INFORMATIQUE THÉORIQUE ET APPLICATIONS

ANDREAS WEBER
HELMUT SEIDL

On finitely generated monoids of matrices with entries in $\mathbb{N}$

Informatique théorique et applications, tome 25, nº 1 (1991), p. 19-38
<http://www.numdam.org/item?id=ITA_1991__25_1_19_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 25, n° 1, 1991, p. 19 à 38)

ON FINITELY GENERATED MONOIDS OF MATRICES WITH ENTRIES IN $\mathbb{N}$ (*)

by Andreas WEBER [1] and Helmut SEIDL [2]

Communicated by J. BERSTEL

Abstract. – *Let* Γ *be a nonempty, finite set of square matrices of size n with entries in the semiring* $\mathbb{N}$. *Consider the matrix-monoid* $\Gamma^* = \bigcup_{\lambda \geq 0} \Gamma^\lambda$ *generated by* Γ. *We show: If* Γ^* *is finite, then* $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ *where* $N = \ulcorner e^2 \cdot n! \urcorner - 2$. *This assertion is false for any N smaller than* 2^{n-2}. *If* Γ *has exactly one member and* Γ^* *is finite, then* $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ *where* $N = \max_{l=0}^{n} (l + g(n-l)) - 1$ *(g denotes Landau's function). In the last assertion N is V minimal.*

Résumé. – *Soit* Γ *un ensemble non vide et fini des matrices carrées de dimension n à entrées dans le semi-anneau* $\mathbb{N}$. *Considérons le monoïde de matrices* $\Gamma^* = \bigcup_{\lambda \geq 0} \Gamma^\lambda$ *engendré par* Γ. *Nous démontrons : Si* Γ^* *est fini, alors* $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ *où* $N = \ulcorner e^2 \cdot n! \urcorner - 2$. *Cette assertion est fausse pour chaque N plus petit que* 2^{n-2}. *Si* Γ *a exactement un élément et* Γ^* *est fini, alors* $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ *où* $N = \max_{l=0}^{n} (l + g(n-l)) - 1$ *(g représente la fonction de Landau). Dans la dernière assertion N est minimal.*

(*) Received August 1988, revised September 1989.

[1] Fachbereich Informatik, J.W. Goethe-Universität, Postfach 111 932, D-6000 Frankfurt am Main 11, R.F.A.

[2] Fachbereich Informatik, Universität des Saarlandes, Im Stadtwald, D-6600 Saarbrücken 11, West Germany.

0. INTRODUCTION

Let $n \in \mathbb{N} \setminus \{0\}$. Let Γ be a nonempty, finite set of $n \times n$-matrices with entries in $\mathbb{N}$ (3). $\Gamma^* = \bigcup_{\lambda \geqq 0} \Gamma^\lambda$ denotes the matrix-monoid generated by Γ. In this paper we deal with the following problem: If Γ^* is finite, – for which N does the identity $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ hold? Note that it is decidable whether or not Γ^* is finite ([MaSi77], [Ja77], [Re77]), – even if the underlying semiring $\mathbb{N}$ is replaced by $\mathbb{Q}$ [MaSi77] or by an arbitrary commutative field [Ja77]. In fact, (for the semiring $\mathbb{N}$) the decision can be made in polynomial time (*see* appendix, *see* also [We87], [Le87], [Ku88]).

The following values for N are known from the literature:

– $N = 2^{3^{3 \cdot n^2 + 1}} - 1$ [MaSi77].

– $N = f(n, \# \Gamma) - 1$ where f is a recursive function [Ja77].

– $N = (\text{entry}(\Gamma))^{n^2 \cdot (n-1)} \cdot 5^{n^3/2} \cdot n^{n^3} + n^2 - 1$ where entry (Γ) is defined as the maximum of 1 and the greatest entry of a matrix in Γ (*see* appendix, *see* also [We87], [Ku88]).

Similar values for N hold true, if the underlying semiring $\mathbb{N}$ is replaced by $\mathbb{Q}$ [MaSi77] or by an arbitrary commutative field [Ja77]. For further results on finitely generated matrix-monoids we refer to [McZ75], [MaSi77], [Ja77], [Re77], [ChI83], chapter 7 of [We87] (presented in the appendix of this paper), [Le87], and [Ku88].

In section 2 of this paper we show: If Γ^* is finite, then $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ where $N = \ulcorner e^2 \cdot n! \urcorner - 2$. For each $n \geqq 2$ there is a set Γ_n of $n \times n$-matrices with entries in $\{0, 1\}$ such that $(\Gamma_n)^*$ is finite and strictly includes $\bigcup_{\lambda=0}^{N} (\Gamma_n)^\lambda$ where $N = 2^{n-2} - 1$. In section 3 we show: If Γ has exactly one member and Γ^* is finite, then $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ where $N = \max_{l=0}^{n} (l + g(n-l)) - 1$ (g denotes Landau's function). For each $n \geqq 1$ there is an $n \times n$-matrix C_n with entries in $\{0, 1\}$ such that $\{C_n\}^*$ is finite and strictly includes $\{(C_n)^0, (C_n)^1, \ldots, (C_n)^N\}$ where $N = \max_{l=0}^{n} (l + g(n-l)) - 2$.

(3) $\mathbb{N}$ denotes the semiring of all nonnegative integers.

The second result of section 3 is essentially due to Ludwig Staiger. Indeed, only recently he slightly improved the corresponding result in a previous version of this paper up to optimality (!) and also exhibited an alternative proof, based on matrix theory, of the first result of section 3 [Sr88]. Using similar methods, another proof of the latter result was obtained by Paavo Turakainen [Tu90].

The function g is considered in number theory. Landau showed: $\lim_{n \to \infty} [\log_e (g(n))/\sqrt{n \cdot \log_e n}] = 1$ ([La09], §61). Further results on the asymptotic behavior of g can be found in [MsNRo88]. In section 3 we mention explicit upper and lower bounds for g due to Massias ([Ms84.1], [Ms84.2]).

It remains as an open problem: Where in the range between 2^{n-2} and $\ulcorner e^2 \cdot n! \urcorner - 2$ is the smallest N such that for each finite monoid Γ^* (of $n \times n$-matrices with entries in $\mathbb{N}$) the identity $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$ holds?

In our proofs we transform the above stated results into assertions on the degree of ambiguity of a finite $\mathbb{N}$-automaton ($\mathbb{N}$-FA). In section 2 we present a "non-ramification" lemma. This lemma allows to shorten an input word of a finitely ambiguous $\mathbb{N}$-FA without changing its ambiguity-behavior. The lemma and its application lead to the first result of section 2 and turn out to be a completion of methods and ideas used in [We87] and in [WeSe88]. In fact, in [Se89] the second author generalizes the above lemma to finite tree automata. In order to prove the second result of section 2 we take advantage of some properties of a finite automaton constructed in [WeSe88]. In section 3 of this paper we use direct methods and constructions.

1. PRELIMINARIES

Let K be a nonempty, finite set. $\mathbb{N}^{K \times K}$ denotes the multiplicative monoid of all square matrices with entries in $\mathbb{N}$ and both rows and columns indexed by K; the matrix multiplication is defined "as usual". Let Γ be a subset of $\mathbb{N}^{K \times K}$: $\Gamma^* := \bigcup_{\lambda \geqq 0} \Gamma^\lambda$ denotes the matrix-monoid generated by Γ. If Γ is finite, then Γ^* is said to be *finitely generated.*

The (i, j)-entry of a matrix $C \in \mathbb{N}^{K \times K}$ is denoted by $C_{i,j}$ $(i, j \in K)$. Let $c_{ij} \in \mathbb{N}$ $(i, j \in \mathrm{K})$, then the (unique) matrix $C \in \mathbb{N}^{K \times K}$ such that for all $i, j \in K$ $C_{i,j} = c_{ij}$ is denoted by $C = (c_{ij})_{i,j \in K}$. We define $\{0, 1\}^{K \times K} := \{C \in \mathbb{N}^{K \times K} \mid \forall i, j \in \mathrm{K}: C_{i,j} \in \{0, 1\}\}$. Let $n \in \mathbb{N}$: $\mathbb{N}^{n \times n}$ denotes the multiplicative monoid $\mathbb{N}^{[n] \times [n]}$ of all $n \times n$-matrices with entries in $\mathbb{N}$ ([4]). We define $\{0, 1\}^{n \times n} := \{0, 1\}^{[n] \times [n]}$.

Following [E74], we define a *finite* $\mathbb{N}$-*automaton* (short form: $\mathbb{N}$-FA) as a 5-tuple $M = (Q, \Sigma, \gamma, Q_I, Q_F)$ where Q and Σ denote nonempty, finite sets of states resp. input symbols, Q_I, $Q_F \subseteq Q$ denote sets of initial resp. final (or accepting) states, and γ is a total function $\gamma: Q \times \Sigma \times Q \to \mathbb{N}$. Σ is called the input alphabet of M, γ is called the multiplicity function of M. Each $(p, a, q) \in Q \times \Sigma \times Q$ denotes a *transition* of M with *multiplicity* $\gamma(p, a, q)$. A transition is called *proper*, if its multiplicity is nonzero. For each $a \in \Sigma$ $\gamma(a) := (\gamma(p, a, q))_{p, q \in Q} \in \mathbb{N}^{Q \times Q}$ denotes the transition matrix for a in M. If $\gamma(Q \times \Sigma \times Q) \subseteq \{0, 1\}$, then M is called a (nondeterministic) *finite automaton* (short form: FA).

The mode of operation of M is described by paths. A *path* π (of length m) for x in M leading from p to q is a word $(q_1, x_1) \ldots (q_m, x_m) q_{m+1} \in (Q \times \Sigma)^m . Q$ so that $(q_1, x_1, q_2), \ldots, (q_m, x_m, q_{m+1})$ are proper transitions of M and the equalities $x = x_1 \ldots x_m$, $p = q_1$ and $q = q_{m+1}$ hold. π is said to consume x. $\gamma(\pi) := \prod_{i=1}^{m} \gamma(q_i, x_i, q_{i+1})$ denotes the *multiplicity* of π. In particular, if $m = 0$, then $\gamma(\pi) = 1$. π is called *accepting*, if $p \in Q_I$ and $q \in Q_F$. The *language recognized* by M, denoted by $L(M)$, is the set of words consumed by all accepting paths in M.

Let $x = x_1 \ldots x_m \in \Sigma^*$ $(x_1, \ldots, x_m \in \Sigma)$, and let $p, q \in Q$: We define $\mathrm{da}_M(p, x, q)$ as the sum of the multiplicities of all paths for x in M leading from p to q. In particular, $\mathrm{da}_M(p, \varepsilon, q) = \#(\{p\} \cap \{q\})$ and, for each $a \in \Sigma$, $\mathrm{da}_M(p, a, q) = \gamma(p, a, q)$. It is easy to show by induction on m: $\mathrm{da}_M(p, x_1 \ldots x_m, q) = (\gamma(x_1) \ldots \gamma(x_m))_{p, q}$ (*see* [E74], chapter VI.6). We will use this result as a second definition of the da_M-operator. The *transition relation* of M is the set $\delta := \delta_M := \{(p, x, q) \in Q \times \Sigma^* \times Q \mid \mathrm{da}_M(p, x, q) \neq 0\}$.

The *degree of ambiguity* of $x \in \Sigma^*$ in M [short form: $\mathrm{da}_M(x)$] is defined as the sum of the multiplicities of all accepting paths for x in M, *i.e.*, $\mathrm{da}_M(x) = \sum_{p \in Q_I} \sum_{q \in Q_F} \mathrm{da}_M(p, x, q)$. The *degree of ambiguity* of M [short form: $\mathrm{da}(M)$] is the supremum of the set $\{\mathrm{da}_M(x) \mid x \in \Sigma^*\}$. M is called *finitely ambiguous*, if $\mathrm{da}(M)$ is finite.

([4]) $[n]$ denotes the set $\{1, \ldots, n\}$.

A state of M is called *useful*, if it appears on some accepting path in M; otherwise, this state is called *useless*. Useless states are irrelevant to the degree of ambiguity in M. If all states of M are useful, then M is called *trim*.

A state $p \in Q$ is said to be *connected* with a state $q \in Q$ (short form: $p \underset{M}{\leftrightarrow} q$), if some paths in M lead from p to q and from q to p. An equivalence class *w. r. t.* the relation "$\underset{M}{\leftrightarrow}$" is called a *strong component* of M. A proper transition (p, a, q) of M is called a *bridge*, if p is not connected with q.

Let $x = x_1 \ldots x_m \in \Sigma^*$ $(x_1, \ldots, x_m \in \Sigma)$. The *graph of accepting paths* for x in M [short form: $G_M(x)$] is the directed multigraph (V, E) where

$$V := \{(q, j) \in Q \times \{0, \ldots, m\} \mid \exists q_I \in Q_I, \exists q_F \in Q_F : (q_I, x_1 \ldots x_j, q) \in \delta \text{ and } (q, x_{j+1} \ldots x_m, q_F) \in \delta\},$$

$$E := \{((p, j-1), i, (q, j)) \in V \times \mathbb{N} \times V \mid j \in [m] \text{ and } (p, x_j, q) \in \delta \text{ and } i \in [\gamma(p, x_j, q)]\}$$

[an edge $((p, j-1), i, (q, j))$ is assumed to lead from vertex $(p, j-1)$ to vertex (q, j).]

Note: The number of all paths in $G_M(x)$ leading from $Q_I \times \{0\}$ to $Q_F \times \{m\}$ equals the degree of ambiguity of x in M. Each vertex of $G_M(x)$ is situated on such a path.

The connection between finite generating sets of matrix-monoids in $\mathbb{N}^{n \times n}$ and finite $\mathbb{N}$-automata with n states is established by the two following propositions:

PROPOSITION 1.1: *Let $\Gamma = \{C_1, \ldots, C_t\}$ be a nonempty, finite subset of $\mathbb{N}^{n \times n}$. We associate to Γ the $\mathbb{N}$-FA $M = ([n], \Sigma, \gamma, [n], [n])$ where $\Sigma := \{a_1, \ldots, a_t\}$ and $(\gamma(i, a_\gamma, j))_{i, j \in [n]} := C_\tau$ $(\tau \in [t])$. Then, the following assertions are true:*

(i) *Γ^* is finite, if and only if M is finitely ambiguous.*

(ii) $\forall \lambda \in \mathbb{N}, \forall C \in \Gamma^\lambda, \exists x \in \Sigma^\lambda : C = (\mathrm{da}_M(i, x, j))_{i, j \in [n]}$.

(iii) $\forall y \in \Sigma^*, \exists D \in \Gamma^{|y|} : (\mathrm{da}_M(i, y, j))_{i, j \in [n]} = D$.

PROPOSITION 1.2: *Let $M = (Q, \Sigma, \gamma, Q, Q)$ be an $\mathbb{N}$-FA. We associate to M the subset $\Gamma := \{\gamma(a) \mid a \in \Sigma\}$ of $\mathbb{N}^{Q \times Q}$. Then, the following assertions are true:*

(i) *M is finitely ambiguous, if and only if Γ^* is finite.*

(ii) $\forall y \in \Sigma^*, \exists D \in \Gamma^{|y|} : (\mathrm{da}_M(p, y, q))_{p, q \in Q} = D$.

(iii) $\forall \lambda \in \mathbb{N}, \forall C \in \Gamma^\lambda, \exists x \in \Sigma^\lambda : C = (\mathrm{da}_M(p, x, q))_{p, q \in Q}$.

Proof of proposition 1.1: By the definition of the da_M-operator we observe:

$\forall \tau_1, \ldots, \tau_m \in [t], \quad \forall i, j \in [n]:$

$$\mathrm{da}_M(i, a_{\tau_1}, \ldots, a_{\tau_m}, j) = (\gamma(a_{\tau_1}) \ldots \gamma(a_{\tau_m}))_{i,j} = (C_{\tau_1} \ldots C_{\tau_m})_{i,j}.$$

From this follows the proposition. $\square$

Proof of proposition 1.2: By the definition of the da_M-operator we know:

$\forall x_1, \ldots, x_m \in \Sigma, \quad \forall p, q \in Q: \qquad \mathrm{da}_M(p, x_1 \ldots x_m, q) = (\gamma(x_1) \ldots \gamma(x_m))_{p,q}.$

From this follows the proposition. $\square$

Landau's function g: $\mathbb{N} \to \mathbb{N}$ (*see* [LO9], § 61) is defined as follows:

$$g(n) := \max\{\mathrm{lcm}(n_1, \ldots, n_k) \mid n_1, \ldots, n_k \in \mathbb{N}\setminus\{0\}, n = n_1 + \ldots + n_k\}.$$

Note that $\mathrm{lcm}(\ ) = 1$, and thus $g(0) = 1$. Clearly, for all $n \in \mathbb{N}$, $g(n) \leqq g(n+1)$.

2. THE GENERAL CASE

In this section we prove the two following theorems:

THEOREM 2.1: *Let* $n \in \mathbb{N}\setminus\{0\}$. *Define* $N := \ulcorner e^2 \cdot n! \urcorner - 2$. *Let* Γ *be a nonempty, finite set of matrices in* $\mathbb{N}^{n \times n}$. *If* Γ^* *is finite, then* $\Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda$.

THEOREM 2.2: *Let* $n \in \mathbb{N}\setminus\{0,1\}$. *Define* $N := 2^{n-2} - 1$. *Then, a set* Γ_n *of at most* $n+2$ *matrices in* $\{0, 1\}^{n \times n}$ *effectively exists such that* $(\Gamma_n)^*$ *is finite and strictly includes* $\bigcup_{\lambda=0}^{N} (\Gamma_n)^\lambda$.

Note that in theorem 2.1 the reversal of the implication is trivially true. Theorem 2.2 means that theorem 2.1 is incorrect for any N less than 2^{n-2}. Thus, 2^{n-2} is a lower bound for the smallest possible N in theorem 2.1.

In order to prove the theorems 2.1 and 2.2 we transform them into assertions on the degree of ambiguity of an $\mathbb{N}$-FA which are stated in the lemmas 2.3 and 2.4. Using the propositions 1.1 and 1.2 we will show that theorem 2.1 resp. 2.2 follows from lemma 2.3 resp. 2.4. After that we will prove these two lemmas, successively.

LEMMA 2.3: *Let $M=(Q, \Sigma, \gamma, Q_I, Q_F)$ be a finitely ambiguous $\mathbb{N}$-FA with n states.*

Define $N := \ulcorner e^2 \cdot n! \urcorner - 2$. Then, the following assertion is true:

$$\forall x \in \Sigma^*, \quad \exists y \in \Sigma^{\leqq N}, \quad \forall q_I \in Q_I, \quad \forall q_F \in Q_F: \quad \mathrm{da}_M(q_I, x, q_F) = \mathrm{da}_M(q_I, y, q_F) \quad (^5).$$

LEMMA 2.4: *Let $n \in \mathbb{N} \setminus \{0,1\}$. Define $N := 2^{n-2} - 1$. Then, a finitely ambiguous FA $M_n = (Q, \Sigma, \gamma, Q, Q)$ with n states and $n+2$ input symbols effectively exists such that the following assertion is true:*

$$\exists y \in \Sigma^*, \quad \exists p, q \in Q, \quad \forall x \in \Sigma^{\leqq N}: \qquad \mathrm{da}_{M_n}(p, x, q) < \mathrm{da}_{M_n}(p, y, q).$$

Proof of theorem 2.1: Let $M = ([n], \Sigma, \gamma, [n], [n])$ be the $\mathbb{N}$-FA associated to Γ in proposition 1.1. We conclude from proposition 1.1 and lemma 2.3:

$$\begin{aligned} \#(\Gamma^*) < \infty \quad &\Rightarrow \quad \mathrm{da}(M) < \infty \\ &\Rightarrow \quad \forall x \in \Sigma^*, \quad \exists y \in \Sigma^{\leqq N}, \quad \forall i, j \in [n]: \qquad \mathrm{da}_M(i, x, j) = \mathrm{da}_M(i, y, j) \\ &\Rightarrow \quad \forall C \in \Gamma^*, \quad \exists D \in \bigcup_{\lambda=0}^{N} \Gamma^\lambda: \qquad C = D \\ &\Rightarrow \quad \Gamma^* = \bigcup_{\lambda=0}^{N} \Gamma^\lambda. \quad \square \end{aligned}$$

Proof of theorem 2.2: Take the FA $M_n = (Q, \Sigma, \gamma, Q, Q)$ whose existence is claimed in lemma 2.4, and consider the subset $\Gamma_n := \{\gamma(a) \mid a \in \Sigma\}$ of $\{0, 1\}^{Q \times Q}$ associated to M_n in proposition 1.2. According to lemma 2.4, $\# Q = n$ and $\#(\Gamma_n) \leqq \# \Sigma = n+2$. Lemma 2.4 claims:

$$\mathrm{da}(M_n) < \infty$$

and

$$\exists y \in \Sigma^*, \quad \exists p, q \in Q, \quad \forall x \in \Sigma^{\leqq N}: \qquad \mathrm{da}_{M_n}(p, x, q) < \mathrm{da}_{M_n}(p, y, q).$$

---

(5) $\Sigma^{\leqq N}$ denotes the set $\bigcup_{\lambda=0}^{N} \Sigma^\lambda$.

By proposition 1.2 this implies:

$$\#(\Gamma_n)^* < \infty$$

and

$$\exists D \in (\Gamma_n)^*, \quad \exists p, q \in Q, \quad \forall C \in \bigcup_{\lambda=0}^{N} (\Gamma_n)^\lambda: \qquad C_{p,q} < D_{p,q}.$$

Thus, $(\Gamma_n)^*$ is finite and strictly includes $\bigcup_{\lambda=0}^{N} (\Gamma_n)^\lambda$. □

In order to prove lemma 2.3 we give some technical definitions and we state a "non-ramification" lemma (lemma 2.5). This lemma guarantees certain pieces of a graph of accepting paths in a finitely ambiguous $\mathbb{N}$-FA to be free from ramifications of edges. This property allows to shorten a sufficiently long input word of such an $\mathbb{N}$-FA without changing its ambiguity-behavior and therefore leads to a proof of lemma 2.3. Having established lemma 2.5, we prove this lemma and lemma 2.3, successively.

Let $M=(Q, \Sigma, \gamma, Q_I, Q_F)$ be an $\mathbb{N}$-FA. Let $x=x_1 \ldots x_m \in \Sigma^* (x_1, \ldots, x_m \in \Sigma)$.

Consider the multigraph $G_M(x)=(V, E)$. Let $j \in \{0, \ldots, m\}$: We define

$$\operatorname{att}(x, j) := \{q \in Q \mid \exists q_I \in Q_I: (q_I, x_1 \ldots x_j, q) \in \delta_M\},$$
$$\operatorname{der}(x, j) := \{q \in Q \mid \exists q_F \in Q_F: (q, x_{j+1} \ldots x_m, q_F) \in \delta_M\},$$
$$\operatorname{set}(x, j) := \{q \in Q \mid (q, j) \in V\} = \operatorname{att}(x, j) \cap \operatorname{der}(x, j).$$

$\operatorname{att}(x, j)$, $\operatorname{der}(x, j)$ and $\operatorname{set}(x, j)$ denote the set of states attainable from Q_I with $x_1 \ldots x_j$, the set of states derivable to Q_F with $x_{j+1} \ldots x_m$, and the set of states at column j in $G_M(x)$, respectively. Let $j_0 \in [m]$: A pair (e_1, e_2) of edges in $G_M(x)$ is called a *ramification of edges* at column j_0 in $G_M(x)$, if e_1 and e_2 are distinct and start at the same vertex in $Q \times \{j_0 - 1\}$, *i.e.*, for some state $p_0 \in Q$ and some distinct $(i_1, q_1), (i_2, q_2) \in \mathbb{N} \times Q$, $e_1, e_2 \in E$ are of the form $e_1 = ((p_0, j_0 - 1), i_1, (q_1, j_0))$ and $e_2 = ((p_0, j_0 - 1), i_2, (q_2, j_0))$. Let $0 \leqq j_1 < j_2 \leqq m$: $G_M(x)$ is said to be *ramification-free* between columns j_1 and j_2, if there is no ramification of edges at any of the columns $j_1 + 1, \ldots, j_2$ in $G_M(x)$.

LEMMA 2.5 (Non-Ramification Lemma): *Let $M=(Q, \Sigma, \gamma, Q_I, Q_F)$ be a finitely ambiguous $\mathbb{N}$-FA. Let $x = x_1 \ldots x_m \in \Sigma^*$ $(x_1, \ldots, x_m \in \Sigma)$. Let $0 \leqq j_1 < j_2 \leqq m$ so that $\operatorname{set}(x, j_1)$ and $\operatorname{set}(x, j_2)$ coincide. Then, $G_M(x)$ is ramification-free between columns j_1 and j_2.*

Proof of lemma 2.5: Let $G_M(x)=(V, E)$. Assume that there is a ramification of edges $(e_1, e_2)\in E^2$ at some column $j_0\in\{j_1+1, \ldots, j_2\}$ in $G_M(x)$. Then, a state $p_0\in Q$ and distinct $(i_1, q_1), (i_2, q_2)\in\mathbb{N}\times Q$ exist such that $e_1=((p_0, j_0-1), i_1, (q_1, j_0))$ and $e_2=((p_0, j_0-1), i_2, (q_2, j_0))$.

Ler $t\in\mathbb{N}\setminus\{0\}$. We define $u:=x_1\ldots x_{j_1}$, $v:=x_{j_1+1}\ldots x_{j_2}$, $w:=x_{j_2+1}\ldots x_m$, and $y:=uv^t w=y_1\ldots y_l$ where $l:=|y|=m+(t-1).(j_2-j_1)$ and $y_1, \ldots, y_l\in\Sigma$. Pumping that segment of $G_M(x)$ which corresponds to v, and which contains e_1 and e_2, yields the directed multigraph $\tilde{G}=(\tilde{V}, \tilde{E})$ and the edges $e_1^{(0)}, e_2^{(0)}, \ldots, e_1^{(t-1)}, e_2^{(t-1)}\in\tilde{E}$ (see *fig.* 1):

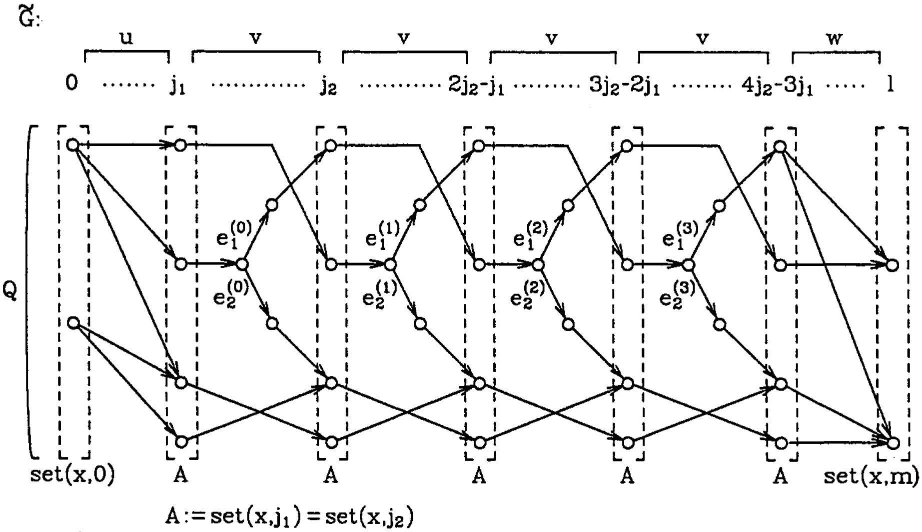

Figure 1.

$$\tilde{V}:=V\cap Q\times\{0, \ldots, j_1\}\cup\bigcup_{\tau=0}^{t-1}\{(q, j+\tau\cdot(j_2-j_1))\mid j_1\leqq j\leqq j_2, (q, j)\in V\}$$

$$\cup\{(q, j+(t-1)\cdot(j_2-j_1))\mid j_2\leqq j\leqq m, (q, j)\in V\},$$

$$\tilde{E}:=\{((p, j-1), i, (q, j))\in\tilde{V}\times\mathbb{N}\times\tilde{V}\mid j\in[1]\wedge(p, y_j, q)\in\delta_M\wedge i\in[\gamma(p, y_j, q)]\},$$

$$e_1^{(\tau)}:=((p_0, (j_0-1)+\tau\cdot(j_2-j_1)), i_1, (q_1, j_0+\tau\cdot(j_2-j_1))),$$

$$e_2^{(\tau)}:=((p_0, (j_0-1)+\tau\cdot(j_2-j_1)), i_2, (q_2, j_0+\tau\cdot(j_2-j_1)))\qquad(\tau=0, \ldots, t-1).$$

Since set (x, j_1) and set (x, j_2) coincide, each vertex of $\tilde{G}$ is situated on some path in that graph leading from $Q_I \times \{0\}$ to $Q_F \times \{1\}$, and $\tilde{G}$ is a subgraph of $G_M(y)$.

We construct pairwise distinct paths $\pi_0, \ldots, \pi_{t-1}$ in $\tilde{G}$ leading from $Q_I \times \{0\}$ to $Q_F \times \{1\}$. Let $\tau \in \{0, \ldots, t-1\}$: Select a path π_τ in $\tilde{G}$ leading from $Q_I \times \{0\}$ to $Q_F \times \{1\}$ such that π_τ runs through $e_1^{(\tau)}$ and does not run through any of the edges $e_1^{(\tau+1)}, \ldots, e_1^{(t-1)}$ [and may run through $e_2^{(\tau+1)}, \ldots, e_2^{(t-1)}$ instead]. Let $0 \leqq \sigma < \tau \leqq t-1$: π_σ and π_τ are distinct, since π_τ runs through $e_1^{(\tau)}$ and π_σ does not.

In conclusion, we know for all $t \in \mathbb{N} \setminus \{0\}$: $\mathrm{da}_M(uv^t w) \geqq t$. Hence, $\mathrm{da}(M)$ is infinite. (Contradiction!) □

In order to prove lemma 2.3 we need the following proposition:

PROPOSITION 2.6: *Let* $n \in \mathbb{N}$. *Then*, $\sum_{A \subseteq B \subseteq [n]} (\# A)! < e^2 \cdot n!$.

Proof: We estimate:

$$\sum_{A \subseteq B \subseteq [n]} (\# A)! = \sum_{d=0}^{n} \binom{n}{d} \cdot 2^{n-d} \cdot d! = n! \cdot \sum_{d=0}^{n} (2^{n-d}/(n-d)!)$$

$$= n! \cdot \sum_{d=0}^{n} (2^d/d!) < e^2 \cdot n!. \quad \square$$

Proof of lemma 2.3: We prove the lemma by induction on the length of $x \in \Sigma^*$. Let $x = x_1 \ldots x_m \in \Sigma^* (x_1, \ldots, x_m \in \Sigma)$.

Base of induction: $|x| \leqq N$. Select $y := x$.

Induction step: Let $|x| \geqq N+1 = \ulcorner e^2 \cdot n! \urcorner - 1$. Consider the multigraph $G_M(x) = (V, E)$. By proposition 2.6, a subset J of $\{0, \ldots, m\}$ and sets A and B with $A \subseteqq B \subseteqq Q$ exist such that $\# J > (\# A)!$ and $\{(\mathrm{set}(x, j), \mathrm{att}(x, j)) \mid j \in J\} = \{(A, B)\}$. Define $j_1 := \min(J)$ and $j_2 := \max(J)$. By lemma 2.5 we observe:

(*) $G_M(x)$ is ramification-free between columns j_1 and j_2.

Let us fix pairwise different states $r_1, \ldots, r_d$ such that $A = \{r_1, \ldots, r_d\}$ (see *fig.* 2).

Let $j \in J$: $\varphi(j)$ is defined to be, according to (*), the uniquely determined d-tuple $(s_1, \ldots, s_d) \in Q^d$ such that $\{s_1, \ldots, s_d\} = \mathrm{set}(x, j) = A$ and for all $i = 1, \ldots, d$ $(r_i, x_{j_1+1} \ldots x_j, s_i) \in \delta_M$. Thus, we have defined a mapping $\varphi: J \to Q^d$ such that $\# \varphi(J) \leqq d!$.

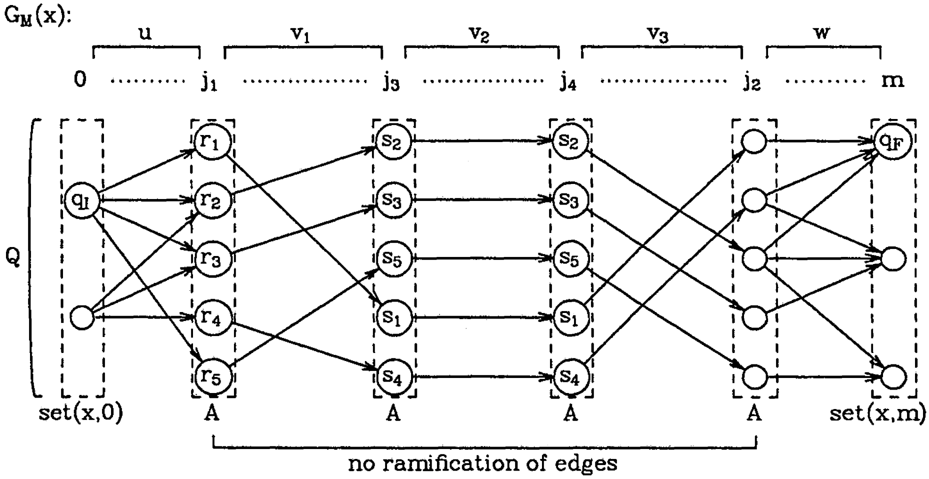

Figure 2.

Since $\# J > (\# A)! = d!$, integers $j_3, j_4 \in J$ and states $s_1, \ldots, s_d \in Q$ exist such that $j_1 \leqq j_3 < j_4 \leqq j_2$ and $\varphi(j_3) = \varphi(j_4) = (s_1, \ldots, s_d)$. Let us consider the decomposition $x = uv_1 v_2 v_3 w$ where $u := x_1 \ldots x_{j_1}$, $v_1 := x_{j_1+1} \ldots x_{j_3}$, $v_2 := x_{j_3+1} \ldots x_{j_4}$, $v_3 := x_{j_4+1} \ldots x_{j_2}$, $w := x_{j_2+1} \ldots x_m$. From the definition of φ and from (*) we derive (*see fig.* 2):

$$\{s_1, \ldots, s_d\} = \operatorname{set}(x, j_3) = \operatorname{set}(x, j_4) = A,$$

$$\forall i_1, i_2 \in [d]: \quad \mathrm{da}_M(s_{i_1}, v_2, s_{i_2}) = \#(\{i_1\} \cap \{i_2\}).$$

From the above we conclude (see *fig.* 2):

$$\forall q_I \in Q_I, \quad \forall q_F \in Q_F: \qquad \mathrm{da}_M(q_I, x, q_F)$$
$$= \sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \mathrm{da}_M(q_I, uv_1, s_{i_1}) . \mathrm{da}_M(s_{i_1}, v_2, s_{i_2}) . \mathrm{da}_M(s_{i_2}, v_3 w, q_F)$$
$$= \sum_{i=1}^{d} \mathrm{da}_M(q_I, uv_1, s_i) . \mathrm{da}_M(s_i, v_3 w, q_F).$$

We define $y := uv_1 v_3 w \in \Sigma^*$. Clearly, $|y| < |x|$. Consider the multigraph $G_M(y) = (\tilde{V}, \tilde{E})$. Since

$$\operatorname{att}(y, j_3) = \operatorname{att}(x, j_3) = B = \operatorname{att}(x, j_4)$$

and $\operatorname{der}(y, j_3) = \operatorname{der}(x, j_4)$, we observe:

$$\operatorname{set}(y, j_3) = \operatorname{att}(y, j_3) \cap \operatorname{der}(y, j_3) = \operatorname{att}(x, j_4) \cap \operatorname{der}(x, j_4)$$
$$= \operatorname{set}(x, j_4) = A = \{s_1, \ldots, s_d\}.$$

From this follows:

$\forall q_I \in Q_I, \quad \forall q_F \in Q_F$:

$$\operatorname{da}_M(q_I, y, q_F) = \sum_{i=1}^{d} \operatorname{da}_M(q_I, uv_1, s_i) \,.\, (\operatorname{da}_M(s_i, v_3 w, q_F).$$

Therefore, the assertion of the lemma follows from the induction hypothesis. □

Let M be an FA. In [We87] and in [WeSe88] the criterion (IDA) is introduced which is proved to characterize the infinite degree of ambiguity of M. The concept of a ramification of edges may be used to furnish an alternative proof of this characterization. The essential idea is to derive the criterion (IDA) from such a ramification of edges in some graph of accepting paths in M which lies between two columns with coinciding sets of states. Note that the relation "$\underset{M}{\leftrightarrow}$" is not used in this alternative proof.

For the proof of lemma 2.4 we adopt from [WeSe88] the two following propositions:

PROPOSITION 2.7 ([WeSe88], lemma 5.2): *For all $n_1, n_2 \in \mathbb{N} \setminus \{0\}$ a trim FA $M := M_{n_1, n_2} = (Q, \Sigma, \gamma, Q_I, Q_F)$ with $n_1 + n_2$ states and $n_1 + n_2 + 2$ input symbols effectively exists such that the following assertions are true*:

(i) *M has two strong components with n_1 and n_2 states, respectively. For some order Q_1, Q_2 of these components, (p_1, p_2), $(q_1, q_2) \in Q_1 \times Q_2$ exist such that $Q_I = \{p_1\}$, $Q_F = \{q_2\}$, and every bridge of M is of the form (q_1, a, p_2) where $a \in \Sigma$.*

(ii) *M is finitely ambiguous.*

(iii) *There is a word $y \in \Sigma^*$ so that $\operatorname{da}_M(y)$ is at least $2^{n_1 + n_2 - 2}$.*

PROPOSITION 2.8 ([WeSe88], assertion (*) in the proof of lemma 4.5): *Let $M = (Q, \Sigma, \gamma, Q_I, Q_F)$ be an FA with the following properties*:

(i) *M has two strong components. For some order Q_1, Q_2 of these components, (p_1, p_2), $(q_1, q_2) \in Q_1 \times Q_2$ exist such that $Q_I = \{p_1\}$, $Q_F = \{q_2\}$, and every bridge of M is of the form (q_1, a, p_2) where $a \in \Sigma$.*

(ii) *For every useful state* $q \in Q$ *and every word* $v \in \Sigma^*$ $\mathrm{da}_M(q, v, q)$ *is at most* 1.

Then, for all $x \in \Sigma^*$, $\mathrm{da}_M(x)$ *is at most* $|x|$.

Proof of lemma 2.4: Choose $n_1, n_2 \in \mathbb{N} \setminus \{0\}$ so that $n = n_1 + n_2$. Consider the FA $M := M_{n_1, n_2} = (Q, \Sigma, \gamma, Q_I, Q_F)$ whose existence is claimed in proposition 2.7. By the assertions (i) and (ii) of proposition 2.7 we can apply proposition 2.8 to M which yields for all $x \in \Sigma^*$: $\mathrm{da}_M(x) \leqq |x|$. Consider $M' := (Q, \Sigma, \gamma, Q, Q)$. M' is a finitely ambiguous FA with n states and $n+2$ input symbols [in fact, since M is trim, $\mathrm{da}(M') \leqq n^2 \cdot \mathrm{da}(M) < \infty$]. Taking $y \in \Sigma^*$ as in assertion (iii) of proposition 2.7, we observe for all $x \in \Sigma^{\leqq N}$:

$$\mathrm{da}_{M'}(p_1, x, q_2) = \mathrm{da}_M(x) \leqq |x| \leqq N = 2^{n-2} - 1 < \mathrm{da}_M(y) = \mathrm{da}_{M'}(p_1, y, q_2).$$

Thus, M' is the FA M_n we are looking for. □

3. MONOIDS GENERATED BY ONE MATRIX

In this section we prove the two following theorems the second of which is essentially due to Staiger [Sr88]:

THEOREM 3.1: *Let* $n \in \mathbb{N} \setminus \{0\}$. *Define* $N := \max_{l=0}^{n} (l + g(n-l)) - 1$. *Let* C *be a matrix in* $\mathbb{N}^{n \times n}$. *If* $\{C\}^*$ *is finite, then* $\{C\}^* = \{C^0, C^1, \ldots, C^N\}$.

THEOREM 3.2: *Let* $n \in \mathbb{N} \setminus \{0\}$. *Define* $N := \max_{l=0}^{n} (l + g(n-l)) - 2$. *Then, a matrix* C_n *in* $\{0, 1\}^{n \times n}$ *effectively exists such that* $\{C_n\}^*$ *is finite and strictly includes* $\{(C_n)^0, (C_n)^1, \ldots, (C_n)^N\}$.

Massias ([Ms84.1], [Ms84.2]) showed: There is a constant $k_1 < 1.053\,14$ such that for all $n \in \mathbb{N} \setminus \{0\}$ $g(n)$ is at most $e^{k_1 \cdot \sqrt{n \cdot \log_e n}}$, where equality holds for $n = 1319766$ (!!). Thus, in theorem 3.1 N can be replaced by $e^{k_1 \cdot \sqrt{n \cdot \log_e n}} + n - 1$. Note that in this theorem the reversal of the implication is trivially true.

Theorem 3.2 means that theorem 3.1 is incorrect for any N less than $\max_{l=0}^{n} (l + g(n-l)) - 1$. Therefore, N in theorem 3.1 is minimal. It is a result of Massias [Ms84.1] that for all $n \geqq 906$ $g(n)$ is at least $e^{\sqrt{n \cdot \log_e n}}$. Thus, in theorem 3.2 N can be replaced by $e^{\sqrt{n \cdot \log_e n}} - 2$ for all $n \geqq 906$.

In order to prove the theorems 3.1 and 3.2 we transform them into assertions on the degree of ambiguity of an $\mathbb{N}$-FA with one input symbol which are stated in the lemmas 3.3 and 3.4. Using the propositions 1.1 and 1.2 we will show that theorem 3.1 resp. 3.2 follows from lemma 3.3 resp. 3.4. After that we will prove the lemmas 3.3 and 3.4, successively. This will be done using direct methods and constructions. We want to point out that there are two other proofs of theorem 3.1, based on matrix theory, by Staiger [Sr88] and by Turakainen [Tu90]. As a side effect, this alternative proof yields a lot of knowledge about the structure of the matrices in question.

LEMMA 3.3: *Let $M=(Q, \{a\}, \gamma, Q_I, Q_F)$ be a finitely ambiguous $\mathbb{N}$-FA with n states.*

Then, for some $N \leqq \max_{l=0}^{n} (l+g(n-l))-1$, the following assertion is true:

$$\forall x \in \{a\}^*, \quad \exists y \in \{a\}^{\leqq N}, \quad \forall q_I \in Q_I, \quad \forall q_F \in Q_F:$$
$$\mathrm{da}_M(q_I, x, q_F)=\mathrm{da}_M(q_I, y, q_F).$$

LEMMA 3.4: *Let $n \in \mathbb{N} \setminus \{0\}$. Define $N:=\max_{l=0}^{n} (l+g(n-l))-2$. Then, an FA $M_n=(Q, \{a\}, \gamma, Q, Q)$ with n states effectively exists such that the following assertions are true:*

(i) *The degree of ambiguity of M_n is at most n.*

(ii) $\exists \mu \in \mathbb{N}, \forall \lambda \in \{0, \ldots, N\}, \exists p, q \in Q: \mathrm{da}_{M_n}(p, a^\lambda, q) \neq \mathrm{da}_{M_n}(p, a^\mu, q)$.

Proof of theorem 3.1: Let $M=([n], \{a\}, \gamma, [n], [n])$ be the $\mathbb{N}$-FA associated to $\Gamma=\{C\}$ in proposition 1.1. We conclude from proposition 1.1 and lemma 3.3:

$$\begin{aligned}
&\#(\{C\}^*)<\infty \quad \Rightarrow \quad \mathrm{da}(M)<\infty \\
&\quad \Rightarrow \quad \forall x \in \{a\}^*, \quad \exists y \in \{a\}^{\leqq N}, \quad \forall i, j \in [n]: \quad \mathrm{da}_M(i, x, j)=\mathrm{da}_M(i, y, j) \\
&\quad \Rightarrow \quad \forall \lambda \in \mathbb{N}, \quad \exists \mu \in \{0, \ldots, N\}: \quad C^\lambda=C^\mu \\
&\quad \Rightarrow \quad \{C\}^*=\{C^0, C^1, \ldots, C^N\}. \quad \square
\end{aligned}$$

Proof of theorem 3.2: Take the FA $M_n=(Q, \{a\}, \gamma, Q, Q)$ with n states whose existence is claimed in lemma 3.4, and consider the matrix $C_n:=\gamma(a)$ in $\{0, 1\}^{Q \times Q}$.

Note that the set $\{C_n\}$ is associated to M_n in proposition 1.2. Lemma 3.4 claims:

$$\mathrm{da}(M_n) \leqq n$$

and

$$\exists \mu \in \mathbb{N}, \quad \forall \lambda \in \{0, \ldots, N\}, \exists p, q \in Q: \qquad \mathrm{da}_{M_n}(p, a^\lambda, q) \neq \mathrm{da}_{M_n}(p, a^\mu, q).$$

According to proposition 1.2 this implies:

$\# \{C_n\}^* < \infty$

and

$$\exists \mu \in \mathbb{N}, \quad \forall \lambda \in \{0, \ldots, N\}, \quad \exists p, q \in Q: ((C_n)^\lambda)_{p, q} \neq ((C_n)^\mu)_{p, q}.$$

Thus, $\{C_n\}^*$ is finite and strictly includes $\{(C_n)^0, (C_n)^1, \ldots, (C_n)^N\}$. □

Proof of lemma 3.3: Let w.l.o.g. M be trim. A strong component U of M is called trivial, if $\gamma_{|U \times \{a\} \times U} = 0$. Let $\{q_1\}, \ldots, \{q_{n_0}\}$ resp. $Q_1, \ldots, Q_k$ be the trivial resp. nontrivial strong components of M. We define $Q_0 := \{q_1, \ldots, q_{n_0}\}$, and $n_i := \# Q_i (i = 1, \ldots, k)$.

First of all, we show:

(1) If $i, j \in [k]$ are distinct, then $\delta_M \cap Q_i \times \{a\}^* \times Q_j = \varnothing$.

(2) For all $i \in [k]$ there is a bijective mapping $\varphi_i: Q_i \to [n_i]$ such that the following holds for all $r, s \in Q_i$:

$$\gamma(r, a, s) = \begin{cases} 1 & \text{if} \quad \varphi_i(s) = \varphi_i(r) + 1 \bmod n_i \\ 0 & \text{else} \end{cases}$$

Proof of (1): Assume that, for some distinct $i, j \in [k]$, $\delta_M \cap Q_i \times \{a\}^* \times Q_j \neq \varnothing$. Choose $r \in Q_i$ and $s \in Q_j$. Then, for some $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{N} \setminus \{0\}$, (r, a^{λ_1}, r), (r, a^{λ_2}, s), $(s, a^{\lambda_3}, s) \in \delta_M$. Since $Q_i \cap Q_j = \varnothing$, this implies for all $t \in \mathbb{N}$: $\mathrm{da}_M(r, a^{\lambda_2}(a^{\lambda_1 \lambda_3})^t, s) \geqq t+1$. Hence, since r and s are useful, $\mathrm{da}(M)$ is infinite. (Contradiction!)

Proof of (2): Let $i \in [k]$ and $r \in Q_i$. Let $s_1, s_2 \in Q_i$ so that (r, a, s_1), $(r, a, s_2) \in \delta_M$. Then, for some $\lambda_1, \lambda_2 \in \mathbb{N}$, (s_1, a^{λ_1}, r), $(s_2, a^{\lambda_2}, r) \in \delta_M$. Assume that s_1 and s_2 are distinct, or that $\gamma(r, a, s_1) \geqq 2$. Then, $\mathrm{da}_M(r, v, r) \geqq 2$ where $v := a^{(1+\lambda_1) \cdot (1+\lambda_2)}$. This implies for all $t \in \mathbb{N}$: $\mathrm{da}_M(r, v^t, r) \geqq 2^t$. Hence, since r is useful, $\mathrm{da}(M)$ is infinite. (Contradiction!) Therefore,

$\sum_{s \in Q_i} \gamma(r, a, s) = 1$. From this follows (2).

Let $\varphi_1, \ldots, \varphi_k$ be as in (2). From (2) follows by induction on λ:

(3)
$$\forall i \in [k], \quad \forall r, s \in Q_i, \quad \forall \lambda \in \mathbb{N}:$$
$$\mathrm{da}_M(r, a^\lambda, s) = \begin{cases} 1 & \text{if} \quad \varphi_i(s) = \varphi_i(r) + \lambda \bmod n_i \\ 0 & \text{else} \end{cases}$$

We define the $\mathbb{N}$-FA $M_0 = (Q, \{a\}, \gamma_0, Q, Q)$:

$$\gamma_0(p, a, q) := \begin{cases} \gamma(p, a, q) & \text{if} \quad \{p, q\} \cap Q_0 \neq \varnothing \\ 0 & \text{else} \end{cases}$$

$(p, q \in Q)$.

Let $m \in \mathbb{N}$. Using $\varphi_1, \ldots, \varphi_k$ introduced in (2) we define:

$$\psi_1(m) := \{(\lambda_1, \lambda_2, i, r, s) \mid \lambda_1, \lambda_2 \in \mathbb{N}, \lambda_1 + \lambda_2 \leqq m, i \in [k], r, s \in Q_i\},$$
$$\psi_2(m) := \{(\lambda_1, \lambda_2, i, r, s) \mid \lambda_1, \lambda_2 \in \mathbb{N}, \lambda_1 + \lambda_2 \leqq n_0, i \in [k], r, s \in Q_i,$$
$$\varphi_i(s) = \varphi_i(r) + m - (\lambda_1 + \lambda_2) \bmod n_i\}.$$

If $m \geqq n_0$, then (1) and (3) imply:

(4) $\forall q_I \in Q_I, \quad \forall q_F \in Q_F: \qquad \mathrm{da}_M(q_I, a^m, q_F)$
$$= \sum_{(\lambda_1, \lambda_2, i, r, s) \in \psi_1(m)} \mathrm{da}_{M_0}(q_I, a^{\lambda_1}, r) \cdot \mathrm{da}_M(r, a^{m-(\lambda_1+\lambda_2)}, s) \cdot \mathrm{da}_{M_0}(s, a^{\lambda_2}, q_F)$$
$$= \sum_{(\lambda_1, \lambda_2, i, r, s) \in \psi_2(m)} \mathrm{da}_{M_0}(q_I, a^{\lambda_1}, r) \cdot \mathrm{da}_{M_0}(s, a^{\lambda_2}, q_F).$$

Let $x, y \in \{a\}^*$. If $|x| = |y| \bmod \mathrm{lcm}(n_1, \ldots, n_k)$, then $\psi_2(|x|) = \psi_2(|y|)$. Therefore, (4) implies:

(5) $\forall x, y \in \{a\}^*: \quad |x|, |y| \geqq n_0 \ ad\ |x| = |y| \bmod \mathrm{lcm}(n_1, \ldots, n_k)$
$$\Rightarrow \quad \forall q_I \in Q_I, \quad \forall q_F \in Q_F: \qquad \mathrm{da}_M(q_I, x, q_F) = \mathrm{da}_M(q_I, y, q_F).$$

Thus, defining $N := n_0 + \mathrm{lcm}(n_1, \ldots, n_k) - 1$, the lemma follows from (5). □

Proof of lemma 3.4: Let $n_0 \in \mathbb{N}$ and $n_1, \ldots, n_k \in \mathbb{N} \setminus \{0\}$ so that

$$n = n_0 + n_1 + \ldots + n_k,$$
$$n_0 + g(n - n_0) = \max_{l=0}^{n} (l + g(n - l)) = N + 2, \text{ and } g(n - n_0) = \mathrm{lcm}(n_1, \ldots, n_k).$$

We construct an FA $M_n = (Q, \{a\}, \gamma, Q, Q)$ with n states:

$$Q := \bigcup_{i=0}^{k} Q_i, \qquad Q_i := \{q_{i,1}, \ldots, q_{i,n_i}\} \quad (i=0, \ldots, k),$$

$$\gamma(q_{i_1,j_1}, a, q_{i_2,j_2}) := \begin{cases} 1 & \text{if} \quad i_1 = i_2 \in [k] \text{ and } j_2 = j_1 + 1 \bmod n_{i_1} \\ & \quad \text{or } i_1 = i_2 = 0 \text{ and } j_2 = j_1 + 1 \\ 0 & \quad \text{else} \end{cases}$$

$$(i_1, i_2 \in \{0, \ldots, k\}, j_1 \in [n_{i_1}], j_2 \in [n_{i_2}])$$

Let $\lambda \in \mathbb{N}$. We observe:

(6) $\forall i_1, i_2 \in \{0, \ldots, k\}, \quad \forall j_1 \in [n_{i_1}], \quad \forall j_2 \in [n_{i_2}]$:

$$\mathrm{da}_{M_n}(q_{i_1,j_1}, a^{\lambda}, q_{i_2,j_2}) = \begin{cases} 1 & \text{if} \quad i_1 = i_2 \in [k] \text{ and } j_2 = j_1 + \lambda \bmod n_{i_1} \\ & \quad \text{or } i_1 = i_2 = 0 \text{ and } j_2 = j_1 + \lambda \\ 0 & \quad \text{else} \end{cases}$$

From (6) follows: $\mathrm{da}_{M_n}(a^{\lambda}) = \max\{0, n_0 - \lambda\} + \sum_{i=1}^{k} n_i \leqq n$. Therefore, M_n has property (i) claimed in the lemma.

Let $\lambda \in \mathbb{N}$ such that for all $p, q \in Q$ $\mathrm{da}_{M_n}(p, a^{\lambda}, q) = \mathrm{da}_{M_n}(p, a^{N+1}, q)$. According to (6) this implies:

$$(\forall i \in [k], \forall j_1, j_2 \in [n_i] : j_2 = j_1 + \lambda \bmod n_i \Leftrightarrow j_2 = j_1 + N + 1 \bmod n_i)$$

and

$$(\forall j_1, j_2 \in [n_0] : j_2 = j_1 + \lambda \Leftrightarrow j_2 = j_1 + N + 1).$$

Since $N + 1 \geqq n_0$, this implies:

$$(\forall i \in [k] : \lambda = N + 1 \bmod n_i) \qquad \text{and} \qquad \lambda \geqq n_0.$$

Hence, $\lambda = N + 1 \bmod g(n - n_0)$ and $\lambda \geqq n_0$. Since $N + 1 = n_0 + g(n - n_0) - 1$, this implies: $\lambda > N$.

Thus, we have shown:

$$\forall \lambda \in \{0, \ldots, N\}, \quad \exists p, q \in Q : \qquad \mathrm{da}_{M_n}(p, a^{\lambda}, q) \neq \mathrm{da}_{M_n}(p, a^{N+1}, q).$$

This proves that M_n has property (ii) claimed in the lemma. $\square$

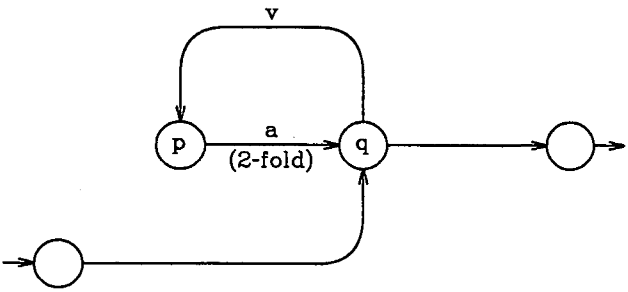

Figure 3.

APPENDIX

For the sake of completeness we report here on chapter 7 of [We87]. Indeed, we apply two basic results on the degree of ambiguity of finite automata presented in [WeSe88] to finitely generated matrix-monoids.

Let $M=(Q, \Sigma, \gamma, Q_I, Q_F)$ be an $\mathbb{N}$-FA with n states. We define entry $(M):=\max(\{1\} \cup \gamma(Q \times \Sigma \times Q))$ and the FA $\underline{M}:=(Q, \Sigma, \underline{\gamma}, Q_I, Q_F)$ where $\underline{\gamma}(p, a, q):=\min\{1, \gamma(p, a, q)\}$ $((p, a, q) \in Q \times \Sigma \times Q)$. Let $Q_1, \ldots, Q_k \subseteqq Q$ be those strong components of M which contain only useful states (note that $k \leqq n$).

Let us assume that, for some $U \in \{Q_1, \ldots, Q_k\}$, $\gamma(U \times \Sigma \times U) \nsubseteq \{0, 1\}$. Let $(p, a, q) \in U \times \Sigma \times U$ and $v \in \Sigma^*$ so that $\gamma(p, a, q) \geqq 2$ and $(q, v, p) \in \delta_M$ (see *fig.* 3). Then, we observe for all $i \in \mathbb{N}$: $\mathrm{da}_M(q, (va)^i, q) \geqq 2^i$. Hence, since q is useful, $\mathrm{da}(M)$ is infinite.

Now we assume that $\bigcup_{i=1}^{k} \gamma(Q_i \times \Sigma \times Q_i) \subseteqq \{0, 1\}$. Let π be an accepting path in M (or equivalently, in $\underline{M}$). π only visits equivalence classes from $\{Q_1, \ldots, Q_k\}$ and each such class at most once. Thus, according to the assumption, π has multiplicity at most $[\text{entry}(M)]^{k-1}$. From this follows for all $x \in \Sigma^*$: $\mathrm{da}_{\underline{M}}(x) \leqq \mathrm{da}_M(x) \leqq [\text{entry}(M)]^{k-1} . \mathrm{da}_{\underline{M}}(x)$. Hence, we know: $\mathrm{da}(\underline{M}) \leqq \mathrm{da}(M) \leqq [\text{entry}(M)]^{k-1} . \mathrm{da}(\underline{M})$.

Summarizing the above, we have shown:

LEMMA A.1: *Let M be an $\mathbb{N}$-FA as above. Then, the following assertions are true:*

(i) $\mathrm{da}(M) < \infty \Rightarrow \mathrm{da}(M) \leqq [\mathrm{entry}(M)]^{k-1} \cdot \mathrm{da}(\underline{M}) < \infty$.

(ii) $\mathrm{da}(M) = \infty \Leftrightarrow$

$$(\exists\, U \in \{Q_1, \ldots, Q_k\}:\ \gamma(U \times \Sigma \times U) \nsubseteq \{0, 1\} \vee da(\underline{M}) = \infty). \quad \square$$

Let $n \in \mathbb{N} \setminus \{0\}$. Let Γ be a nonempty, finite set of matrices in $\mathbb{N}^{n \times n}$. We define entry $(\Gamma) := \max(\{1\} \cup \{C_{i,j} \mid C \in \Gamma, i, j \in [n]\})$ and $\|\Gamma^*\| := \sup\left\{\sum_{i,j=1}^{n} C_{i,j} \mid C \in \Gamma^*\right\}$. From lemma A.1, proposition 1.1, and from the theorems 2.1 and 3.2 of [WeSe88] follows:

THEOREM A.2 ([We87], theorems 7.1-7.3; see also [Ku88]): *Let $\Gamma \subseteq \mathbb{N}^{n \times n}$ be as above. Then, the following assertions are true*:

(i) *If Γ^* is finite, then $\|\Gamma^*\|$ is at most $[entry(\Gamma)]^{n-1} \cdot 5^{n/2} \cdot n^n$.*

(ii) *It is decidable in time $O(n^6 . \# \Gamma)$ whether or not Γ^* is infinite.*

(iii) *If Γ^* is finite, then $\# (\Gamma^*)$ is at most $[entry\ (\Gamma)]^{n^2 \cdot (n-1)} \cdot 5^{n^3/2} \cdot n^{n^3} + n^2$.*

ACKNOWLEDGMENTS

We thank Joachim Herzog for communicating to us results on Landau's function. We gratefully appreciate Ludwig Staiger's interest in this work. In particular, he sharpened our original lower bound for N in the one matrix case up to optimality. A clever observation of the referee inspired us to improve the upper bound in the general case.

REFERENCES

ChI83. T.-H. Chan and O. Ibarra, On the Finite-Valuedness Problem for Sequential Machines, *TCS*, 1983, *23*, pp. 95-101.

E74. S. Eilenberg, Automata, Languages, and Machines, Academic Press, New York, N.Y., 1974, *A*.

Hs78. M. Harrison, Introduction to Formal Language Theory, Addison-Wesley, Reading, Mass., 1978.

HoU79. J. Hopcroft and J. Ullman, Introduction to Automata Theory, Languages and Computation, Addison-Wesley, Reading, Mass., 1979.

Ja77. G. Jacob, Un algorithme calculant le cardinal, fini ou infini, des demi-groupes de matrices, *T.C.S.*, 1977, *5*, pp. 183-204.

Ku88. W. Kuich, Finite Automata and Ambiguity, Report 253 of the IIG, Technische Universität Graz, 1988.

La09. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Teubner, Leipzig, 1909.

Le87. H. Leung, An Algebraic Method for Solving Decision Problems in Finite Automata Theory, *Ph. D. Thesis*, The Pennsylvania State University, 1987.

MaSi77. A Mandel and I. Simon, On Finite Semigroups of Matrices, *T.C.S.*, 1977, *5*, pp. 101-111.

Ms84.1. J.-P. Massias, Ordre maximum d'un élément du groupe symétrique et applications, *Thèse 3e cycle*, Université de Limoges, 1984.

Ms84.2. J.-P. Massias, Majoration explicite de l'ordre maximum d'un élément du groupe symétrique, *Annales Faculté des Sciences Toulouse*, 1984, *VI*, pp. 269-281.

MsNRo88. J.-P. Massias, J.-L. Nicolas and G. Robin, Evaluation asymptotique de l'ordre maximum d'un élément du groupe symétrique, *Acta Arithmetica*, 1988, *50*, pp. 221-242.

McZ75. R. McNaughton and Y. Zalcstein, The Burnside Problem for Semi-groups, *J. Algebra*, 1975, *34*, pp. 292-299.

Re77. C. Reutenauer, Propriétés arithmétiques et topologiques de séries rationnelles en variables non commutatives, *Thèse 3e cycle*, Université Paris-VI, 1977.

Se89. H. Seidl, On the Finite Degree of Ambiguity of Finite Tree Automata, *Acta Informatica*, 1989, *26*, pp. 527-542.

Sr88. L. Staiger, personal communication.

WeSe86. A. Weber and H. Seidl, On the Degree of Ambiguity of Finite Automata, *Proc. M.F.C.S.*, 1986, in: *L.N.C.S. 233*, Springer-Verlag, pp. 620-629.

Tu90 P. Turakainen, On the Finitness of the Multiplicative Monoid generated by a Nonnegative Matrix. Bull. EATCS, 1990, *40*, pp. 270-272.

We87. A. Weber, Über die Mehrdeutigkeit und Wertigkeit von endlichen Automaten und Transducern, *Dissertation*, Goethe-Universität Frankfurt am Main, 1987.

WeSe88. A. Weber and H. Seidl, On the Degree of Ambiguity of Finite Automata, Preprint, Goethe-Universität Frankfurt am Main, 1988, *T.C.S.* (to appear).