SIAM J. COMPUT.
Vol. 20, No. 4, pp. 669-685, August 1991

005

MINIMIZATION OF RATIONAL WORD FUNCTIONS*

CHRISTOPHE REUTENAUER† AND MARCEL-PAUL SCHUTZENBERGER‡

Abstract. Rational functions from a free monoid into another are characterized by the finiteness of the index of some congruence naturally associated with the function. A sequential bimachine is constructed computing the function, which is completely canonical, and in some sense minimal. This generalizes the Nerode criterion and the minimal automaton of a rational language, and similar results for sequential functions.

Key words. rational function, sequential bimachine

AMS(MOS) subject classification. 68D15

1. Introduction. Sequential machines appear as a ubiquitous tool in data processing and in basic software, since they constitute the most general algorithm between words that can be executed in real time by a finite device. Their theory is one of the earliest well-developed chapters of Automata Theory [8], and their natural generalization, i.e., the rational functions from a free monoid A^* (set of input words) to another B^* (output words) plays a basic role in the study of context-free languages and compilation [1]. The present paper is a contribution to the understanding of rational functions.

Here and in the sequel, we follow Eilenberg's terminology as used in his treatise [7]. In particular, by *function* we mean a partial mapping, and we recall that a rational function α from a semigroup S into a semigroup T is a function such that its graph $\{(s, \alpha(s)) \mid s \in \mathrm{dom}(\alpha)\}$ is a *rational* subset of the product semigroup $S \times T$. This definition is not the most convenient for our present purposes, and we shall use other equivalent definitions, by means of automata and machines. In order to understand the concepts which motivate the study of these objects, we begin with an informal presentation of the topic.

Recall that a sequential automaton is a two-tape machine reading the input tape from left to right, and writing on the output tape from left to right; no left move, nor ε-move, is allowed. A sequential function is by definition a function $\alpha: A^* \to B^*$ which is realized by some sequential automaton. Sequential functions are closed under functional composition.

Strictly speaking, what we have just described are left sequential objects and one could consider right sequential ones in a symmetric way (read and write from right to left). However, the associated functions are quite different. For instance, in a fixed integer base, multiplication by a given integer can be carried out by a sequential automaton if and only if it reads from right to left, while it is the reverse that is true for the division.

This leads to a more intuitive definition of rational function as the closure under composition of left and right sequential functions. An early theorem of Elgot and Mezei on general rational relations (see [1, Chap. 4, Thm. 5.2]) shows that any rational function can be obtained by composing one left and one right sequential function. This is expressed in more compact fashion by the concept of a bimachine [12] according

* Received by the editors November 20, 1989; accepted for publication (in revised form) October 23, 1990.
† Département de Mathématique et d'Informatique, Université du Québec à Montréal, Montréal, Québec, Canada H3C 3P8.
‡ Académie des Sciences, Paris, France, and 97 rue du Ranelagh, 75016, Paris, France.

to Eilenberg's terminology [7]. A further basic property that we shall make use of is that if α is an injective rational function of its domain, its inverse α^{-1} is again a rational function. For instance, morphisms $\varphi: A^* \to B^*$ may be the simplest rational functions. They are both left and right sequential functions. Another way of stating that a morphism φ is injective is the condition that the image $\varphi(A)$ of the input alphabet is a code, and in this case the decoding function φ^{-1} has been intensively studied (see [2]).

The main result of this paper is a characterization of rational functions, which extends to functions the classical definition of recognizable languages in terms of finiteness of the index of a certain congruence (Theorem 1). As a byproduct, this shortens considerably the proof of a Hankel-like characterization of rational functions [13]. The second main result (Theorem 2) shows that it is possible to associate to a rational function α a bimachine that is completely canonical, up to the choice of a certain left congruence on A^* which must be compatible with the left *adjacency* relation of α. Among these congruences, there is one, the *syntactic congruence*, which is canonical. When α is a total function, the bimachine that we construct is minimal in the following sense: it has the minimum number of left states among all bimachines computing α and having the set of right states corresponding to the given congruence. In general, it is not true that α has a unique minimal device realizing it (see, for instance, [3] for the case of decoding functions) but our result is the first step in this direction. The existence of a canonical machine is far from being trivial because, in view of the two-sided action, there is an unbounded number of ways by which one can realize the necessary trade-off between the spaces of left and right states.

Of course, the construction of a canonical bimachine gives a decision procedure for the equivalence of two rational functions (the fact that this is decidable was already known, see [1]). One can expect that, similar to the close relation between combinatorial aspects of rational languages and algebraic properties of their syntactic monoid, there should exist connections between properties of a rational function and its canonical bimachine (see the open problems at the end of this paper).

2. Preliminary results. Recall that a subset of a monoid M is called *rational* if it may be obtained from the finite subsets of M by a finite sequence of the following three operations: union $\mathbf{K} \cup \mathbf{L}$, product $\mathbf{KL}$, star $\mathbf{K}^* = \bigcup_{n \geq 0} \mathbf{K}^n$ = the submonoid generated by $\mathbf{K}$ (see [1], [6]).

We prefer the terminology "rational" to "regular," because the former emphasizes the analogy with the theory of rational functions of classical analysis and of rational power series in noncommuting variables.

We consider here partial functions from a finitely generated free monoid into another. If $\alpha: A^* \to B^*$ is such a function, then it is called *rational* if its graph $\#\alpha = \{(u, v) \in A^* \times B^* \mid u \in \operatorname{dom}(\alpha), v = \alpha(u)\}$ is a rational subset of the product monoid $A^* \times B^*$.

In the sequel, we identify each word w and the subset $\{w\}$. We write $\alpha(w) = \varnothing$, if w is not in the domain of α.

A more effective characterization is the following: the function α is rational if and only if there exists a matrix representation (monoid homomorphism) $\mu: A^* \to (2^{B^*})^{n \times n}$, where 2^{B^*} is the boolean semiring of subsets of B^* (with union and product), a row vector λ, and a column vector ρ of length n with entries in the same semiring, such that for any word w, one has $\alpha(w) = \lambda \mu(w) \rho$ ((see [1, Chap. 3, Prop. 7.3]); the fact that α is a function forces each entry of μ, λ, ρ to be empty or a singleton, once the unnecessary states have been removed). The latter characterization shows that a

rational function has the following property, which is called the *Hankel property*, because it concerns the Hankel matrix $(\alpha(uv))_{u,v\in A^*}$.

LEMMA 1 (Hankel property). *For any rational function α, there exists an integer n and $2n$ functions $\beta_1, \cdots, \beta_n, \gamma_1, \cdots, \gamma_n : A^* \to B^*$ such that for any words x, y in A^**

$$\alpha(xy) = \bigcup_{1\leqq i\leqq n} \beta_i(x)\gamma_i(y).$$

Here and in the sequel, we consider each word, and $\varnothing$, to be embedded in the boolean semiring 2^{B^*}, with union and product; thus the previous equation means that for each i with $x \in \mathrm{dom}\,(\beta_i)$, $y \in \mathrm{dom}\,(\gamma_i)$, one has $\alpha(xy) = \beta_i(x)\gamma_i(y)$, and that $\alpha(xy) = \varnothing$ if for no i one has $x \in \mathrm{dom}\,(\beta_i)$ and $y \in \mathrm{dom}\,(\gamma_i)$.

Proof. Let μ, λ, ρ be as in the characterization before the lemma. Then

$$\alpha(xy) = \lambda\mu(xy)\rho = \lambda\mu x\mu y\rho = \bigcup_{1\leqq i\leqq n} (\lambda\mu x)_i(\mu y\rho)_i = \cup \beta_i(x)\gamma_i(y).$$

To conclude, note that if $|\beta_i(x)| \geqq 2$ for some x (in case β_i is not a function), one must have $\gamma_i = \varnothing$, because α is a function; so this index i can be omitted (the case is similar if some γ_i is not a function). □

A result of Schützenberger shows that the converse also holds [13]. We shall give a new proof of it in the next section. For the moment, let us point out what this Hankel property means in the case of characteristic functions, i.e., functions whose image is contained in $\{\varnothing, 1\}$ (we denote by 1 the empty word).

LEMMA 2. *Let $\alpha : A^* \to B^*$ be the characteristic function of its domain L. The following conditions are equivalent*:

(i) *α has the Hankel property.*

(ii) *$c(L)$ is a finite union $\cup H_i \times K_i$, where $c(w) = \cup_{w=xy} (x, y) \subset A^* \times A^*$.*

(iii) *L is a rational language.*

Note that (ii) is a Hopf-algebra-like characterization of rational languages.

Proof. (i)⇒(ii): Let $H_i = \mathrm{dom}\,(\beta_i)$ and $K_i = \mathrm{dom}\,(\gamma_i)$, where β_i and γ_i satisfy $\alpha(xy) = \cup_{1\leqq i\leqq n} \beta_i(x)\gamma_i(y)$. Then clearly $c(L) = \cup H_i \times K_i$.

(ii)⇒(iii): this is evident by "Nerode's criterion": if the set $\{x^{-1}L \mid x \in A^*\}$ is finite, then L is rational, where $x^{-1}L = \{y \mid xy \in L\}$. Now, $x^{-1}L$ is the union of the K_i's for which $x \in H_i$. Hence the $x^{-1}L$ are finite in number.

(iii)⇒(i) is a particular case of Lemma 1. □

The next lemma shows the functorial properties of the Hankel property.

LEMMA 3. (i) *If α and α' satisfy the Hankel property, then so does $\alpha' \circ \alpha$.*

(ii) *If α satisfies the Hankel property, then* $\mathrm{dom}\,(\alpha)$ *is rational.*

(iii) *If α satisfies the Hankel property, then α^{-1} preserves rationality.*

Proof. (i) We have

$$\alpha' \circ \alpha(xy) = \alpha'\left(\bigcup_i \beta_i(x)\gamma_i(y)\right) = \bigcup_i \alpha'(\beta_i(x)\gamma_i(y))$$

$$= \bigcup_i \bigcup_{i'} \beta_{i'}(\beta_i(x))\gamma_{i'}(\gamma_i(y)) = \bigcup_{i,i'} (\beta_{i'} \circ \beta_i)(x)(\gamma_{i'} \circ \gamma_i)(y).$$

(ii) In this case, the characteristic function of $\mathrm{dom}\,(\alpha)$ satisfies the Hankel property, so it is rational by Lemma 2.

(iii) Let L be a rational language in B^*, and let $\alpha : A^* \to B^*$ satisfy the Hankel property. Let α' be the characteristic function of L. Then by Lemma 2 and (i), $\alpha' \circ \alpha$ satisfies the Hankel property, hence by (ii), $\mathrm{dom}\,(\alpha' \circ \alpha)$ is rational. But $\mathrm{dom}\,(\alpha' \circ \alpha) = \alpha^{-1}(L)$. □

This lemma will enable us to prove the following implication: if α has the Hankel property, then α is a rational function. Proving it is much more difficult than in the case of characteristic functions (Lemma 2). It depends on a Nerode-like characterization of rational functions (the main result of § 3), and on Choffrut's theorem, which characterizes subsequential functions, and which is itself a generalization of the Ginsburg–Rose theorem on sequential functions. In order to state this theorem, define the *left distance* between two words by

$$\|u, v\| = |u| + |v| - 2|u \wedge v|,$$

where $|u|$ is the length of u and $u \wedge v$ the longest common left factor of u and v. In other words, $\|u, v\| = |s| + |t|$ where $u = ps$, $v = pt$, and $p = u \wedge v$. This can also be expressed by the equality $\|u, v\|$ = length of the reduced word (in the free group) $u^{-1}v$, or equivalently $v^{-1}u$. From this last fact, it is immediate that $\|u, v\|$ satisfies the triangular inequality. Hence, it is a distance (see also [1, Chap. 4, § 2, p. 104]).

A function $\alpha : A^* \to B^*$ will be said to be *uniformly bounded* if for any integer k, there exists an integer K such that for all $x, y \in \text{dom}(\alpha)$, $\|x, y\| \leqq k \Rightarrow \|\alpha(x), \alpha(y)\| \leqq K$. The terminology stems from the fact that such a function maps each bounded subset of dom (α) into a bounded subset of B^*, in a uniform way. Thus we do not use the terminology "bounded variation" of [4].

We shall give a formal definition of *subsequential functions* in § 4, but it seems advisable to recall now the following result.

THEOREM (Choffrut [4] or [1, Chap. 4, Thm. 2.7]). *A function α is subsequential if and only if it is uniformly bounded and α^{-1} preserves rationality.*

We say that two functions $\alpha, \beta : A^* \to B^*$ are *adjacent* if

$$\sup \{\|\alpha(f), \beta(f)\|, f \in \text{dom}(\alpha) \cap \text{dom}(\beta)\} < \infty.$$

The next result is a decidability result, which will imply that every construction in this paper is effective.

PROPOSITION 1. *If $\alpha, \alpha' : A^* \to B^*$ are rational functions, then one can decide if they are adjacent. In this case, the function $\alpha \wedge \alpha'$ defined by: $(\alpha \wedge \alpha')(f)$ equals the longest common left factor of $\alpha(f)$ and $\alpha'(f)$ when $f \in \text{dom}(\alpha) \cap \text{dom}(\alpha')$, and otherwise, $(\alpha \wedge \alpha')(f) = \alpha(f) \cup \alpha'(f)$, is rational and can be computed effectively.*

Remark 1. If α_1, α_2 are rational but not adjacent, then $\alpha_1 \wedge \alpha_2$ is not rational, in general. Define them, indeed, to be the homomorphisms $\{a_1, a_2\}^* \to t^*$ such that $\alpha_i(a_i) = t$, $\alpha_i(a_j) = 1$ for $j \neq i$.

Then $(\alpha_1 \wedge \alpha_2)(f)$ is equal to $t^{n(f)}$, where $n(f) = \inf(|f|_{a_1}, |f|_{a_2})$, which implies that $\alpha_1 \wedge \alpha_2$ is not rational (indeed, the inverse image of $(t^2)^*$, by the pumping lemma for finite automata, is not rational).

We shall need the following lemma, which is an easy consequence of a theorem of Fine and Wilf (see [9, Chap. 1, Prop. 3.5]).

LEMMA 4. *Let u, v, w, u', v', w' be words such that* $\sup \{\|uv^n w, u'v'^n w'\|, n \in \mathbb{N}\} < \infty$. *Then one has:*

(1) *For some word t, either $u' = ut$ and $tv' = vt$, or $u = u't$ and $tv = v't$.*

One of the referees pointed out that the lemma easily follows from the preliminary remark that $|v| = |v'|$.

Proof of Proposition 1. (1) Without loss of generality, we may assume that α and α' have the same domain and that $\alpha(1) = \alpha'(1) = \varnothing$. Indeed, we may restrict α and α' to dom $(\alpha) \cap$ dom $(\alpha') \setminus \{1\}$ and test the adjacency of these new functions. In this case, there exist transducers T and T' for α and α', with set of states Q, Q', initial states q_0, q_0', and unique final states q_f, q_f' (see [1, Chap. 3, Thm. 7.1]).

Define the "Kronecker product" of T and T': it is the "transducer" $\bar{T}$, with set of states $\bar{Q} = Q \times Q'$, inputs in A^*, and outputs in $B^* \times B^*$; there is a path $(p, p') \xrightarrow{x/(u,u')} (q, q')$ in $\bar{T}$ if and only if there is a path $p \xrightarrow{x/u} q$ in T and $p' \xrightarrow{x/u'} q'$ in T'; moreover, all the unnecessary states of $\bar{T}$ are removed, so that all states of $\bar{T}$ are accessible and coaccessible, with initial state $\bar{q}_0 = (q_0, q'_0)$ and final state $\bar{q}_f = (q_f, q'_f)$.

A *simple* path is a path without repetition of states, and a *simple* circuit is a closed path with no repetition of internal states.

We show that α and α' are adjacent if and only if $\bar{T}$ satisfies the following condition:

(C) For any simple path $(q_0, q'_0) \xrightarrow{x/(u,u')} (q, q')$ and any simple circuit $(q, q') \xrightarrow{y/(v,v')} (q, q')$, we have equation (1) of Lemma 4.

Clearly, if α, α' are adjacent, and with the notations of (C), there exists a path

$$(q, q') \xrightarrow{z/(w,w')} (q_f, q'_f).$$

Then $\alpha(xy^n z) = uv^n w$ and $\alpha'(xy^n z) = u'v'^n w'$. As α, α' are adjacent, Lemma 4 shows that (1) holds.

Conversely, suppose that (C) holds. Then, for each long enough word m in $\mathrm{dom}(\alpha) = \mathrm{dom}(\alpha')$, there is a factorization $m = xyz$, a simple path and a simple circuit as in (C) above, and a path $(q, q') \xrightarrow{z/(w,w')} (q_f, q'_f)$.

Then $\alpha(m) = uvw$, $\alpha'(m) = u'v'w'$. By (1), we have, e.g., $u' = ut$ and $tv' = vt$. Then $u'v'w' = utv'w' = uvtw'$, hence $\|\alpha(m), \alpha'(m)\| = \|uvw, uvtw'\| = \|w, tw'\| = \|uw, utw'\| = \|uw, u'w'\| = \|\alpha(xz), \alpha'(xz)\|$, which allows us to conclude by induction on the length of m.

Clearly, condition (C) is decidable, which completes the first part of the proof.

(2) We construct now a transducer for $\alpha \wedge \alpha'$, which will imply that it is a rational function. This construction is a rather classical covering construction, so we shall not be very formal.

We call a path in $\bar{T}$ *elementary* if it starts from (q_0, q'_0) and if only the last vertex is allowed to appear more than once, and in this case, only twice. Hence, such a path is either a simple path, or the concatenation of a simple path with a simple circuit, as in condition (C).

Denote by $u \wedge v$ the longest common left factor of the words u and v. We construct a tree T^* having the set of elementary paths in $\bar{T}$ as a set of nodes; there is an edge from π to π' in T^* if $\pi' = \pi e$, with e an edge in $\bar{T}$. Note that π, π' correspond to paths $(q_0, q'_0) \xrightarrow{x/(u,u')} (p, p')$ and $(q_0, q'_0) \xrightarrow{xa/(v,v')} (q, q')$, with u (respectively, u') a left factor of v (respectively, v'); so we have an equation $v \wedge v' = (u \wedge u')s$, for some word s in B^*: then the previously created edge in T^* will be labelled by a/s.

Call an elementary path *complete* if its last state is repeated. Now, in T^*, merge the node corresponding to such a state with its first occurrence in the path: in this way, we obtain a transducer S; let β be the function computed by S.

We show that $\beta = \alpha \wedge \alpha'$. Clearly, $\beta(m) = (\alpha \wedge \alpha')(m)$ for any word m such that there is in $\bar{T}$ an elementary path $(q_0, q'_0) \xrightarrow{m/\cdots} (q_f, q'_f)$.

It follows that this equality is true for each short enough word m. Now, let m be such that there is a nonelementary path $(q_0, q'_0) \xrightarrow{m/\cdots} (q_f, q'_f)$. Then this path may be decomposed as

$$(q_0, q'_0) \xrightarrow{x/(u,u')} (q, q') \xrightarrow{y/(v,v')} (q, q') \xrightarrow{z/(w,w')} (q_f, q'_f),$$

where the first two factors form an elementary path, for some factorizations $m = xyz$, $\alpha(m) = uvw$, $\alpha'(m) = u'v'w'$. Moreover, $\alpha(xz) = uw$ and $\alpha'(xz) = u'w'$.

This corresponds in S to a path

$$(q_0, q_0') \xrightarrow{x/\bar{u}} (q, q') \xrightarrow{y/\bar{v}} (q, q') \xrightarrow{z/\bar{w}} (q_f, q_f').$$

By construction, we have $\bar{u} = u \wedge u'$, $\bar{u}\bar{v} = uv \wedge u'v'$. By induction on $|m|$, we also have $(\alpha \wedge \alpha')(xz) = \beta(xz) = \bar{u}\bar{w}$. Now, condition (C) holds, so we have, e.g., $u' = ut$ and $tv' = vt$. Hence, $\bar{u}\bar{w} = uw \wedge u'w' = uw \wedge utw' = u(w \wedge tw')$. As $\bar{u} = u \wedge u' = u$, we obtain $\bar{w} = w \wedge tw'$. Moreover, $\alpha(m) \wedge \alpha'(m) = uvw \wedge u'v'w' = uvw \wedge uvtw' = uv(w \wedge tw') = uv\bar{w}$. Now, we have also $\bar{u}\bar{v} = uv \wedge u'v' = uv \wedge uvt = uv$, so that $\alpha(m) \wedge \alpha'(m) = \bar{u}\bar{v}\bar{w} = \beta(xyz) = \beta(m)$, which had to be shown. □

3. A characterization of rational functions. We give a characterization of rational functions, which has some formal analogy with the Nerode criterion for rational languages and which is related to Choffrut's theorem (see § 2).

As we consider partial functions, it will be convenient to use symbol $\varnothing$, and the distance will be extended by setting

$$\|\varnothing, \varnothing\| = 0, \qquad \|\varnothing, u\| = \|u, \varnothing\| = \infty$$

for any word u. By convention, we have $n < \infty$ for any number n and $n + \infty = \infty$. Then, the triangular inequality remains valid. Now, let α be a fixed (partial) function $A^* \to B^*$, where A, B are finite alphabets. Define a relation

$$u \sim v$$

on A^* by the condition

$$\sup \{\|\alpha(fu), \alpha(fv)\|, f \in A^*\} < \infty.$$

Note that, by the above conventions, $u \sim v$ implies that $\alpha(fu) = \varnothing$ if and only if $\alpha(fv) = \varnothing$. This implies, by the triangular inequality, that $\sim$ is transitive. Moreover, it is clearly reflexive and symmetric and it is not difficult to show that $\sim$ is left compatible, i.e., $u \sim v \Rightarrow xu \sim xv$ for any word x. Hence $\sim$ is a left congruence of A^*.

We call it the *syntatic left congruence* of α. The terminology is justified by the following observation: if α is the characteristic partial function of a language L (i.e., $\alpha(w) = 1$ if $w = L$, $= \varnothing$ if $w \notin L$), then its syntactic left congruence is the usual syntactic left congruence of L. One could, of course, also define the right syntactic congruence in a symmetric way.

The main result of this section is given in the following theorem.

THEOREM 1. *A partial function $\alpha : A^* \to B^*$ is rational if and only if its syntactic left congruence is of finite index and if $\alpha^{-1}(L)$ is rational for any rational language $L \subset B^*$.*

A consequence of this result is a new proof of the Hankel-like characterization of [13].

COROLLARY. *A partial function $\alpha : A^* \to B^*$ is rational if and only if there exists an integer n and partial functions $\beta_i, \gamma_i : A^* \to B^*$, $1 \leqq i \leqq n$, such that for any words x, y*

$$\alpha(xy) = \bigcup_{1 \leqq i \leqq n} \beta_i(x)\gamma_i(y). \tag{2}$$

Proof. We prove the theorem and its corollary at the same time by showing that α rational $\Rightarrow$ α satisfies the Hankel property $\Rightarrow$ $\sim$ of finite index and α^{-1} preserves rationality $\Rightarrow$ α rational. The first implication is Lemma 1 and one-half of the second is Lemma 3. So, assuming (2), we show that the syntactic congruence $\sim$ of α is of finite index.

We show that the condition

$$\forall i, \quad 1 \leqq i \leqq n: \gamma_i(u) \neq \varnothing \quad \text{iff } \gamma_i(v) \neq \varnothing \tag{3}$$

implies $u \sim v$: this will imply that the index of $\sim$ is less than or equal to 2^n. So, let (3) be satisfied and define N to be some integer greater than the lengths of the words $\gamma_i(u)$, $\gamma_i(v) \neq \emptyset$, $1 \leqq i \leqq n$. Let f be any word; we show that $\|\alpha(fu), \alpha(fv)\| < 2N$. Indeed, if $\alpha(fu) = \emptyset$, then by (2), for any i, either $\beta_i(f) = \emptyset$ or $\gamma_i(u) = \emptyset$. By (3) we obtain: for all i, $\beta_i(f)$ or $\gamma_i(v) = \emptyset$, and again by (2), $\alpha(fv) = \emptyset$. In this case, $\|\alpha(fu), \alpha(fv)\| = 0 < 2N$. On the other hand, if $\alpha(fu) \neq \emptyset$, then there exists by (2) an i such that $\alpha(fu) = \beta_i(f)\gamma_i(u)$ and $\beta_i(f) \neq \emptyset \neq \gamma_i(u)$. Hence, by (3), we have $\gamma_i(v) \neq \emptyset$, which implies by (2) that $\alpha(fv) = \beta_i(f)\gamma_i(v)$. Hence

$$\|\alpha(fu), \alpha(fv)\| = \|\beta_i(f)\gamma_i(u), \beta_i(f)\gamma_i(v)\| = \|\gamma_i(u), \gamma_i(v)\| < 2N.$$

Finally, we have $\sup\{\|\alpha(fu), \alpha(fv)\|, f \in A^*\} < \infty$ and thus $u \sim v$.

We now show the last implication: if $\sim$ is of finite index and if α^{-1} preserves rationality, then α is a rational function.

Since $\sim$ is a left congruence of finite index on A^*, the set

$$Q = A^*/\sim$$

is a finite set with a left action $(w, q) \mapsto wq$ of A^* on Q. Consider the finite alphabet $A \times Q$ and define a length-preserving function

$$\gamma : A^* \to (A \times Q)^*$$

by

$$\gamma(a_n \cdots a_1) = (a_n, q_{n-1}) \cdots (a_2, q_1)(a_1, q_0),$$

where $a_i \in A$, q_0 is the class of 1 mod $\sim$ and where $q_i = a_i q_{i-1}$ for $i = 1, \cdots, n-1$. This function γ is clearly sequential from right to left, and hence a rational function (see [1, Chap. 4, Cor. 2.3]). Clearly, γ is injective, hence γ^{-1} is a partial function. Actually, $\gamma^{-1} = \pi | \mathrm{Im}\,(\gamma)$, where π is the canonical projection

$$\pi : (A \times Q)^* \to A^*.$$

Define $\beta = \alpha \circ \gamma^{-1} : (A \times Q)^* \to B^*$. We have $\alpha = \beta \circ \gamma$ since γ is a total function. We show that β is a subsequential function, hence it is rational (see [1, Chap. 4, Prop. 2.4]); this will imply that α is rational, as a product of rational functions. (See [1, Chap. 3, Thm. 4.4 and Def., § 1].)

We use Choffrut's theorem, stated in § 2. As β^{-1} clearly preserves rationality (because $\beta^{-1} = \gamma \circ \alpha^{-1}$ and γ and α^{-1} both preserve rationality), it is enough to show that β is uniformly bounded.

CLAIM. *If $FU \in \mathrm{Im}(\gamma)$ with $F \neq 1$, then the last letter of F is of the form (a, uq_0) where $u = \pi(U)$.*

This is immediate from the definition of γ.

Let k be an integer. Define K to be some integer greater than $\|\alpha(fu), \alpha(fv)\|$ for any word f and any words u, v such that $u \sim v$ and $|u| + |v| \leqq k$, and greater than $\|\beta(X), \beta(Y)\|$ for $|X| + |Y| \leqq k$ and $X, Y \in \mathrm{dom}\,(\beta)$.

This is possible by the definition of $\sim$ and the fact that the words u, v with $|u| + |v| \leqq k$ (respectively, the words X, Y with $|X| + |Y| \leqq k$) are finite in number.

We show that

$$\forall X, Y \in \mathrm{dom}\,(\beta), \qquad \|X, Y\| \leqq k \Rightarrow \|\beta(X), \beta(Y)\| \leqq K, \tag{4}$$

which will imply that β is uniformly bounded. By the definition of K, it is enough to prove (4) for $|X| + |Y| > k$.

So, let X, Y with X, $Y \in \mathrm{dom}(\beta)$, $|X|+|Y|>k$, $\|X, Y\| \leqq k$. We may write $X = FU$, $Y = FV$, where F is the longest common left factor of X and Y. Since $\|X, Y\| \leqq k$, we have $|U|+|V| \leqq k$. Since $|X|+|Y|>k$, we also have $F \neq 1$.

Let $u = \pi(U)$, $v = \pi(V)$, $f = \pi(F)$. Since X, $Y \in \mathrm{dom}(\beta)$, we have X, $Y \in \mathrm{Im}(\gamma)$; hence, by the claim, the last letter of F is $(a, uq_0) = (a, vq_0)$, and thus $uq_0 = vq_0$, which implies $u \sim v$. By the definition of β, we have $\beta(X) = \alpha(fu)$ and $\beta(Y) = \alpha(fv)$. Since $|u|+|v| = |U|+|V| \leqq k$, we have by the definition of K, $\|\alpha(fu), \alpha(fv)\| \leqq K$, i.e., $\|\beta(X), \beta(Y)\| \leqq K$, which proves (4). □

4. A canonical bimachine. We modify slightly the definition of a generalized bimachine, as given in [1] and [7]. One of the reasons for this is that we want to give an arbitrary image to the empty word under the function computed by the bimachine.

A bimachine is given by

- A finite set L of *left states*, with right action $L \times A^* \to L$, $(l, w) \mapsto lw$, and a *left initial state* l_0.
- A finite set R of *right states*, with a left action $A^* \times R \to R$, $(w, r) \mapsto wr$, and with a *right initial state* r_0.
- An *output function* $\omega : \mathbf{L} \times A \times \mathbf{R} \to B^*$.
- A *final left function* $\lambda : \mathbf{R} \to B^*$ and a *final right function* $\rho : L \to B^*$.

The output function is extended to $L \times A^* \times R$ by the formula

$$\omega(l, uv, r) = \omega(l, u, vr)\omega(lu, v, r). \tag{5}$$

In particular, $w(l, 1, r) = 1$. The function computed by the bimachine is $\alpha : A^* \to B^*$ defined by

$$\alpha(w) = \lambda(wr_0)\omega(l_0, w, r_0)\rho(l_0 w). \tag{6}$$

If $w = a_1 \cdots a_n (a_i \in A)$, this may be written more algorithmically (using (5)) as

$$\alpha(a_1 \cdots a_n) = \lambda(a_1 \cdots a_n r_0) \cdot \prod_{i=1}^{n} \alpha(l_0 a_1 \cdots a_{i-1}, a_i, a_{i+1} \cdots a_n r_0) \times \rho(l_0 a_1 \cdots a_n). \tag{7}$$

When R is reduced to a single element, then a bimachine is simply a subsequential transducer, as in [1] (a subsequential transducer is sometimes called a generalized sequential machine with endmarker, see [5, Thm. 2.2]). A bimachine in the sense of [1], [7] is a bimachine as above, where λ and ρ are constant functions equal to 1.

Let $\alpha : A^* \to B^*$ be a function. We define on A^* a relation, which will be reflexive, symmetric, compatible with left multiplication, *but not transitive* in general. We call it the (*left*) *syntactic adjacency relation* of α, denoted by

$$u \leftrightarrow v.$$

It is defined by

$$\sup\{\|\alpha(fu), \alpha(fv)\|, f \in A^*, \alpha(fu) \neq \varnothing \neq \alpha(fv)\} < \infty. \tag{8}$$

Note that, in view of the definition of adjacent functions (§ 2), one has $u \leftrightarrow v$ if and only if the two functions $f \mapsto \alpha(fu)$ and $f \mapsto \alpha(fv)$ are adjacent. It is also easy to see that α is uniformly bounded if and only if $u \leftrightarrow v$ for any words u and v. Note, moreover, that if $\mathrm{dom}(\alpha) = A^*$, then $\leftrightarrow$ is transitive and equal to the left syntactic congruence of α.

We call a left congruence $\sim$ on A^* *compatible* with $\leftrightarrow$ if for any words u, v,

$$u \sim v \Rightarrow u \leftrightarrow v.$$

In terms of their graphs, this means that $\#(\sim)$ is contained in $\#(\leftrightarrow)$. Recall that when $\sim$ is a left congruence, then $R = A^*/\sim$ is naturally equipped with a left action $A^* \times R \to R$.

THEOREM 2. *Let $\alpha : A^* \to B^*$ be a rational function. Let $\sim$ be a left congruence of finite index on A^* and $R = A^*/\sim$ and r_0 the class of* 1 mod $\sim$. *The following conditions are equivalent*:

(i) *$\sim$ is compatible with the syntactic adjacency relation of α.*

(ii) *R, together with the natural left action and r_0 as initial right state, is the set of right states of some bimachine computing α.*

It will turn out that the bimachine that we obtain in the proof is completely canonical, once $\sim$ is given. Moreover, one may choose for $\sim$ the congruence considered in the previous section, thus obtaining a completely canonical bimachine. On the other hand, we shall verify that this bimachine is minimal, in the sense stated in the introduction, when α is a total function.

Proof of Theorem 2 (first part). (ii)$\Rightarrow$(i): Let R be the set of right states of a bimachine computing α. We have, by the definition of R,

$$u \sim v \Leftrightarrow ur_0 = vr_0.$$

We have to show that $u \sim v$ implies (8). Suppose that $u \sim v$, that is, $ur_0 = vr_0 = r$, for some r in R. Let N be some integer greater than the lengths of the words (if defined) $\omega(l, u, r_0)\rho(l')$ and $\omega(l, v, r_0)\rho(l')$, for l, l' in L. We have, by (5) and (6),

$$\begin{aligned}\alpha(fu) &= \lambda(fur_0)\omega(l_0, fu, r_0)\rho(l_0fu)\\ &= \lambda(fur_0)\omega(l_0, f, ur_0)\omega(l_0f, u, r_0)\rho(l_0fu)\\ &= \lambda(fr)\omega(l_0, f, r)\omega(l_0f, u, r_0)\rho(l_0fu).\end{aligned}$$

Similarly,

$$\alpha(fv) = \lambda(fr)\omega(l_0, f, r)\omega(l_0f, v, r_0)\rho(l_0fv).$$

If $\alpha(fu) \neq \varnothing \neq \alpha(fv)$, then $\alpha(fu)$ and $\alpha(fv)$ have $\lambda(fr)\omega(l_0, f, r)$ as a common left factor, hence

$$\|\alpha(fu), \alpha(fv)\| < 2N.$$

This shows (8), and thus $\sim$ is compatible with the left adjacency of α. □

Before continuing the proof, we need several lemmas.

LEMMA 5. *If α is a rational function and $\sim$ is a left congruence on A^* of finite index, then there exist nonempty rational functions β_i, γ_i, $1 \leqq i \leqq n$, such that*

(i) $\forall u, v \in A^*$, $\alpha(uv) = \bigcup_{1 \leqq i \leqq n} \beta_i(u)\gamma_i(v)$.

(ii) *Each set* dom (γ_i) *is contained in a single class* mod $\sim$.

Proof. (i) follows from Lemma 1 and its proof, which show that β_i, γ_i may be chosen rational. Now, note that each class mod $\sim$ is a rational language, and that the restriction of a rational function to a rational language is still rational. So, replacing in (i) each γ_i by the union of its restrictions to each class mod $\sim$, we obtain (ii). □

Remark 2. Using this lemma, it is easy to prove that *the graph of the syntactic adjacency relation of a rational function is a recognizable subset of $A^* \times A^*$* (in the sense of [1, Chap. 3, Thm. 1.5] and [7], i.e., a finite union of sets $K \times L$, where K, L are rational languages).

Indeed, define $i \leftrightarrow j$ if the functions β_i, β_j are adjacent. Now, for $I, J \subset \{1, \cdots, n\}$, define $I \leftrightarrow J$ if for any i in I, j in J, one has $i \leftrightarrow j$. Finally, let $I(u) = \{i \mid u \in \text{dom}\,(\gamma_i)\}$.

Then one shows that $u \leftrightarrow v$ if and only if $I(u) \leftrightarrow I(v)$. This implies that the graph of $\leftrightarrow$ is equal to

$$\bigcup_{I \leftrightarrow J} \{u \in A^* \mid I(u) = I\} \times \{v \in A^* \mid I(v) = J\},$$

which is recognizable.

We need to define the operator "longest common left factor" for sets of words rather than only pairs of words. For technical reasons, it should also be defined on the empty set. Each singleton set will be identified with its element. So, for a nonempty language L, let $\bigwedge L$ denote the longest common left factor of the words in L equal to $\emptyset$ if $L = \emptyset$.

For $x_1, \cdots, x_n \in A^* \cup \{\emptyset\}$, we define $x_1 \wedge \cdots \wedge x_n$ to be $\bigwedge L$, where L is the underlying set of the sequence. So $x_1 \wedge \cdots \wedge x_n \neq \emptyset$ if and only if at least one x_i is not equal to $\emptyset$. Note that if L is a language, then $\bigwedge L = \bigwedge L'$ for some sublanguage L' of cardinality less than or equal to 2 (indeed, if $|L| \geqq 2$, there exist words u, v in L such that $u \wedge v = \bigwedge L$). If $\alpha_1, \cdots, \alpha_n$ are functions $A^* \to B^*$, then the function $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_n$ will be defined by $\alpha(f) = \alpha_1(f) \wedge \cdots \wedge \alpha_n(f)$. Note that $\mathrm{dom}\,(\alpha) = \bigcup_{1 \leqq i \leqq n} \mathrm{dom}\,(\alpha_i)$, in view of the definitions.

We shall use the easily verified identities

$$\bigwedge \left(\bigcup_{i \in I} L_i \right) = \bigwedge_{i \in I} \left(\bigwedge L_i \right)$$

for any languages L_i, $i \in I$, and

$$\bigwedge (gL) = g\left(\bigwedge L\right)$$

for any language L and g in $A^* \cup \{\emptyset\}$.

LEMMA 6. *Let $\alpha_1, \cdots, \alpha_n : A^* \to B^*$ be pairwise adjacent functions such that each α_i^{-1} preserves rationality.*

(i) *For any words g_1, g_2 in B^*, the language*

$$\{f \mid \exists w \in B^*,\ \alpha_1(f) = wg_1,\ \alpha_2(f) = wg_2\}$$

is rational.

(ii) *If the functions α_i are, moreover, rational, then $\alpha_1 \wedge \cdots \wedge \alpha_n$ is rational.*

Note that this gives an alternative proof of the following: α, α' rational and adjacent implies $\alpha \wedge \alpha'$ rational (see Proposition 1).

Remark 3. Let α_1, α_2 be as in Remark 1. Then the language $\{f \in A^* \mid \alpha_1(f) = \alpha_2(f)\}$ (this is the case $g_1 = 1 = g_2$ of the lemma) is equal to $\{f \in A^*, |f|_{a_1} = |f|_{a_2}\}$, and hence is not rational. This shows that the adjacency hypothesis is not superfluous in Lemma 6.

Proof. (i) Let p be an integer such that $|g_1|, |g_2| < p$ and that for any f in A^*, $\alpha_1(f)$ and $\alpha_2(f)$, if defined, differ only by a right factor of length less than p.

We show that for f in A^*, the condition

(a) $\exists w \in B^*, \quad |w| \geqq p, \quad \alpha_1(f) = wg_1 \quad \text{and} \quad \alpha_2(f) = wg_2$

is equivalent to the condition

(b) $\exists i \in \{0, \cdots, 2p-1\}, \quad \exists u \in B^p \ \text{ such that } \ \alpha_1(f) \in B^i(B^{2p})^* u g_1$
and $\alpha_2(f) \in B^i(B^{2p})^* u g_2$.

Suppose that this is proved. Then the language L of the lemma is equal to $L_1 \cup L_2$, where $L_1 = \{f \in A^* \mid f \text{ satisfies (a)}\}$ and $L_2 = \{f \in L, |\alpha_1(f)| \leqq 2p \text{ or } |\alpha_2(f)| \leqq 2p\}$. By the hypothesis that the α_i^{-1} preserve rationality and by (b), L_1 is rational. Moreover, if $\alpha_1(f)$ is short, then so is $\alpha_2(f)$ and vice versa. Hence, L_2 is contained in a finite union of languages of the form $L_w = \{f \in A^* \mid \alpha_1(f) = wg_1 \text{ and } \alpha_2(f) = wg_2\}$, which are also rational; since each L_w is contained in L, we conclude that L is rational.

It is clear that (a) implies (b). Suppose that (b) holds, that is, $\alpha_1(f) = s_1 u g_1$, $\alpha_2(f) = s_2 u g_2$ with $|s_1|, |s_2| \equiv i \bmod 2p$. We must show that $s_1 = s_2$. By adjacency, we have $\alpha_1(f) = th_1$, $\alpha_2(f) = th_2$ with $|h_1|, |h_2| < p$. As $|g_1|, |g_2| < p$, the difference between $|s_1 u|$ and $|t|$ is less than p. The case is similar for $|s_2 u|$ and $|t|$. Thus $\|s_1| - |s_2\| = \|s_1 u| - |s_2 u\| < 2p \Rightarrow |s_1| = |s_2|$.

Now, $|h_1| \leqq p \leqq |ug_1|$, which implies, by $s_1 u g_1 = th_1$, that $|s_1| \leqq |t|$, hence s_1 is a left factor of t. Similarly, s_2 is a left factor of t. As they are of equal length, they are equal.

(ii) We have, by a previous formula,

$$\alpha_1 \wedge \cdots \wedge \alpha_n = (\alpha_1 \wedge \alpha_2) \wedge \alpha_3 \wedge \cdots \wedge \alpha_n,$$

hence we may assume that $n = 2$, because each α_i is adjacent to $\alpha_1 \wedge \alpha_2$.

Without loss of generality, we may assume that $\mathrm{dom}\,(\alpha_1) = \mathrm{dom}\,(\alpha_2) = D$. Then D is a finite union of languages $D(g_1, g_2)$, where $D(g_1, g_2) = \{f \in A^* \mid \exists w \in B^*, \alpha_1(f) = wg_1, \alpha_2(f) = wg_2\}$ and where g_1 and g_2 have no common left factor. Each of these languages is rational by (i), and if $f \in D(g_1, g_2)$, then $(\alpha_1 \wedge \alpha_2)(f) = \alpha_1(f) g_1^{-1}$. Hence, the restriction of $\alpha_1 \wedge \alpha_2$ to $D(g_1, g_2)$ is rational, and finally $\alpha_1 \wedge \alpha_2$ is rational, as the union of a finite number of rational functions. □

LEMMA 7. *Let $\alpha : A^* \to B^*$ be a rational function, and $\sim$ a left congruence on A^* of finite index that is compatible with the left syntactic adjacency relation of α. Let $R = A^*/\sim$ and define for each r in R a function α_r by*

$$\alpha_r(f) = \bigwedge \{\alpha(fu) \mid u \in A^*, ur_0 = r\},$$

where r_0 is the class of 1 mod $\sim$. *Then there exists a finite language L_r such that*

(i) $u \in L_r \Rightarrow ur_0 = r$,

(ii) $\alpha_r(f) = \bigwedge_{u \in L_r} \alpha(fu)$.

As a consequence, the function α_r is rational.

The point of the lemma is that L_r does not depend on f (otherwise, it is immediate, using a previous remark on $\bigwedge L$).

Proof. Suppose there exists a finite language L_r such that (i) and (ii) are satisfied. By (i) and compatibility of $\sim$, the words in L_r are pairwise in relation $\leftrightarrow$, that is, the functions $f \mapsto \alpha(fu)$ are, for u in L_r, pairwise adjacent. Since these functions are rational, we obtain by (ii) and Lemma 6(ii) that α_r is a rational function.

In order to prove that there exists a finite language L_r satisfying (i) and (ii), take β_i, γ_i as in Lemma 5. By condition (ii) of this lemma, there exists for each i a unique $r(i)$ such that $u \in \mathrm{dom}\,(\gamma_i) \Rightarrow ur_0 = r(i)$. We know that for each i, there exists a finite language $L_i \subset \mathrm{dom}\,(\gamma_i)$ such that $\bigwedge \gamma_i(A^*) = \bigwedge \gamma_i(L_i)$. Let $L_r = \bigcup_{r(i)=r} L_i$. We thus have $\bigwedge \{\gamma_i(u) \mid u \in \mathrm{dom}\,(\gamma_i)\} = \bigwedge \{\gamma_i(u) \mid u \in L_r\}$. Moreover, (i) holds by definition. We have also

$$\begin{aligned}
\alpha_r(f) &= \bigwedge \{\alpha(fu) \mid ur_0 = r\} \\
&= \bigwedge \{\beta_i(f)\gamma_i(u) \mid ur_0 = r, 1 \leqq i \leqq n \text{ and } u \in \mathrm{dom}\,(\gamma_i)\} \\
&= \bigwedge \{\beta_i(f)\gamma_i(u) \mid r(i) = r, u \in \mathrm{dom}\,(\gamma_i)\} \\
&= \bigwedge_{r(i)=r} (\bigwedge \{\beta_i(f)\gamma_i(u) \mid u \in \mathrm{dom}\,(\gamma_i)\}) \\
&= \bigwedge_{r(i)=r} \beta_i(f)(\bigwedge \{\gamma_i(u) \mid u \in \mathrm{dom}\,(\gamma_i)\}) \\
&= \bigwedge_{r(i)=r} \beta_i(f)(\bigwedge \{\gamma_i(u) \mid u \in L_r\}) \\
&= \bigwedge_{r(i)=r} (\bigwedge \{\beta_i(f)\gamma_i(u) \mid u \in L_r\}) \\
&= \bigwedge \{\beta_i(f)\gamma_i(u) \mid r(i) = r, u \in L_r\} \\
&= \bigwedge \{\alpha(fu) \mid u \in L_r\} \quad (\text{because } u \in L_r \text{ and } u \in \mathrm{dom}\,(\gamma_i) \Rightarrow r(i) = r) \\
&= \bigwedge_{u \in L_r} \alpha(fu).
\end{aligned}$$

□

LEMMA 8 (Notations of Lemma 7). *There exist a function* $\omega : A^* \times A^* \times R \to B^*$ *and a function* $\rho : A^* \to B^*$ *such that*

(i) *For any words* f, g *in* A^* *and state* r *in* R

$$\alpha_r(fg) = \alpha_{gr}(f)\omega(f, g, r);$$

(ii) *For any word* f *in* A^*

$$\alpha(f) = \alpha_{r_0}(f)\rho(f).$$

Proof. The second assertion is immediate, because by definition, $\alpha_{r_0}(f)$ is a left factor of $\alpha(f)$. If $\alpha(f) \neq \emptyset$, we define $\rho(f) = (\alpha_{r_0}(f))^{-1}\alpha(f)$. If $\alpha(f) = \emptyset$, we pose $\rho(f) = \emptyset$. Note that the set

$$\{\alpha(fgu) \mid u \in A^*, ur_0 = r\}$$

is contained in the set

$$\{\alpha(fv) \mid v \in A^*, vr_0 = gr\}.$$

Hence, by definition, $\alpha_{gr}(f)$ is a left factor of $\alpha_r(fg)$. If $\alpha_r(fg) \neq \emptyset$, we define $\omega(f, g, r) = (\alpha_{gr}(f))^{-1}\alpha_r(fg)$. If $\alpha_r(fg) = \emptyset$, we pose once again $\omega(f, g, r) = \emptyset$. □

LEMMA 9 (Notations of Lemma 7). *Define a relation* $\equiv$ *on* A^* *by*

$$f \equiv g$$

if and only if

$$\omega(fu, a, r) = \omega(gu, a, r)$$

for any word $u \in A^*$, *letter* $a \in A$, *and state* r *in* R, *and if*

$$\rho(fu) = \rho(gu)$$

for any word u. *Then* $\equiv$ *is a right congruence of finite index.*

Proof. Recall that when $\delta_1, \cdots, \delta_p$ are functions $A^* \to B^*$ such that

(i) Each $\delta_i(A^*)$ is finite;

(ii) For each g in B^* and i, $\delta_i^{-1}(g)$ is a rational language;

then by Nerode's criterion, the right congruence on A^*, defined by $f \equiv g$ if and only if $\delta_i(fu) = \delta_i(gu)$ for any i and u, is of finite index.

Hence, it is enough to show that the functions $\omega(\cdot, a, r) : f \mapsto \omega(f, a, r)$ and ρ have finite image and that for any a, r, g in B^*, the languages $\{f \in A^* \mid \omega(f, a, r) = g\}$ and $\{f \in A^* \mid \rho(f) = g\}$ are rational.

For this, it is enough, in view of Lemma 6(i) and Lemma 7, to show that the functions $f \mapsto \alpha_r(fa)$ and $f \mapsto \alpha_{ar}(f)$ are adjacent for any $a \in A$ and $r \in R$, and that the functions α and α_{r_0} are adjacent.

By Lemma 7, we have $\alpha_r(fa) = \bigwedge_{u \in L_r} \alpha(fau)$ and $\alpha_{ar}(f) = \bigwedge_{v \in L_{ar}} \alpha(fv)$.

Note that $u \in L_r$ and $v \in L_{ar}$ implies that $ur_0 = r \Rightarrow aur_0 = ar$, and $vr_0 = ar$. Hence $au \sim v$, which implies $au \leftrightarrow v$ and the functions $f \mapsto \alpha(fau)$ and $f \mapsto \alpha(fv)$ are adjacent. Moreover, for $w, w' \in L_r$, one has $w \sim w'$ (by Lemma 7 (i)), hence $w \leftrightarrow w'$ (by compatibility of $\sim$), hence the functions $f \mapsto \alpha(fw)$ and $f \mapsto \alpha(fw')$ are adjacent. This shows that the functions $f \mapsto \alpha_r(fa)$ and $f \mapsto \alpha_{ar}(f)$ are adjacent, because of the following easily verified fact: if $\alpha_1, \cdots, \alpha_n$ (respectively, $\beta_1, \cdots, \beta_p$) are pairwise adjacent, and if each α_i is adjacent to each β_j, then $\alpha_1 \wedge \cdots \wedge \alpha_n$ is adjacent to $\beta_1 \wedge \cdots \wedge \beta_p$.

Moreover, $\alpha_{r_0}(f) = \bigwedge_{u \in L_{r_0}} \alpha(fu)$ and a similar proof shows that this function is adjacent to α. □

Proof of Theorem 2 (Second part). Let $L = A^*/\equiv$, where $\equiv$ is the right congruence of Lemma 9. Then L is finite, and equipped with a right action $L \times A^* \to L$. For l in L, a in A, and r in R, we may define $\omega(l, a, r) = \omega(f, a, r)$ and $\rho(l) = \rho(f)$, where f is a representative of $l \bmod \equiv$.

Let l_0 be the class of 1 mod $\equiv$. Define a function $\lambda : R \to B^*$ by $\lambda(r) = \alpha_r(1)$.

With these pointed sets (L, l_0), (R, r_0) and functions ω, λ, ρ, we obtain a bimachine for which we have only to verify that it computes α, that is, formulas (5) and (6). For this, it is enough to show that the functions ω and ρ of Lemma 8 satisfy

$$\omega(f, gh, r) = \omega(f, g, hr)\omega(fg, h, r) \tag{9}$$

and

$$\alpha(f) = \lambda(fr_0)\omega(1, f, r_0)\rho(f). \tag{10}$$

But we have, by Lemma 8,

$$\alpha_r(fgh) = \alpha_{ghr}(f)\omega(f, gh, r)$$

and

$$\begin{aligned}\alpha_r(fgh) &= \alpha_{hr}(fg)\omega(fg, h, r)\\ &= \alpha_{ghr}(f)\omega(f, g, hr)\omega(fg, h, r).\end{aligned}$$

So (9) is true as soon as $\alpha_{ghr}(f) \neq \varnothing$. When $\alpha_{ghr}(f) = \varnothing$, then $\alpha_r(fgh) = \varnothing$, and by the definition of ω, we have $\omega(fg, h, r) = \varnothing = \omega(f, gh, r)$. So, (9) is also true.

For (10), we have, by Lemma 8,

$$\alpha(f) = \alpha_{r_0}(f)\rho(f) = \alpha_{r_0}(1 \cdot f)\rho(f) = \alpha_{fr_0}(1)\omega(1, f, r_0)\rho(f) = \lambda(fr_0)\omega(1, f, r_0)\rho(f),$$

which proves (10). □

Remark 4. (1) Note that when r_0 in R is replaced by r, and ρ by the constant function ρ' equal to 1, then this new bimachine computes α_r. Indeed, by Lemma 8,

$$\alpha_r(f) = \alpha_r(1 \cdot f) = \alpha_{fr}(1)\omega(1, f, r) = \lambda(fr)\omega(l_0, f, r)\rho'(l_0 f).$$

(2) When α is a subsequential function, then its left syntactic adjacency is universal (i.e., $u \leftrightarrow v$ for any word u, v), hence a left congruence. If one takes this congruence for $\sim$ in Theorem 2, then the bimachine constructed in the proof is exactly the minimal subsequential transducer of α, as constructed by Choffrut [4] (see also [11]).

5. Example, remarks, and open problems. (a) Let $A = \{a, b\}$ and $\alpha : A^* \to A^*$ be the function which removes odd runs in a word. More formally, if

$$w = a^{i_1}b^{j_1} \cdots a^{i_k}b^{j_k}$$

where the exponents are greater than or equal to 1, except possibly i_1 and j_k, then define

$$i'_s = \begin{cases} i_s & \text{if } i_s \text{ is even} \\ 0 & \text{otherwise;} \end{cases}$$

$$j'_s = \begin{cases} j_s & \text{if } j_s \text{ is even} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\alpha(w) = a^{i'_1}b^{j'_1} \cdots a^{i'_k}b^{j'_k}$. Moreover, $\alpha(1) = 1$.

We leave to the reader the verification of the following facts.

(1) The left syntactic congruence $\sim$ of α is generated by the relations

$$a^2 \sim 1, \qquad b^2 \sim 1, \qquad ab \sim a, \qquad ba \sim b.$$

(2) Identify $R = A^*/\sim$ with $\{1, a, b\}$. The functions $\alpha_1, \alpha_a, \alpha_b$ are defined by $\alpha_1 = \alpha$, $\alpha_a(f) = \alpha(fa)$, $\alpha_b(f) = \alpha(fb)$.

(3) The function ρ is constant and equal to 1, and

$$\omega(f, a, 1) = \omega(f, b, 1) = \omega(f, a, b) = \omega(f, b, a) = 1.$$

Moreover,

$$\omega(f, a, a) = \begin{cases} a^2 & \text{if the last run of } f \text{ is an even run of } a\text{'s or if} \\ & f \text{ does not end with } a. \\ 1 & \text{otherwise;} \end{cases}$$

$$\omega(f, b, b) = \begin{cases} b^2 & \text{if the last run of } f \text{ is an even run of } b\text{'s or if} \\ & f \text{ does not end with } b. \\ 1 & \text{otherwise.} \end{cases}$$

(4) The right congruence $\equiv$ is generated by the relations

$$a^2 \equiv 1, \qquad b^2 \equiv 1, \qquad ab \equiv b, \qquad ba \equiv a.$$

Actually, it is the *right* syntactic congruence of α (this is not a general fact, even for everywhere-defined functions).

(5) The function λ is constant equal to 1 and if $L = A^*/\equiv$ is identified with $\{1, a, b\}$, then ω is described by the following tables

l \ r	1	a	b
1	1	a^2	1
a	1	1	1
b	1	a^2	1

$\omega(l, a, r)$

l \ r	1	a	b
1	1	1	b^2
a	1	1	b^2
b	1	1	1

$\omega(l, b, r)$

5.1. Minimization. We verify that, when α is a *total function*, then the bimachine constructed in the proof of Theorem 2 has the minimum number of left states among all bimachines computing α, with R as a set of right states (with its natural left action), with r_0 as initial right state.

So let α be computed by the bimachine B' with a set of left states L', initial left state l'_0, set of right states R, initial right state r_0, output function ω', final left function λ', and final right function ρ'.

We show that for any words g, f in A^*, the equality $l'_0 f = l'_0 g$ implies $f \equiv g$ (where $\equiv$ is the right congruence of Lemma 9). This will imply that $L = A^*/\equiv$ has fewer elements than $l'_0 A^*$, hence fewer than L' (because $l'_0 A^* \subset L$).

We work in the free group generated by A. With the notations of Lemma 7, we have

$$\alpha_r(f) = \bigwedge \{\alpha(fu) \mid u \in A^*, ur_0 = r\}. \tag{11}$$

By (5) and (6) applied to bimachine B', we have

$$\alpha(fu) = \lambda'(fur_0)\omega'(l'_0, f, ur_0)\omega'(l'_0 f, u, r_0)\rho'(l'_0 fu).$$

This, along with (11), implies that

$$\alpha_r(f) = \lambda'(fr)\omega'(l'_0, f, r)\beta(l'_0 f, r) \tag{12}$$

where $\beta\colon L' \times R \to B^*$ is the function defined by

$$\beta(l', r) = \bigwedge \{\omega'(l', u, r_0)\rho'(l'u) \mid ur_0 = r\}.$$

From (12) we deduce

$$\begin{aligned}\alpha_r(fg) &= \lambda'(fgr)\omega'(l'_0, fg, r)\beta(l'_0 fg, r)\\ &= \lambda'(fgr)\omega'(l'_0, f, gr)\omega'(l'_0 f, g, r)\beta(l'_0 fg, r),\end{aligned} \tag{13}$$

where we have used (5) again. From (12) again, we deduce

$$\alpha_{gr}(f)\omega(f, g, r) = \lambda'(fgr)\omega'(l'_0, f, gr)\beta(l'_0 f, gr)\omega(f, g, r). \tag{14}$$

Recall that we have, by Lemma 8(i),

$$\alpha_r(fg) = \alpha_{gr}(f)\omega(f, g, r).$$

Using this and comparing (13) and (14), we therefore deduce that

$$\omega(f, g, r) = \beta(l'_0 f, gr)^{-1}\omega'(l'_0 f, g, r)\beta(l'_0 fg, r). \tag{15}$$

Indeed, α is a total function, so α_r and α_{gr} are total functions as well, and every factor in (13) and (14) is defined; we thus may simplify by $\lambda'(fgr)\omega'(l'_0, f, gr)$, and multiply (in the free group) by $\beta(l'_0 f, gr)^{-1}$.

By Lemma 8(ii), we have

$$\alpha(f) = \alpha_{r_0}(f)\rho(f).$$

As α is computed by β', and by (12), we thus obtain

$$\lambda'(fr_0)\omega'(l'_0, f, r_0)\rho'(l_0 f) = \lambda'(fr_0)\omega'(l'_0, f, r_0)\beta(l'_0 f, r_0)\rho(f).$$

Thus, we deduce

$$\rho(f) = \beta(l'_0 f, r_0)^{-1}\rho'(l_0 f). \tag{16}$$

Now, let f, g, u, a, r be as in Lemma 9, and suppose that $l'_0 f = l'_0 g$. Then by (15), used twice (with $f \to fu$, $g \to a$, and after $f \to gu$, $g \to a$), we obtain

$$\begin{aligned}\omega(fu, a, r) &= \beta(l'_0 fu, ar)^{-1}\omega'(l'_0 fu, a, r)\beta(l'_0 fua, r)\\ &= \beta(l'_0 gu, ar)^{-1}\omega'(l'_0 gu, a, r)\beta(l'_0 gua, r) = \omega(gu, a, r).\end{aligned}$$

Moreover, by (16), we have

$$\begin{aligned}\rho(fu) &= \beta(l'_0 fu, r_0)^{-1}\rho'(l_0 fu)\\ &= \beta(l'_0 gu, r_0)^{-1}\rho'(l_0 gu) = \rho(gu).\end{aligned}$$

This shows, by Lemma 9, that $f \equiv g$, which was to be shown.

5.2. Counterexample. We show that when α is not a total function, then the minimization result of § 5.1 is no longer valid. This is a mystery which should be elucidated elsewhere.

Let $\alpha : a^* \to a^*$ be defined by $\alpha(a^{2n}) = a^{2n}$, $\alpha(a^{2n+1}) = \varnothing$. Take $R = a^*/a^2 \sim 1$ ($\sim$ is the syntactic left congruence) and identify R with $\{1, a\}$. Then

$$\begin{aligned}
\alpha_1(1) &= \bigwedge \{\alpha(u), u.1 = 1\} \\
&= \bigwedge \{\alpha(a^{2n}), n \in \mathbb{N}\} = 1 \\
\alpha_1(a) &= \bigwedge \{\alpha(au), u.1 = 1\} \\
&= \bigwedge \{\alpha(aa^{2n}), n \in \mathbb{N}\} = \varnothing \\
\alpha_a(a^2) &= \bigwedge \{\alpha(a^2u), u.1 = a\} \\
&= \bigwedge \{\alpha(a^2a^{2n+1}), n \in \mathbb{N}\} = \varnothing \\
\alpha_a(a) &= \bigwedge \{\alpha(au), u.1 = a\} \\
&= \bigwedge \{\alpha(aa^{2n+1}), n \in \mathbb{N}\} = a^2.
\end{aligned}$$

Using Lemma 8, we have $\alpha_a(a) = \alpha_1(1)\ \omega(1, a, a)$ and $\alpha_a(a^2) = \alpha_1(a)\ \omega(a, a, a)$. Hence, $\omega(1, a, a) = a^2$, and $\omega(a, a, a) = \varnothing$ (see the proof of Lemma 8). We deduce, by Lemma 9, that $a \not\equiv 1$.

However, the function is subsequential in both directions, hence, it may be computed with R as a set of right states, and a trivial set of left states (i.e., a singleton).

The reader may find it instructive to compare the previous example to the two following ones:

$$\begin{cases} a^{2n} \to a^{2n} \\ a^{2n+1} \to b^{2n+1} \end{cases} \qquad \begin{cases} a^{2n} \to 1 \\ a^{2n+1} \to a \end{cases}.$$

The first function is not subsequential, in either direction, while the second is subsequential in both directions.

5.3. Open problem. A theory of morphisms between bimachines computing the same function α should be developed, keeping in mind the following possible conjecture: there are only a finite number of minimal bimachines computing α (minimal would mean universally attractive in the category of these bimachines).

One cannot expect a single minimal bimachine: evidence for this is given by the rational languages; there is no "morphic" relation between the left and the right minimal automaton.

5.4. Open problem. A bimachine has two sets of states, hence there are two finite monoids attached to it. Call a bimachine *aperiodic* such that these monoids are aperiodic (i.e., with trivial subgroups, or period equal to 1). Characterize the rational functions α, which are computed by some aperiodic bimachine. A tentative conjecture could be: α is as above if and only if for any rational language L, the period of $\alpha^{-1}(L)$ divides that of L (recall that p is a period of L if the cardinality of each cyclic subgroup of the syntactic monoid of L divides p).

More generally, a theory of varieties of rational functions could be made, as has been done for rational languages and finite monoids [10]. A first step would be to study sequential and subsequential functions.

5.5. Open problem. Characterize rational functions which are both left-to-right and right-to-left subsequential. These functions simultaneously generalize rational languages (by their characteristic function) and biprefix codes (by their decoding functions).

An answer in the case of numerical functions (i.e., with image in a cyclic free monoid) has been given by Choffrut and Schützenberger [6].

Acknowledgments. We want to thank the two referees for many valuable comments and suggestions.

REFERENCES

[1] J. BERSTEL, *Transductions and Context-Free Languages*, Teubner, Stuttgart, Germany, 1979.
[2] J. BERSTEL AND D. PERRIN, *Theory of Codes*, Academic Press, New York, 1985.
[3] J.-M. BOË, J. BOYAT, J.-P. BORDAT, AND Y. CÉSARI, *Une caractérisation des sous-monoïdes libérables*, in Théorie des codes, D. Perrin, ed., Laboratoire d'Informatique Théorique et de Programmation, Paris, (1979), pp. 9-20.
[4] C. CHOFFRUT, *A generalization of Ginsburg and Rose's characterization of g.-s.-m. mappings*, Lecture Notes in Computer Science 71, Springer-Verlag, Berlin, New York, 1979, pp. 88-103.
[5] C. CHOFFRUT AND K. CULIK, *Properties of finite and push-down transducers*, SIAM J. Comput., 12 (1983), pp. 300-315.
[6] C. CHOFFRUT AND M. P. SCHUTZENBERGER, *Counting with rational functions*, Theoret. Comput. Sci., 58 (1988), pp. 81-101.
[7] S. EILENBERG, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974.
[8] S. GINSBURG, *An Introduction to Mathematical Machine Theory*, Addison-Wesley, Reading, MA, 1962.
[9] M. LOTHAIRE, *Combinatorics on Words*, Addison-Wesley, Reading, MA, 1983.
[10] J.-E. PIN, *Variétés de langages formels*, Masson, Paris, 1984.
[11] C. REUTENAUER, *Subsequential functions: Characterizations, minimization, examples*, in Proc. International Meeting of Young Computer Scientists, Lecture Notes in Computer Science, J. Kelemen, ed., to appear.
[12] M. P. SCHUTZENBERGER, *A remark on finite transducers*, Inform. and Control, 4 (1961), pp. 185-196.
[13] ———, *Une propriété de Hankel des relations fonctionnelles entre monoïdes libres*, Adv. in Math., 24 (1977), pp. 274-280.