free cyclic monoid $t^*$ generated by the single symbol $t$. Then for every $u \in A^*$ the length of its image by $\alpha$, i.e., $|u\alpha|$, is an integer and therefore "counts" something. More precisely, we are interested in characterizing those rational functions which count occurrences of some rational subset $H$ of $A^*$: we call such functions *counting*. A necessary condition on $H$ in order to ensure the linear growth of the image $u\alpha$ is that there do not exist an infinite chain (in the ordering "being-factor-of") of distinct words. This again is equivalent to $H$ being a finite union of rational *semaphores*, i.e., of rational subsets for which two elements may not be a proper factor of each other (cf. [2, Chapter II, 5]).

The characterization requires a notion close to continuity of functions as used in analysis. We say that a function $\alpha : A^* \to t^*$ satisfies the Lipschitz condition if there exists an integer $k > 0$ such that

$$-kn \leqslant |u\alpha| - |v\alpha| \leqslant kn$$

holds for all $u, v \in A^*$, where $n$ is the minimum number of letters which have to be erased in $u$ and $v$ in order to obtain a common subword. Then our main result (cf. Theorem 5.1) states that (after a possible partition of $A^*$ into finitely many rational subsets, which we ignore here for the simplicity of the exposition) counting and the Lipschitz condition are essentially the same notions for rational functions.

In Section 2, basics on rational functions and their transducers are presented. Some important distances based on the notions of prefix, factor and subword are defined on free monoids, and the Lipschitz condition, extending the classical notion of real metric spaces, is precisely stated. Section 3 establishes a characterization of the rational functions which satisfy the Lipschitz condition with respect to the subword distance in terms of their transducers. Section 4 deals with rational semaphores and counting functions. A few closure properties of these functions are proved and one implication of our theorem, to wit: "all counting functions satisfy the Lipschitz condition", is established. Section 5 is devoted to the main theorem which is proved by induction on the cardinality of the finite transition monoid underlying the transducer. We have gathered in the Appendix all the technical results concerning the combinatorics of words which may be omitted at first reading.

## 2. Preliminaries

### 2.1. Free monoids

Let $A$ be a finite nonempty set—or *alphabet*—: let $A^*$ and $A^+$ respectively be the free monoid and the free semigroup which it generates. An element $w$ of $A^*$ is a *word* and its *length* is denoted by $|w|$. The identity of $A^*$ or *empty* word is denoted by 1: $A^+ = A^* - \{1\}$. The elements of $A$ are called *letters*.

Given a factorization $w = w_1 w_2 w_3$ we say that $w_2$ is a *factor*, $w_1$ is a *prefix* and $w_3$ is a *suffix* of $w$. A *subword* $v$ of $w$ is a word obtained from $w$ by erasing some letters in $w$. Thus, *abb* is a subword of *aababa* (*aababa*).

Given a subset $X \subseteq A^+$, the number of occurrences of elements of $X$ in a word $w$ is denoted by $|w|_X$. In particular, we have $|w| = \sum_{a \in A} |w|_a$.

**Example 2.1.** If $A = \{a, b\}$ and $X = ab^*a$, then $|w|_X = \max(0, |w|_a - 1)$.

### *2.2. Rational relations*

Given arbitrary sets $X$ and $Y$, we consider a relation $\rho$ from $X$ to $Y$, denoted $\rho : X \to Y$, as an application of $X$ into the powerset of $Y$. Its *domain* is the subset $\mathrm{dom}\,\rho = \{x \in X \mid x\rho \neq \emptyset\}$ and its *graph* is the subset $\#\rho = \{(x, y) \in X \times Y \mid y \in x\rho\}$. We will consider any function $\alpha : X \to Y$ as a relation from $X$ to $Y$ where $x\alpha$ contains at most one element.

Assume now $X$ and $Y$ are respectively the free monoids $A^*$ and $B^*$. A relation $\rho : A^* \to B^*$ is *rational* if its graph is a rational subset of the product monoid $A^* \times B^*$ (cf., e.g., [5, p. 236]). Designating by $\mathrm{Rat}\,M$ the semiring of all rational subsets of an arbitrary monoid $M$, the rational relations are characterized as follows (cf., e.g., [1, Theorem 7.1] or [12]).

**Theorem 2.2.** *Let $\rho : A^* \to B^*$ be a relation. The following conditions are equivalent:*

(i) *$\rho$ is rational;*

(ii) *there exist a finite nonempty set $Q$, a morphism $\mu$ of $A^*$ into the multiplicative monoid $(\mathrm{Rat}\,B^*)^{Q \times Q}$ of $Q \times Q$-matrices with entries in $\mathrm{Rat}\,B^*$, and a $Q$-row and a $Q$-column vectors $\lambda$ and $\gamma$ with entries in $\mathrm{Rat}\,B^*$ such that $u\alpha = \lambda u \mu \gamma$ holds for all $u \in A^*$.*

We shall say that the triple $(\lambda, \mu, \gamma)$ or simply $\mu$ is a *transducer* realizing $\rho$. Then $Q$ is the set of *states* and its cardinality $|Q|$ is the *dimension* of the transducer.

From now on, unless otherwise stated, we only deal with rational relations which are functions. Then it can be easily verified that, without loss of generality, we may assume the following to hold:

**(1)** all entries in $\lambda$, $\gamma$ and $a\mu$ $(a \in A)$ are in $B^* \cup \{\emptyset\}$.

**(2)** all states $q \in Q$ are useful in the sense that there exist $u, v \in A^*$ such that

$$(\lambda u \mu)_q (v \mu \gamma)_q \neq \emptyset.$$

We define the *norm* $\|\mu\|$ of a transducer $\mu$ realizing a function, as the maximum length of all non-zero entries in the matrices $a\mu$, where $a \in A$. We set $\|\mu\| = 0$ whenever all $a\mu$ are the zero matrix.

It is convenient to consider a transducer $\mu$ as an ordinary finite automaton with outputs in $B^*$. Its transitions are the triples $(q, a, q')$ where $q, q' \in Q$, $a \in A$ and

$a\mu_{qq'} \neq \emptyset$. With such a triple the output $x = a\mu_{qq'}$ is associated. Then $u\mu_{qq'} = x$, where $u \in A^*$, $x \in B^*$, and $q, q' \in Q$, can be interpreted as saying that there exists a path from $q$ to $q'$ with label $u$ and output $x$.

We shall use the notation:

$$q \xrightarrow[x]{u} q' \quad \text{in place of} \quad u\mu_{qq'} = x$$

and, more generally,

$$q_0 \xrightarrow[x_1]{u_1} q_1 \xrightarrow[x_2]{u_2} q_2 \to \cdots \xrightarrow[x_r]{u_r} q_r$$

will stand for the product:

$$u_1 u_2 \ldots u_r \mu_{q_0 q_r} = u_1 \mu_{q_0 q_1} \ldots u_r \mu_{q_{r-1} q_r} = x_r \ldots x_r.$$

We shall need two operations on the class of functions of $A^*$ into $B^*$. Given two functions $\alpha, \beta : A^* \to B^*$ having disjoint domains, we define their *disjoint union* as the function $\gamma : A^* \to B^*$ whose graph is the union of the graphs of $\alpha$ and $\beta$: $\#\gamma = \#\alpha \cup \#\beta$.

Now, for arbitrary, not necessarily rational functions $\alpha, \beta : A^* \to B^*$, we define their *product* as the function $\gamma : A^* \to B^*$ satisfying $u\gamma = u\alpha u\beta$ for all $u \in A^*$. In particular, the domain of $\gamma$ is the intersection of the domains of $\alpha$ and $\beta$. Beware that our definition differs from the usual product of composition of relations and from the componentwise product of relations.

**Example 2.3.** Let $A = \{a\}$, $B = \{a, b\}$ and consider the two functions defined by $a^n\alpha = a^n b$ and $a^n\beta = ba^n$. The componentwise product of $\alpha$ and $\beta$ is not a function since it assigns $\{a^i b^2 a^{n-i} \mid 0 \leq i \leq n\}$ to $a^n$.

### *2.3. Distances—Lipschitz functions*

Based on the notions of prefix, factor and subword, three distances over $A^*$ may be defined by setting:

$$d_i(u, v) = |u| + |v| - 2L_i(u, v) \quad i = 1, 2, 3$$

where $L_1(u, v)$ (respectively $L_2(u, v)$, $L_3(u, v)$) is the maximum length of a prefix (respectively factor, subword) common to $u$ and $v$.

Indeed, let us verify the triangular inequality $d_i(u, v) \leq d_i(u, w) + d_i(w, v)$. We may consider a word $u \in A^*$ as a mapping of the interval $[1, |u|]$ into $A$ and denote by $u(i)$ the $i$th occurrence of $u$: $u = u(1) \ldots u(|u|)$. For an arbitrary subset $I \subseteq [1, |u|]$ we set $u(I) = u(i_1) \ldots u(i_r)$, where $1 \leq i_1 < \cdots < i_r \leq |u|$ is the set of elements of $I$.

For $i = 1, 2, 3$ the triangular inequality is equivalent to $L_i(u, w) + L_i(w, v) \leq |w| + L_i(u, v)$. Let $I$, $J$ be two subsets of $[1, |w|]$ and $x = w(I)$, $y = w(J)$ be the two subwords of $w$ satisfying $L_i(u, w) = |x|$ and $L_i(w, v) = |y|$. Then we obtain

$$L_i(u, w) + L_i(w, v) = |x| + |y| = \operatorname{card}(I \cup J) + \operatorname{card}(I \cap J).$$

Since $z = w(I \cap J)$ is a subword of $x$ and $y$ we have $\text{card}(I \cap J) = |z| \leq L_i(u, w)$. Furthermore, $\text{card}(I \cap J) \leq |w|$, thus completing the verification.

We say that $d_1$ (respectively $d_2$, $d_3$) is the *prefix-* (respectively *factor-*, *subword-*) distance. The following inequalities are straightforward:

$$d_3(u, v) \leq d_2(u, v) \leq d_1(u, v), \tag{3}$$

$$d_3(xuy, zvt) - |xyzt| \leq d_3(u, v) \leq d_3(xuy, zvt) + |xyz|, \tag{4}$$

$$\text{if } A \text{ is reduced to one letter, then } d_1(u, v) = d_2(u, v) = d_3(u, v). \tag{5}$$

The following technical result shows that if two words are close relatively to the subword distance, then they have a large common factor.

**Lemma 2.4.** *Let $u, v \in A^*$ be two words such that $|u| + |v| = L$ and $d_3(u, v) = k$. then $u$ and $v$ have a common factor of length $(L-k)/(2(k-1))$.*

**Proof.** Assume $u = u_0 w_1 u_1 \dots w_r u_r$ and $v = v_0 w_1 v_1 \dots w_r v_r$ where $w_1 \dots w_r$ is a maximal subword common to $u$ and $v$, i.e.,

$$d_3(u, v) = |u_0 \dots u_r| + |v_0 \dots v_r| = k.$$

Let $w_i$ satisfy $|w_i| \geq |w_j|$ for $j = 1, \dots, r$. Then we have

$$L = |u| + |v| \leq k + 2rF \quad \text{where } F = |w_i|.$$

Thus, $F \geq (L-k)/2r$. Since $u_j v_j \neq 1$ holds for $j = 0, \dots, r$, we have $k = \sum_{0 \leq j \leq r} |u_j v_j| \geq r+1$ which yields $(L-k)/(2(k-1))$. □

The previous notions of distances are meant to help studying the functions of a free monoid $A^*$ into another $B^*$.

Assume $A^*$ and $B^*$ are equipped with the distances $d_i$ and $d_j$ respectively, $i, j \in \{1, 2, 3\}$. Then we say that $\alpha : A^* \to B^*$ is a *Lipschitz function* whenever there exists an integer $k > 0$ satisfying $d_j(u\alpha, v\alpha) \leq k d_i(u, v)$ for all $u, v \in \text{dom } \alpha$. The next section will give an example of such functions.

### *2.4. Subsequential functions*

Among the class of rational functions, particularly important are the subsequential functions introduced in [14] where they were shown to satisfy a noticeable functional equation. *Subsequential* functions are realized by transducers $(\lambda, \mu, \gamma)$ where $\lambda$ has all entries equal to 0 except one equal to 1, and where all $a\mu$ for $a \in A$ are row monomial, i.e., have at most one non-zero entry in each row. These functions are a natural generalization of the *sequential* functions studied by Ginsburg and Rose (cf. [6]). Indeed, a subsequential transducer is *sequential* if all entries of $\gamma$ are 0 or 1. Taking advantage of the "monomiality" of the matrices, the following more concise notations are useful.

We first define a *transition* function $Q \times A^* \to Q$ by setting

$$q.u = q' \text{ iff } u\mu_{qq'} \neq 0.$$

Similarly, we define an *output* function $Q \times A^* \to B^*$ by setting

$$q * u = \begin{cases} u\mu_{qq'} & \text{if for some } q' \in Q,\ q.u = q', \\ 0 & \text{otherwise.} \end{cases}$$

It can be verified easily that the transition and output functions are perfectly determined by their values on $Q \times A$ and the induction rules for all $q \in Q$, $u \in A^*$ and $a \in A$:

(i) $q.1 = q$,
(ii) $q * 1 = 1$,
(iii) $q.ua = (q.u)a$,
(iv) $q * ua = (q * u)((q.u) * a)$.

Denoting by $q_-$ the index of the non-zero entry of $\lambda$ we have for all $u \in A^*$

$$\lambda u \mu \gamma = (q_- * u)\gamma_{q_-.u}.$$

The following characterization of subsequential functions will be useful (cf. [4]).

**Theorem 2.5.** *A function $\alpha : A^* \to B^*$ is subsequential iff the following two conditions hold:*

(1) *for all $L \in \mathrm{Rat}\, B^*$ we have $L\alpha^{-1} \in \mathrm{Rat}\, A^*$;*
(2) *$\alpha$ is a Lipschitz function for the prefix distance.*

## 3. Transducers of Lipschitz functions

In this section, unless otherwise stated, we shall assume that all free monoids are equipped with the subword distance which we shall denote by $d$ and to which the term Lipschitz refers. Furthermore, all transducers are supposed to satisfy conditions (1) and (2).

We shall establish the following characterization of the transducers realizing rational Lipschitz functions.

**Theorem 3.1.** *Let $\alpha : A^* \to B^*$ be a rational function realized by a transducer $\mu$. Then the following conditions are equivalent:*

(i) *$\alpha$ is a Lipschitz function;*
(ii) *there exists an integer $k > 0$ such that for all $w \in A^*$ and for any two entries $x, x' \in B^*$ of the matrix $w\mu$ we have $d(x, x') < k$;*
(iii) *for all $w \in A^*$ and for all diagonal entries $x = w\mu_{qq} \in B^*$, $x' = w\mu_{q'q'} \in B^*$, the words $x$ and $x'$ are conjugate, i.e., $xy = yx'$ for some $y \in B^*$.*

**Proof.** By standard arguments akin to Eilenberg's normalization procedure (cf. [5, p. 138]), we may assume that $\mathrm{dom}\, \alpha \subseteq A^+$ and that for some $q_- \neq q_+ \in Q$ we have $w\alpha = w\mu_{q_-q_+}$ for all $w \in A^*$.

(i) $\Rightarrow$ (ii): Let $w\mu_{qp} = x$ and $w\mu_{q'p'} = x'$ be two entries different from $\emptyset$. There exist words $u$, $u'$, $v$, $v' \in A^*$ of length less than or equal to $|Q|$ and words $z$, $z'$, $t$, $t' \in B^*$ such that

$$\left.\begin{aligned} u\mu_{q_-q} &= z, & v\mu_{pq_+} &= t, \\ u'\mu_{q_-q'} &= z', & v'\mu_{p'q_+} &= t'. \end{aligned}\right\} \tag{6}$$

Then we have

$$d(zxt, z'x't') \leqslant kd(uwv, u'wv') \leqslant 4k|Q|.$$

By condition (4) this implies $d(x, x') \leqslant 4|Q|(k + \|\mu\|)$.

(ii) $\Rightarrow$ (i): Let $u$, $v \in A^*$ be two words of $\operatorname{dom}\alpha$ and let $u = u_0 w_1 u_1 \dots w_r u_r$, $v = v_0 w_1 u_1 \dots w_i v_r$ be two factorizations such that $w_1 \dots w_r$ is a common subword of maximal length. Consider the two paths:

$$q_- \xrightarrow[x_0]{u_0} q_1 \xrightarrow[z_1]{w_1} p_1 \xrightarrow[x_1]{u_1} q_2 \to \cdots \to q_r \xrightarrow[z_r]{w_r} p_r \xrightarrow[x_r]{u_r} q_+,$$

$$q_- \xrightarrow[y_0]{v_0} q'_1 \xrightarrow[t_1]{w_1} p'_1 \xrightarrow[y_1]{v_1} q'_2 \to \cdots \to q'_r \xrightarrow[t_r]{w_r} p'_r \xrightarrow[y_r]{v_r} q_+.$$

By (4) we have

$$d(u\alpha, v\alpha) \leqslant |x_0 \dots x_r| + \sum_{1 \leqslant i \leqslant r} d(z_i, t_i).$$

Since $u_i v_i \neq 1$, $i = 0, \dots, r$, we obtain

$$d(u, v) = |u_0 \dots u_r| + |v_0 \dots v_r| = \sum_{0 \leqslant i \leqslant r} |u_i v_i| \geqslant r + 1,$$

which yields

$$d(u\alpha, v\alpha) \leqslant (\|\mu\| + k) d(u, v) - k.$$

(ii) $\Rightarrow$ (iii): By hypothesis, using notations (6) with $q = p$ and $q' = p'$, for all integers $i > 0$ we have $d(zx^i t, z'x''^i t') < k$. In view of (6) this implies $d(x^i, x''^i) < k + |zz'tt'|$. In virtue of Lemma 2.4 for some large enough $i$, $x^i$ and $x''^i$ have a common factor greater than or equal to $|x| + |x'|$, which by Proposition A.5 of the Appendix implies that $x$ and $x'$ are conjugate.

(iii) $\Rightarrow$ (ii): Assume we have $w\mu_{qp} = x \in B^*$ and $w\mu_{q'p'} = x' \in B^*$. Intuitively, what we want to prove is that there exist two paths leading from $q$ to $p$ and from $q'$ to $p'$ labelled by $w$ and admitting the same factorization of their label, such that almost all occurrences of $w$ belong to a loop.

More formally, we claim that there exist an integer $r > 0$ and a factorization of $w = u_0 v_1 u_1 \dots v_r u_r$ such that

$$|u_0 \dots u_r| < |Q^2|, \tag{7}$$

and that there exist two paths

$$q = q_0 \xrightarrow[z_0]{u_0} q_1 \xrightarrow[t_1]{v_1} q_1 \xrightarrow[z_1]{u_1} q_2 \to \cdots \to q_r \xrightarrow[t_r]{v_r} q_{r+1} = p,$$

$$q' = q'_0 \xrightarrow[z'_0]{u_0} q'_1 \xrightarrow[t'_1]{v_1} q'_1 \xrightarrow[z'_1]{u_1} q'_2 \to \cdots \to q'_r \xrightarrow[t'_r]{v_r} q'_{r+1} = p'.$$

Indeed, assume (7) is not verified. Then there exists a factorization

$$u_0 u_1 \ldots u_r = w_0 w_1 w_2 \quad \text{where } w_1 \neq 1$$

and two paths (omitting the outputs):

$$q \xrightarrow{w_0} s \xrightarrow{w_1} s \xrightarrow{w_2} p,$$

$$q' \xrightarrow{w_0} s' \xrightarrow{w_1} s' \xrightarrow{w_2} p'.$$

This yields a factorization

$$u = u'_0 v'_0 \ldots v'_{r'} v'_{r'} \quad \text{with } |u'_0 \ldots u'_{r'}| < |u_0 \ldots u_r|$$

and we may conclude by minimality.

Now, observe that by Theorem A.1 of the Appendix there exists an integer $k > 0$ such that $t\mu_{qq} = x$, $t\mu_{q'q'} = x'$ imply $d(x, x') < k$. Furthermore, because of $|u_i| \neq 0$ for $i = 1, \ldots, r-1$, we have $r < |Q^2| + 1$. Then we compute

$$d(z_0 t_0 z_1 \ldots t_r z_r, z'_0 t'_0 z'_1 \ldots t'_r z'_r) \leq |z_0 \ldots z_r| + |z'_0 \ldots z'_r| + \sum_{1 \leq i \leq r} d(t_i, t'_i)$$

$$< 2\|\mu\| |Q^2| + k(|Q^2| + 1),$$

completing the proof. □

As a consequence we have the following proposition.

**Proposition 3.2.** *Let $\alpha : A^* \to B^*$ be a rational function. Assume $B^*$ is equipped with the subword distance. Then $\alpha$ is a Lipschitz function when $A^*$ is equipped with the subword distance iff it is a Lipschitz function when $A^*$ is equipped with the factor distance.*

**Proof.** In view of (3) we may only prove that the condition is necessary. Let $\mu$ be a transducer realizing $\alpha$ and satisfying the same conditions as those of the previous theorem. Consider a word $w \in A^*$ and two entries $w\mu_{qp} = x$, $w\mu_{q'p'} = x'$. Then there exist two paths

$$q_- \xrightarrow[y]{u} q \xrightarrow[x]{w} p \xrightarrow[z]{v} q_+, \qquad q_- \xrightarrow[y']{u'} q' \xrightarrow[x']{w'} p' \xrightarrow[z']{v'} q_+,$$

where $u$, $u'$, $v$, $v'$ have length less than or equal to $|Q|$. By hypothesis, there exists an integer $k > 0$ depending only on $4|Q|$ such that

$$d_3(yxz, y'x'z') \leq d_2(yxz, y'x'z') \leq k,$$

i.e.,

$$d_3(x, x') \leq 4|Q| \|\mu\| + k$$

which, by condition (ii) of the previous theorem, completes the proof. □

## 4. Counting functions

### *4.1. Semaphores*

By a *semaphore* we mean a subset $H \subseteq A^*$ containing none of its proper factors $H \cap (A^*HA^+ \cap A^+HA^*) = \emptyset$. Thus, a semaphore is a biprefix code.

For any integer $k \geq 1$ we define the semaphore $H^{(k)}$ consisting of all the words starting and ending in $H$ and having exactly $k$ occurrences of $H$. Formally, we first introduce the family $L_k$, $k > 0$, by setting

$$L_1 = HA^* \cap A^*H \quad \text{and} \quad L_{k+1} = HA^* \cap A^+L_k \quad k \geq 1.$$

Then we have $H^{(k)} = L_k - L_{k+1}$ thus showing that $H^{(k)}$ is rational if $H$ is. Clearly, $H^{(1)} = H$.

**Example 4.1.** If $H = \{a\}$, then $H^{(k)} = [a(A-a)^*]^{k-1}a$. If $H = \{aa\}$, then $w \in H^{(k)}$ iff $w$ has a factorization

$$w = a^{n_0}u_1a^{n_1} \ldots u_ra^{n_r}$$

where $u_i \in (A-a)A^* \cap A^*(A-a) - A^*a^2A^*$ for $i = 1, \ldots, r$, $n_i \geq 2$ for $i = 0, \ldots, r$ and $n_0 + \cdots + n_r = k + r + 1$.

The following trivial statements will be useful later on:

$$|w|_{H(k)} = \max(0, |w|_H - k + 1), \tag{8}$$

$$|u|_H \leq |xuy|_H \leq |x| + |u|_H + |y|. \tag{9}$$

As a consequence of (8), assume $u$ is a maximal factor common to $w_1$ and $w_2$: $w_1 = x_1uy_1$ and $w_2 = x_2uy_2$. Then we have

$$-d_2(w_1, w_2) \leq |w_1|_H - |w_2|_H \leq d_2(w_1, w_2). \tag{10}$$

### *4.2. Counting functions*

From now on we assume that $B$ consists of the unique element $t$: $B = \{t\}$.

A function $\alpha : A^* \to t^*$ with rational domain $X$ is an *elementary counting* function if one of the following two conditions is satisfied:

**(11)** there exists an $s \in \mathbb{N}$ such that $|w\alpha| = s$ for all $w \in X$;

**(12)** there exist a rational semaphore $H \in A^*$ and a rational number $r > 0$ such that $|w\alpha| = r|w|_H$ for all $w \in X$. Furthermore, it is required that $\max\{|w|_H \mid w \in X\} = \infty$.

In the second case we say that $\alpha$ *counts* $H$.

A function $\alpha: A^* \to t^*$ with rational domain $X$ is a *counting* function if there exists an integer $n>0$ and, for $i=1,\ldots,n$, there exist a rational semaphore $H_i \in$ Rat $A^*$ and a rational number $r_i \in Q$, and if there exists a partition $X = X_1 \cup \cdots \cup X_m$ of $X$ into $m$ rational subsets and, for $j=1,\ldots,m$, there exists a rational number $s_j \in Q$ such that, for all $w \in X_j$, we have

$$|w\alpha| = s_j + \sum_{1 \leq i \leq n} r_i |w|_{H_i}.$$

**Example 4.2.** Assume $A=\{a, b\}$ and on the subset $X=\{w \in A^* \,||w|_a \equiv 1[3]$ and $|w|_b \equiv 1[2]\}$ consider the function $\alpha: A^* \to t^*$ defined by

$$|w\alpha| = \tfrac{1}{3}|w|_a + \tfrac{1}{2}|w|_b - \tfrac{5}{6}.$$

Then $\alpha$ is counting function.

The following result shows how the elementary counting functions generate all counting functions.

**Proposition 4.3.** *Every counting function is a finite disjoint union of products of elementary counting functions.*

**Proof.** Clearly, it suffices to consider the case where, in the previous definition, $m=1$. For all $w \in X$ we have

$$|w\alpha| = s + \sum_{1 \leq i \leq n} r_i |w|_{H_i}, \tag{14}$$

where $s, r_i \in Q$.

Without loss of generality we may assume

$$r_i > 0, \quad i=1,\ldots,n. \tag{15}$$

Indeed, set $I=\{1 \leq i \leq n \mid r_i < 0\}$. Then, for all $i \in I$, $|w|_{H_i} \leq K$ for some fixed integer $K$. We may partition $X$ into finitely many rational subsets over which $|w|_{H_i}$ is a constant $0 \leq \vartheta_i \leq K$; i.e., over which we have

$$|w\alpha| = \left(s + \sum_{i \in I} r_i \vartheta_i\right) + \sum_{i \notin I} r_i |w|_{H_i}.$$

If $r_i = 0$ for all $i \notin I$, then $\alpha$ is a constant. Otherwise, we may assume $r_i > 0$ if $i \notin I$, possibly after deleting some $H_i$'s.

Set $s = s'/N$ and $r_i = r_i'/N$ where $s' \in \mathbb{Z}$ and $N, r_i'$ are positive integers. Without loss of generality we may add to condition (15):

$$|w|_{H_i} \equiv 0 \bmod N. \tag{16}$$

Then $X$ may be partitioned into finitely many rational subsets over which $|w|_{H_i} \equiv \lambda_i$ for some fixed $0 \leqslant \lambda_i < N$, $i = 1, \ldots, n$. Over each of these subsets, by (8), we have

$$|w|_{H_i}(\lambda_i + 1) \equiv 0 \bmod N.$$

Under condition (16) $s$ is an integer. It finally suffices to prove that we may assume that $s$ is positive. If $n = 1$, then (14) reduces to

$$|w\alpha| = s + r_1|w|_{H_1}$$

Let $\lambda N$ be the least integer such that $s + \lambda r_1 N = s' \geqslant 0$. Then we have $|w|_{H_1} \geqslant N\lambda$ for all $w \in X$; thus,

$$|w\alpha| = s' + r_1(|w|_{H_1} - \lambda N) = s' + r_1|w|_{H_1}(\lambda N + 1).$$

More generally, for an arbitrary $n > 1$ we have

$$|w\alpha| = (s + r_1|w|_{H_1}) + \sum_{2 \leqslant i \leqslant n} r_i|w|_{H_i}.$$

Let $\lambda N$ be the greatest integer (if it exists) such that $s + \lambda r_1 N < 0$. By induction hypothesis, the result holds for all restrictions of $X$ to the subsets of words containing $\vartheta N$ $(0 \leqslant \vartheta \leqslant \lambda)$ occurrences of $H_1$. Over the subset of words $w$ containing more than $\lambda N$ occurrences of $H_1$ we have

$$|w\alpha| = (s + (\lambda + 1)r_1 N) + r_1|w|_{H_1}(\lambda + 1)N + 1 + \sum_{2 \leqslant i \leqslant n} r_i|w|_{H_i}. \qquad \square$$

**Corollary 4.2.** *Let $\alpha : A^* \to t^*$ have finite image. Then $\alpha$ is rational iff $\alpha$ is a counting function.*

**Proof.** Assume $A^*\alpha = \{x_1, \ldots, x_n\} \subseteq t^*$. If $\alpha$ is rational, then $x_i\alpha^{-1} = X_i \in \operatorname{Rat} A^*$ for $i = 1, \ldots, n$ (cf., e.g., [1, Corollary 4.2]). Then $\alpha$ is the finite union of the constant functions

$$w\alpha_i = x_i \quad \text{for all } w \in X_i.$$

Conversely, if $\alpha$ is a counting function, by the previous proposition, it is a finite union of constant functions with rational domain. Moreover, such functions are rational since their graphs are of the form

$$X \times x \in \operatorname{Rat} A^* \times t^* \quad (X \in \operatorname{Rat} A^*, x \in t^*). \qquad \square$$

The following result proves our main Theorem 5.1 in one direction. Because of Proposition 3.2 the term "Lipschitz" refers indifferently to the subword or to the factor distance.

**Proposition 4.5.** *If $\alpha : A^* \to t^*$ is a counting function, then it is a rational Lipschitz function.*

**Proof.** Let $w_1, w_2$ belong to the domain of $\alpha$. For some maximal factor $u$ common to $w_1$ and $w_2$, we have $w_1 = x_1 u y_1$ and $w_2 = x_2 u y_2$. Then by (10) and (13) we obtain

$$d(w_1\alpha, w_2\alpha) \leq s + rd(w_1, w_2),$$

where $r = \sum_{1 \leq i \leq n} |r_i|$ and $s = \max\{|s_j| \mid j = 1, \ldots, m\}$.

In order to prove that $\alpha$ is rational, let us first verify that all elementary counting functions are rational. By Corollary 4.4, it suffices to consider the case $w\alpha = r|w|_H$ as in (12). We first prove a lemma (a subset $X$ is *suffix* if $A^+X \cap X = \emptyset$).

**Lemma 4.6.** *Let $H \in \mathrm{Rat}\, A^*$ be suffix. Then there exists a rational function $\beta : A^* \to t^*$ such that $|w\beta| = |w|_H$.*

**Proof.** Let $\mathfrak{A} = (Q, i, T)$ be the minimal automaton recognizing the left ideal $A^*H$. We transform it into a sequential transducer by defining

$$q^*a = \begin{cases} t & \text{if } q.a \in T, \\ 1 & \text{otherwise.} \end{cases}$$

The resulting rational function $\alpha : A^* \to t^*$ satisfies $|w\beta| = \mathrm{card}\{u \in A^*H \mid w \in uA^*\}$. Since $H$ is suffix, this last integer equals $|w|_H$. □

**Proof of Proposition 4.5** (*continued*). We now return to the proof of Proposition 4.5. If $r = p/n$ where $p, n \in \mathbb{N}$, then the previous lemma shows that $\alpha$ is the composition of the three rational functions $\beta$ (as in the lemma ) and $\gamma, \delta : t^* \to t^*$ respectively defined by their graphs $\#\gamma = (t^n, t)^*$ and $\#\delta = (t, t^p)^*$. Since rational relations are closed under composition (cf., e.g., [1, Theorem 4.4.), any elementary counting function is rational.

Now, because of the characterization of Theorem 2.2, a rational function $\alpha : A^* \to t^*$ may be viewed as a rational series in the noncommutative unknowns. $A$ over the commutative semiring Rat $t^*$. The Hadamard product of such series (corresponding to the product of functions defined in Section 2.2.) is a rational series, thus a rational function (cf., e.g., [1, 2] or [3, Theorem 1, p. 21]). Then the result follows from Proposition 4.3. □

### *4.3. Some closure properties of counting functions*

We are mainly concerned here with closure properties of counting functions under certain compositions.

**Proposition 4.7.** *Let $A = A_1 \cup A_2$ be a partition and $\pi$ the projection of $A^*$ onto $A_1^*$. If $\beta : A_1^* \to t^*$ is a counting function then $\alpha = \pi\beta : A^* \to t^*$ is itself a counting function.*

**Proof.** By Proposition 4.3, we may assume, without loss of generality, that $w\beta = r|w|_H + s$ holds for all $w \in X = \mathrm{dom}\,\alpha$, with $r \in Q_+$, $s \in \mathbb{N}$ and $H \in \mathrm{Rat}\, B^*$ a

semaphore. Set $K = H\pi^{-1} - (A_2^* A^* \cup A^* A_2)$. Clearly, $K$ is a rational semaphore. Since $|w|_K = |w\pi|_H$ holds for all $w \in A^*$, we obtain

$$|w\alpha| = |w\pi\beta| = r|w\beta|_H + s = r|w|_K + s \qquad \square$$

**Proposition 4.8.** *Let $A$, $B$ be two alphabets and $\beta : B^* \to t^*$ a counting function. Consider a partition $A = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$, and a surjective mapping $\gamma$ of $Y = A_1^* A_2$ onto $B$ such that $b\gamma^{-1} \in \mathrm{Rat}\, A^*$ for all $b \in B$. Extend $\gamma$ to $A^*$ by setting $w\gamma = w_1\gamma \ldots w_n\gamma$ if $n > 0$, $w = w_1 \ldots w_n w_{n+1}$, $w_i \in Y$, for $i = 1, \ldots, n$ and $w_{n+1} \in A_1^*$. Then $\alpha = \gamma\beta : A^* \to t^*$ is itself a counting function.*

**Proof.** As in the previous proposition, we may assume, without loss of generality, that $w\beta = r|w|_H + s$ holds (with the same meaning for $r$, $s$ and $H$). The subset $K = A_2(H\gamma^{-1})$ is a rational semaphore. We define

$$X_1 = (H\gamma^{-1})A^* \cap X \text{ and } X_2 = X - (H\gamma^{-1})A^*.$$

Then we have

$$|u\gamma|_H = \begin{cases} |u|_K + 1 & \text{if } u \in X_1, \\ |u|_K & \text{if } u \in X_2. \end{cases}$$

This yields

$$|u\alpha| = |u\gamma\beta| = \begin{cases} r|u\gamma|_H + s = r|u|_K + s + 1 & \text{if } u \in X_1, \\ r|u\gamma|_H + s = r|u|_K + s & \text{if } u \in X_2. \end{cases} \qquad \square$$

**Proposition 4.9.** *Let $A = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$ be a partition and $\beta : A_1^* \to t^*$ a counting function. Define $\alpha : A^* \to t^*$ by $w\alpha = u_1\beta \ldots u_n\beta$ where $n > 0$, $u_1, \ldots, u_n \in A_1^*$ and $w \in u_1 A_2 u_2 A_2 \ldots u_{n-1} A_2 u_n$. Then $\alpha$ is a counting function.*

**Proof.** Clearly, if $X$ is the domain of $\beta$, then $(XA_2)^*X$ is the domain of $\alpha$. We adopt the notations of the definition of a counting function. Thus the following holds:

$$|w\alpha| = \sum_{1 \leq i \leq n} r_i |w|_{H_i} + \sum_{1 \leq j \leq m} s_j |w|_{A_2 X_j A_2} + \lambda,$$

where

$$\lambda = \begin{cases} s_j & \text{if } w \in X_j \cup X_j A_2 \cup A_2 X_j, \\ s_j + s_k & \text{if } w \in X_j (A_2 X)^* A_2 X_k. \end{cases}$$

It then suffices to observe that $A_2 X_j A_2$ is a rational semaphore. $\square$

**Proposition 4.10.** *If $\alpha$, $\alpha' : A^* \to t^*$ are counting functions, so is their product $u\beta = u\alpha u\alpha'$.*

**Proof.** Without loss of generality we may assume that $\mathrm{dom}\, \alpha = \mathrm{dom}\, \alpha' = X$, and that both functions admit the same decomposition $X = \bigcup_{1 \leq j \leq m} H_j$. If $\alpha$, $\alpha'$ satisfy

for all $w \in X_j$,

$$|w\alpha| = \sum_{1 \leq i \leq n} r_i |w|_{H_i} + s_j, \qquad |w\alpha'| = \sum_{1 \leq i \leq n'} r_i |w|_{H'_i} + s'_j,$$

then we obtain

$$|w\beta| = \left( \sum_{1 \leq i \leq n} r_i |w|_{H_i} + \sum_{1 \leq i \leq n'} r_i |w|_{H'_i} \right) + (s_j + s'_j). \qquad \square$$

## 5. The main theorem

The purpose of this section is to establish our main theorem which characterizes the counting functions of $A^*$ into a free cyclic monoid, i.e., a free monoid generated by a single element $t$. Because of (5) and Proposition 3.2, the term "Lipschitz" refers indifferently to the factor or the subword distance.

**Theorem 5.1.** *A function $\alpha : A^* \to t^*$ is counting iff it is a rational Lipschitz function.*

Besides the results of Section 4, the proof requires further preliminary results.

### *5.1. A congruence of finite index*

We first prove the following proposition.

**Proposition 5.2.** *If a function $\alpha : A^* \to t^*$ is a rational Lipschitz function, then it is subsequential.*

**Proof.** It suffices to verify the two conditions of Theorem 2, 5. The first one is a general property of rational relations (cf., e.g., [1], Corollary 4.2]). The second condition follows from (3). $\square$

Let us denote by $\mathcal{M}$ the monoid of row monomial $Q \times Q$-matrices with entries in $t^* \cup \{\emptyset\}$. Because of the previous result, a rational Lipschitz function $\alpha : A^* \to t^*$ can be realized by a subsequential transducer $(\lambda, \mu, \gamma)$ where $\mu : A^* \to \mathcal{M}$.

We introduce the following notation. For all non-zero matrices $m \in \mathcal{M}$, $m\rho$ denotes the shortest non-zero entry of $m$ and we define the matrix $m\pi$ by the equality $m = m\rho m\pi$. The following identities are straightforward:

$$m_1 m_2 \pi = (m_1 \pi m_2 \pi)\pi, \tag{17}$$

$$m_1 m_2 \rho = m_1 \rho m_2 \rho (m_1 \pi m_2 \pi)\rho. \tag{18}$$

As a consequence, given a morphism $\mu : A^* \to \mathcal{M}$, the relation $u \sim_\mu v$ (or, more simply, $u \sim v$ when $\mu$ is understood), defined for all $u, v \in A^*$ such that $u\mu\pi = v\mu\pi$, is a congruence.

We denote by $\sigma$ the morphism of $A^*$ into the monoid of $Q \times Q$-matrices with boolean entries which to every $u \in A^*$ assigns its *support*:

$$u\sigma_{qq'} = \begin{cases} 1 & \text{if } u\mu_{qq'} \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have the following proposition.

**Proposition 5.3.** *Let $\alpha : A^* \to t^*$ be a rational function realized by some subsequential transducer $\mu$. Then $\alpha$ is a Lipschitz function iff the congruence $\sim_\mu$ has finite index.*

**Proof.** Assume $\alpha$ is a Lipschitz function. Then, by Ramsey's Theorem, for some integer $N$ all words $w \in A^*$ of length greater than or equal to $N$ can be factorized into $w = w_1 w_2 w_3 w_4$, where $|w_2||w_3| \neq 0$ and $w_2\sigma = w_3\sigma$ is an idempotent. Because of condition (iii) of Proposition 3.1, there exists $k > 0$ such that $w_2 w_3 \mu = t^k . w_2 \mu$, i.e., $w_1 w_2 w_3 w_4 \mu = t^k w_1 w_2 w_4 \mu$. This implies $w \sim_\mu w_1 w_2 w_4$, thus showing that the congruence has finite index.

Conversely, assume $\alpha$ is not a Lipschitz function. By theorem 3.1, there exist two indices $q, q' \in Q$, and a word $u \in A^*$ such that $u\mu_{qq} \neq u\mu_{q'q'}$. Then all $u^n \mu \pi$, $n > 0$, are different. □

*5.2. Proof of the theorem*

In view of Proposition 4.5, it suffices to prove that every rational Lipschitz function is a counting function. Thus let $\alpha : A^* \to t^*$ be a rational Lipschitz function. By Proposition 5.2, $\alpha$ is realized by some subsequential transducer $(\lambda, \mu, \gamma)$. As in Section 2.4, with $\mu$ we associate its transition and output functions denoted by $(q, u) \to q.u$ and $(q, u) \to q * u$ respectively.

Furthermore, given any $u \in A^*$, we denote by $u\sigma$ the support of the matrix $u\mu$ as defined in the previous section. The semigroup $S = A^*\sigma$ may be viewed as acting on the set $Q$. Indeed if $x \in S$, then $q.x = q'$ iff $q.u = q'$ for some $u\sigma = x$.

With the notations of the previous section we have the identity $u\alpha = u\mu\rho\ (\lambda u \mu \pi \gamma)$. Since, by Proposition 5.3, $\lambda u \mu \pi \gamma$ is a rational function of finite image, $\alpha$ is a counting function iff $\mu\rho$ is.

Our result thus amounts to proving the following claim:

**(19)** *if $\mu : A^* \to \mathcal{M}$ satisfies condition* (iii) *of Theorem* 3.1, *then $u \to u\mu\rho$ is a counting function.*

We shall prove (19) by induction on the cardinality of the semigroup $S = A^*\sigma$ via the following result due to Krohn and Rhodes (cf., e.g., [9, Lemma 7.2.7]).

**Proposition 5.4.** *Given a morphism $\sigma$ of $A^*$ into a finite semigroup $S$, one of the following cases occurs*:

(1) *$S$ is cyclic*;

(2) *$S$ consists of a unique $\mathcal{L}$-class and (possibly) the identity*;

(3) *there exists a partition $A = A_1 \cup A_2$ such that $(A_1^* A_2)^*\sigma$ and $A_1^*\sigma$ are proper subsemigroups of $S$.*

**Proof** (*Basis of the induction*): We first prove our result under the following hypothesis:

**(20)** *all elements of S have the same minimal image* $P \subseteq Q$: $\forall x \in S, \lim_{n\to\infty} Qx^n = P$.

Let $I \subseteq A^*$ be the right ideal of all words whose image is $P$:

$$I = \{w \in A^* \mid Qw = P\}.$$

Its minimal generator set $W = I - IA^+$ is finite and we have

$$A^* = WA^* \cup W_0,$$

where $W_0$ is the set of all proper prefixes of the words in $W$.

Let $w \in W$ be a fixed element and denote by $\varphi : A^* \to \mathbb{Q}$ the additive morphism defined for all $a \in A$ by

$$a\varphi = \frac{1}{|P|} \sum_{p\in P} |p * a|.$$

We shall prove the identity

$$uv\varphi = \frac{1}{|P|} \sum_{p\in P} |p * uv|. \tag{21}$$

Indeed, since the word $u$ induces a permutation on $P$ we have

$$\frac{1}{|P|} \sum_{p\in P} |p * uv| = \frac{1}{|P|} \sum_{p\in P} (|p * u| + |(p.u) * v|)$$

$$= \frac{1}{|P|} \sum_{p\in P} |p * u| + \frac{1}{|P|} \sum_{p.u\in P} |(p.u) * v| = u\varphi + v\varphi = uv\varphi.$$

We now claim that the following holds:

$$p.u = p \in P \text{ implies } |p * u| = u\varphi. \tag{22}$$

Indeed, let $n$ be the order of the permutation induced by $u$ on $P$. Since $q.u^n = q$ holds for all $q \in P$ and since $q * u^n$ does not depend on $q \in P$, (21) yields

$$|p * u| = \frac{1}{n} |p * u^n| = \frac{1}{n|P|} \sum_{q\in P} |qxu^n| = \frac{1}{n} u^n\varphi = u\varphi.$$

Set $p = q_-.w$ and consider an arbitrary word $u \in A^*$. There exists $v \in A^*$ such that $p.uv = p$. Because of (22) we have

$$|q_- * wuv| = |q_- * w| + |p * uv| = |q_- * w| + u\varphi + v\varphi = |q_- * wu| + |(p.u) * v|;$$

thus

$$|wu\alpha| = \sum_{a\in A} a\varphi |u|_a + (|q_- * w| + v\varphi - |(p.u) * v|)$$

which completes the proof in the present case.

Now we claim that the previous verification covers the cases when $S$ satisfies either condition (1) or (2) of Proposition 5.4. Indeed, assume the alphabet may be

partitioned into $A = A_1 \cup A_2$ where $a \in A_1$ iff it induces the identity on $Q$, and where $A_2^* \sigma$ satisfies (20).

For all $u \in A^*$ we denote by $u_1$ and $u_2$ its projections over $A_1^*$ and $A_2^*$ respectively. Since $a\mu$ is a diagonal matrix with the same non-zero entry for all $a \in A_1$, we have $u\mu = u_1\mu u_2\mu$ and thus, $u\mu\rho = u_1\mu\rho u_2\mu\rho$.

Now the function $u_1 \to u_1\mu\rho$ is a morphism of $A_1^*$ into $t^*$, thus it is a counting function. By the previous study, $u_2 \to u_2\mu\rho$ is a counting function from $A_2^*$ into $t^*$. Applying twice Proposition 4.7 and then Proposition 4.10 completes the verification.

*Induction step*

We now assume that there exists a partition $A = A_1 \cup A_2$ such that $S_1 = A_1^* \sigma$ and $s_2 = (A_1^* A_2)^* \sigma$ are proper subsemigroups of $S$. Let $\gamma$ be a surjective mapping of $A_1^* A_2$ onto a (finite) set $B$ defined for all $u, u' \in A_1^* a$, $a' \in A_2$ by

$$ua\gamma = u'a'\gamma \quad \text{iff} \quad a = a' \text{ and } u\mu\pi = u'\mu\pi.$$

Define a morphism $\mu_2 : B^* \to \mathcal{M}$ by setting $b\mu_2 = u\mu\pi$ for all $u \in b\gamma^{-1}$. Let $u_1, \ldots, u_n \in A_1^* A_2$ and $b_i = u_i\gamma$ for all $i = 1, \ldots, n$. Then we have

$$b_1 \ldots b_n\mu_2 = b_1\mu_2 \ldots b_n\mu_2 = u_1\mu\pi \ldots u_n\mu\pi$$

and thus,

$$(b_1 \ldots b_n\mu_2)\pi = (u_1\mu\pi \ldots u_n\mu\pi)\pi = u_1 \ldots u_n\mu\pi.$$

Therefore, $\mu_2$ satisfies condition (iii) of Theorem 3.1 and so does obviously the restriction $\mu_1$ of $\mu$ to $A_1^*$. Thus, we may apply the induction hypothesis to $\mu_1$ and $\mu_2$. Consider the partition

$$A^* = A_1^* \cup \left[ \bigcup_{x \in A_1^* \mu\pi} A^* A_2 x (\mu\pi)^{-1} \right].$$

We shall verify that the restriction of $\alpha$ to every subset of the partition is a counting function and that all semaphores are counted with the same coefficient.

Clearly, by induction hypothesis, the restriction of $\alpha$ to $A_1^*$ is a counting function. Now let $x \in A_1^* \mu\pi$ be a fixed element. Every word $w \in A^* A_2 (\mu\pi)^{-1}$ has a unique factorization $w = v_1 a_1 \ldots v_r a_r v_{r+1}$ where $a_i \in A_2$, $i = 1, \ldots, r$, $v_1 \in A_1^*$, $i = 1, \ldots, r+1$. We have

$$\begin{aligned} w\mu &= v_1 a_1 \mu \ldots v_r a_r \mu v_{r+1}\mu \\ &= v_1 a_1 \mu\rho \ldots v_r a_r \mu\rho \; v_{r+1}\mu\rho \; v_1 a_1 \mu\pi \ldots v_r a_r \mu\pi \; v_{r+1}\mu\pi. \end{aligned} \tag{23}$$

We set $b_i = v_i a_i \gamma$, $i = 1, \ldots, r$. Because of

$$v_i a_i \mu = v_i a_i \mu\rho \; v_i a_i \mu\pi = v_i\mu\rho \; a_i\mu\rho \; v_i\mu\pi \; a_i\mu\pi,$$

there exists $z \in t^*$ depending only on $b_i$ such that $z \, v_i\mu\rho = v_i a_i \mu\rho$. Setting $b_i\tau = z$ defines a morphism $\tau : B^* \to t^*$. Then (23) yields

$$w\mu\rho = (v_1\mu_1\rho \ldots v_{r+1}\mu_1\rho)(b_1 \ldots b_r\tau)(b_1 \ldots b_r\mu_2)\rho\lambda, \tag{24}$$

where $\lambda$ has finite image.

Let us verify that each of the three terms in (24) is a counting function of $A^*$ into $t^*$. Since $\mu_1\rho$ is a counting function by induction hypothesis, by virtue of Proposition 4.9, $w \to v_1\mu_1\rho \dots v_r\mu_1\rho$ is also a counting function. Now $\tau$ is a morphism, thus a counting function of $B^*$ into $t^*$. By Proposition 4.8, $w \to b_1 \dots b_r\tau$ is a counting function of $A^*$ into $t^*$. Finally, $\mu_2\rho : B^* \to t^*$ is a counting function by induction hypothesis. Then Proposition 4.8 shows that $w \to (b_1 \dots b_r\mu_2)\rho$ is also a counting function from $A^*$ into $t^*$. □

## A. Appendix

In this section we shall prove the following result.

**Theorem A.1.** *Let $M$ be a submonoid of $A^* \times A^*$. The following conditions are equivalent:*

(i) *for all $(u, v) \in M$, $u$ and $v$ are conjugate;*

(ii) *there exists a $t \in A^*$ such that $ut = tv$ holds for all $(u, v) \in M$.*

We first recall a few basic results on free monoids.

### A.1. *Primitivity—conjugacy*

A word $w \in A^*$ is *primitive* if it is not a power of some shorter word: $w = u^n$ implies $n = 1$. The following result shows that each word $w \in A^+$ is the power of some unique primitive word called its *root* and denoted by $\sqrt{w}$. By convention we set $\sqrt{1} = 1$ although the empty word is not primitive (cf. [10, Lemmas 3 and 4]).

**Proposition A.2.** *Given $u, v \in A^*$, the following conditions are equivalent:*

(i) $uv = vu$;

(ii) *there exists $w \in A^+$ and $i, j > 0$ such that $u = w^i$ and $v = w^j$;*

(iii) *there exist $n, m > 0$ such that $u^n = v^m$.*

As a consequence, restricted to the free semigroup $A^+$, the relation of commutation is an equivalence relation.

By analogy with groups, two words $u$, $v$ are *conjugate* if there exists a *conjugacy factor* $w \in A^*$ such that $uw = wv$. Conjugate words are characterized by the following result (cf. [10, Theorem 3]).

**Proposition A.3.** *Given $u, v \in A^+$ and $w \in A^*$, the following conditions are equivalent:*

(i) $uv = wv$;

(ii) *there exist $x, y \in A^*$ and $i \geqslant 0$ such that $u = xy$, $w = (xy)^i x$ and $v = yx$;*

(iii) *there exist two unique integers $i, j \geqslant 0$ and two words $x, y \in A^*$, $y \neq 1$, such that $xy$ is primitive and $u = (xy)^i$, $w = (xy)^j x$ and $v = (yx)^i$.*

In particular, two words $u$, $v$ are conjugate iff there exist $x$, $y$ such that $u = xy$ and $v = yx$. As a result, the relation of conjugacy is an equivalence relation and we write $u \sim v$ iff $u$ and $v$ satisfy either of the last three conditions.

The next result implies that two words $u$ and $v$ are conjugate iff $\sqrt{u}$ and $\sqrt{v}$ are conjugate.

**Corollary A.4.** *Given $u, v \in A^+$ and $w \in A^*$, the following conditions are equivalent:*

(i) *there exist $n, m > 0$ such that $u^n w = wv^m$;*

(ii) *there exist $u_1, v_1 \in A^+$ and integers $p, q > 0$ satisfying $np = mq$, $u = u_1^p$, $v = u_1^q$ and $u_1 w = w v_1$.*

**Proof.** Clearly, (ii)$\Rightarrow$(i).

Conversely, assume $u^n w = wv^m$ holds. By Proposition A.3(iii), we have $u^n = (xy)^i$ and $v^m = (yx)^i$, where $xy$ is primitive and $y \neq 1$. Then $yx$ is also primitive since $yx = t^r$ implies $xy = t'^r$ for some word $t'$ which is a conjugate of $t$. By Proposition A.2, there exist $p$ and $q$ such that $u = (xy)^p$ and $v = (yx)^q$. It then suffices to set $u_1 = xy$ and $v_1 = yx$. □

The following is a sharp characterization of conjugate words (cf. [10, Theorem 4]).

**Proposition A.5.** *Two words $u, v \in A^*$ are conjugate iff, for some $n, m > 0$, the powers $u^n$ and $v^m$ have a common factor of length $|u| + |v| - \gcd(|u|, |v|)$.*

We end this section with a technical result which will be of very convenient use in the sequel.

**Lemma A.6.** *If $xy$ is a primitive word such that $y \neq 1$ and if, for some $z \in A^*$, $zyx$ is a prefix of a power $(xy)^n$, then $z = (xy)^i x$ for some $i \geqslant 0$.*

**Proof.** Let us set $xy = u$ and $zyxt = u^n$. There exist an integer $0 \leqslant i < n$ and a factorization $u = u_1 u_2$, $u_2 \neq 1$ such that $z = u^i u_1$. Arguing on the lengths we obtain $yx = u_2 u_1$, i.e., by Proposition A.2(ii), $x = u_1$ and $y = u_2$, completing thus the proof. □

*A.2. Proof of Theorem* A.1

Before proving Theorem A.1 we examine a special degenerate case.

**Lemma A.7.** *Let $u_1, u_2, v_1, v_2 \in A^+$ satisfy $u_1 \sim v_1$ and $u_2 \sim v_2$. The following four conditions are equivalent:*

(i) *$u_1 u_2 = u_2 u_1$ and $v_1 v_2 = v_2 v_1$;*

(ii) *$u_1 u_2 = u_2 u_1$ and $u_1 u_2 \sim v_1 v_2$;*

(iii) *$v_1 v_2 = v_2 v_1$ and $v_1 v_2 \sim u_1 u_2$;*

(iv) *for all $t \in A^*$, $u_1 t = t v_1$ iff $u_2 t = t v_2$;*

(v) *there exist two distinct elements $t_1, t_2 \in A^*$ such that*

$$u_1 t_1 = t_1 v_1, \qquad u_2 t_1 = t_1 v_2, \qquad u_1 t_2 = t_2 v_1, \qquad u_2 t_2 = t_2 v_2.$$

**Proof.** Assume first that condition (i) is satisfied and denote by $w$ and $z$ the roots which are common to $u_1$ and $u_2$ and to $v_1$ and $v_2$ respectively:

$$u_1 = w^i, \qquad u_2 = w^j, \qquad v_1 = z^i, \qquad v_2 = z^j.$$

(i)$\Rightarrow$(ii): Since $w$ and $z$ are conjugate, so are $u_1u_2 = w^{i+j}$ and $v_1v_2 = z^{i+j}$ by Corollary A.4.

(i)$\Rightarrow$(iv): In view of Corollary A.4. we have $u_1t = tv_1$ iff $wt = tz$ iff $u_2t = tv_2$

(ii)$\Rightarrow$(i): There exist two conjugate words $w$ and $z$ such that $u_1 = w^i$, $u_2 = w^j$ and $v_1v_2 = z^{i+j}$. The equalities $|u_1| = |v_1|$ and $|u_2| = |v_2|$ imply $v_1 = z^i$ and $v_2 = z^j$; i.e., $v_1v_2 = v_2v_1$.

(iv)$\Rightarrow$(v): Trivial.

(v)$\Rightarrow$(i): By Proposition A.3, without loss of generality, we may assume that $t_1$ is a suffix of $t_2$: $t_2 = zt_1$. Then we obtain

$$u_1zt_1 = u_1t_2 = t_2v_1 = zt_1v_1 = zu_1t_1$$

and

$$u_2zt_1 = u_2t_2 = t_2v_2 = zt_1v_2 = zu_2t_1.$$

This implies $u_1z = zu_1$ and $u_2z = zu_2$, i.e., $u_1u_2 = u_2u_1$ since $z \neq 1$. Similarly, $v_1v_2 = v_2v_1$ holds.

Finally, by symmetry, (i) and (iii) are equivalent. □

We now turn to the proof of Theorem A.1.

Clearly, only (i)$\Rightarrow$(ii) requires a verification. In the first place, we shall establish that any two pairs $(u_1, v_1)$, $(u_2, v_2) \in M$ have a common conjugacy factor $u_1t = tv_1$ and $u_2t = tv_2$.

Because of Lemma A.7, we may assume without loss of generality that $u_1u_2 \neq u_2u_1$ and $v_1v_2 \neq v_2v_1$. Two cases need be considered.

*Case* 1: $u_1$ and $u_2$ are not powers of two conjugate words.

Let $n, m > 0$ satisfy the inequality:

$$|u_1^n| = |u_2^m| \geqslant \max(|u_1|, |u_2|) + 2\min(|u_1| + |u_2|). \tag{A.1}$$

Since $u_1^nu_2^m$ and $v_1^nv_2^m$ are conjugate, we have

$$u_1^nu_2^m = tz \quad \text{and} \quad v_1^nv_2^m = zt.$$

We claim that the following holds:

$$u_1^nu_2^mt = tv_1^nv_2^m, \qquad u_1t = tv_1 \quad \text{and} \quad u_2t = tv_2. \tag{A.2}$$

Indeed, since $u_1$ and $v_1$ are conjugate, there exist two words $x \in A^*$ and $y \in A^+$, and an integer $i \geqslant 0$ such that $xy$ is a primitive word and $u_1 = (xy)^i$ $v_1 = (yx)^i$. Arguing on the symmetry of conjugacy we may assume $|t| \leqslant |u_1^n| = |u_2^m|$. In particular, $u_1^n$ and $v_2^m$ have a common factor of length $|t|$. By Proposition A.5, this implies $|t| \leqslant |t_2| = |u|_1 + |v|_2| = |u_1| + |u_2|$. Now,

$$u_1^nu_2^mt = tv_1^nv_2^m \tag{A.3}$$

holds, which shows that $tyx$ is a prefix of $u_1^n = (xy)^{ni}$. By virtue of Lemma A.6, this implies $t = (xy)^j x$ for some $j \geq 0$, i.e., $u_1 t = t v_1$. Simplifying (A.3) by $u_1^n$ to the left yields $u_2^m t = t v_2^m$, thus $u_2 t = t v_2$ by Corollary A.4.

*Case* 2: $u_1$ and $u_2$ are powers of two distinct conjugate words.

We set $u_3 = u_1^2 u_2^2$ and $v_3 = v_1^2 v_2^2$. Then $u_2$ and $u_3$ are not powers of two conjugate words since otherwise we would have $u_2 u_3 = u_3 u_2$ and thus $u_1 u_2 = u_2 u_1$. According to the previous case there exists a word $t$ such that $u_2 t = t v_2$ and $v_3 t = t v_3$. Then,

$$t v_1^2 v_2^2 = t v_3 = u_3 t = u_1^2 u_2^2 t = u_1^2 t v_2^2.$$

After simplification we get $t v_1^2 = u_1^2 t$ i.e. $t v_1 = u_1 t$.

It now suffices to prove that there exists a conjugacy factor which is common to all elements $(u, v) \in M$.

By Lemma A.7, we may assume that there exist $(u_1, v_1)$ and $(u_2, v_2)$ such that $u_1 u_2 \neq u_2 u_1$ and $v_1 v_2 \neq v_2 v_1$. Possibly after using the same trick as in Case 2, we may further assume that $u_1$ and $u_2$ are not provers of two conjugate words. For some unique $t$ the following holds: $u_1 t = t v_1$ and $u_2 t = t v_2$. Let $(u_3, v_3) \in M$. Without loss of generality we may assume $u_3 v_3 \neq v_3 u_3$. Choosing $n$ and $m$ as in (A.1) we get

$$u_1^n u_2^m t' = t' v_1^n v_2^m \quad \text{and} \quad u_3 t' = t' v_3.$$

Then (A.1) shows that $t$ and $t'$ are equal, thus completing the proof. □

## References


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

[2] J. Berstel and D. Perrin, *Theory of Codes* (Academic Press, New York, 1985).

[3] J. Berstel and C. Reutenauer, *Les Séries Rationnelles et leurs Langages*, (Masson, Paris, 1984).

[4] C. Choffrut, A generalization of Ginsburg and Rose's characterization of *g-sm* mappings, in: *Proc. 6th ICALP*, Lecture Notes in Computer Science (Springer, Berlin, 1979) 88–103.

[5] S. Eilenberg, *Automata, Languages and Machines, Vol. A* (Academic Press, New York, 1974).

[6] S. Ginsburg and G.F. Rose, A characterization of machine mappings, *Canad. J. Math.* **18** (1966) 381–388.

[7] J.H. Johnson, Formal models for string similarity, Ph.D. Thesis, University of Waterloo, 1983.

[8] D.E. Knuth, *The Art of Computer Programming, Vol. 3* (Addison-Wesley, Reading, MA, 1973).

[9] G. Lallement, *Semigroups and Combinatorial Applications* (Wiley Interscience, New York, 1979).

[10] A. Lentin and M.P. Schützenberger, A combinatorial problem in the theory of free monoids, in: B. Bowling, ed., *Proc. Conf. held at the University of North-Carolina at Chapel Hill*, Chapel Hill (1967) 128–144.

[11] R.C. Lyndon and M.P. Schützenberger, The equation $a^m = b^n c^p$ in a free group, *Michigan Math. J.* **9** (1962) 289–298.

[12] M. Nivat, Transductions des langages de Chomsky, *Ann. Inst. Fourier* **18** (1968) 339–456.

[13] M.P. Schützenberger, On a theorem of R. Jungen, *Proc. Amer. Math. Soc.* **13** (1962) 885–889.

[14] M.P. Schützenberger, Sur une variante des fonctions séquentielles, *Theoret. Comput. Sci.* **4** (1977) 243–259.