

# JOURNAL

de Théorie des Nombres  
de BORDEAUX

*anciennement Séminaire de Théorie des Nombres de Bordeaux*

Julien CASSAIGNE et Marion LE GONIDEC

**Propriétés et limites de la reconnaissance d'ensembles d'entiers par automates dénombrables**

Tome 22, n° 2 (2010), p. 307-338.

<[http://jtnb.cedram.org/item?id=JTNB\\_2010\\_\\_22\\_2\\_307\\_0](http://jtnb.cedram.org/item?id=JTNB_2010__22_2_307_0)>

© Université Bordeaux 1, 2010, tous droits réservés.

L'accès aux articles de la revue « Journal de Théorie des Nombres de Bordeaux » (<http://jtnb.cedram.org/>), implique l'accord avec les conditions générales d'utilisation (<http://jtnb.cedram.org/legal/>). Toute reproduction en tout ou partie cet article sous quelque forme que ce soit pour tout usage autre que l'utilisation à fin strictement personnelle du copiste est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

cedram

*Article mis en ligne dans le cadre du*  
*Centre de diffusion des revues académiques de mathématiques*  
<http://www.cedram.org/>

# Propriétés et limites de la reconnaissance d'ensembles d'entiers par automates dénombrables

par JULIEN CASSAIGNE et MARION LE GONIDEC

RÉSUMÉ. Nous étudions dans cet article deux familles d'ensembles d'entiers reconnaissables par des automates finis ou dénombrables. Les résultats concernant ces deux notions de reconnaissabilité qui sont présentés ici étendent de manière naturelle les résultats structurels usuels de la famille des ensembles  $k$ -reconnaissables. Le cas particulier de l'ensemble des nombres premiers est également abordé.

ABSTRACT. *Properties and limits of recognition of sets of integers by countable automata* We study two families of sets of integers recognized by countable automata. Presented results on these two recognition notions naturally extend usual structural results about the family of  $k$ -recognizable sets. The particular case of the set of prime numbers is also detailed.

## 1. Introduction

L'étude de familles et de suites d'entiers engendrées par des algorithmes simples peut être liée à divers problèmes en mathématiques et en informatique théorique ; c'est le cas par exemple des ensembles d'entiers  $k$ -reconnaissables, étudiés en particulier par A. Cobham [9], qui interviennent entre autres en dynamique symbolique ou en analyse diophantienne par le biais des mots automatiques. Nous renvoyons le lecteur au livre de J.-P. Allouche et J. Shallit [2] pour un panorama sur les mots automatiques.

Devant la richesse des résultats concernant la famille des ensembles d'entiers reconnaissables, il est assez naturel de se poser la question de l'extension de ces résultats à d'autres familles d'ensembles d'entiers, non plus associées à des langages rationnels et à des automates finis mais à d'autres types de langages plus généraux ou à d'autres types de machines. Dans cette optique, quelques papiers concernent les familles d'entiers reconnaissables par les automates à pile. Ils font suite au problème soulevé dans l'article de Minsky et Papert [21] de trouver, pour des ensembles d'entiers, des critères

de reconnaissabilité ou de non-reconnaissabilité par automate à pile. Citons les articles de Hartmanis et Shank [17] et M.-P. Schützenberger [26] ainsi que les travaux de J. Berstel [7, 8] sur ce sujet. D'autres voies de généralisation, utilisant les nombreuses caractérisations des suites automatiques, comme les suites  $k$ -régulières introduites par J.-P. Allouche et J. Shallit [3, 4] ou les suites  $S$ -reconnaissables introduites par M. Rigo [23] et basées sur des langages rationnels, ont également ouvert des champs de recherche fructueux pour l'étude d'ensembles d'entiers engendrés par des algorithmes simples.

Cet article se situe également dans la problématique d'extension de résultats sur les ensembles  $k$ -reconnaissables. Le cadre choisi ici est celui des ensembles d'entiers reconnaissables par une large famille d'automates finis ou dénombrables ayant peu de contraintes structurelles. Nous considérons des automates dénombrables de degré entrant borné ayant un ensemble d'acceptation fini. Ce cadre peut être aussi motivé par les travaux effectués sur les mots  $k^\infty$ -automatiques (mots engendrés par des automates dénombrables) dans [18, 20] puisque les ensembles d'entiers supports des lettres de ces mots font partie des ensembles étudiés ici. En particulier, nous verrons que contrairement au cas des ensembles reconnaissables, il est nécessaire d'introduire deux notions de reconnaissabilité par automates dénombrables, selon le sens de lecture des représentations choisi ; la notion de reconnaissabilité dans le cadre dénombrable introduite par C. Mauduit [20] correspondant au cas de la lecture de gauche à droite.

Nous montrerons également que les familles d'ensembles d'entiers définies par ces deux notions ont des propriétés structurelles différentes. Cela permet de mettre en évidence certains phénomènes strictement spécifiques au cas des automates finis et certains autres qui s'étendent au cas dénombrable dans un sens de lecture ou dans les deux.

## 2. Ensembles d'entiers $k^\infty$ -reconnaissables DG et GD

L'ensemble  $\llbracket a, b \rrbracket$  désigne l'ensemble  $[a, b] \cap \mathbb{N}$ .

Pour tout entier  $k \geq 2$ , on note  $\rho_k(n) = n_l \cdots n_1 n_0$  la représentation normalisée en base  $k$  de l'entier  $n = \sum_{i=0}^l n_i k^i$  avec  $n_l \neq 0$  et  $n_i \in \llbracket 0, k-1 \rrbracket$ . Par convention, on note  $\rho_k(0) = \varepsilon$ , où  $\varepsilon$  désigne le mot vide. Pour tout ensemble d'entiers  $E$ , on note aussi  $\rho_k(E)$  le langage des représentations normalisées en base  $k$  des éléments de  $E$ .

$$\rho_k(E) = \{\rho_k(n) \mid n \in E\}.$$

Pour tout mot fini  $w = w_0 \cdots w_l$  sur l'alphabet  $\llbracket 0, k-1 \rrbracket$ , on note  $\pi_k(w)$  l'entier défini par  $\pi_k(w) = \sum_{i=0}^{|w|-1} w_i k^{|w|-1-i}$  de sorte que pour tout  $n \in \mathbb{N}$ ,  $\pi_k(\rho_k(n)) = n$ .

## 2.1. Cadre.

**Définition.** On appelle *k-automate dénombrable* la donnée d'un quadruplet  $\mathcal{A} = (Q, \phi, q_0, F)$  où :

- $Q$  est un ensemble dénombrable, appelé *ensemble des états*,
- $\phi : \llbracket 0, k-1 \rrbracket \times Q \rightarrow Q$  est appelée *fonction de transition*,
- $q_0 \in Q$  est appelé *état initial*,
- $F$  est un sous-ensemble fini de  $Q$  appelé *ensemble d'acceptation*.

On supposera ici que tous les automates considérés sont de *degré fini*, c'est-à-dire de degré entrant fini : pour tout état  $q \in Q$ ,  $\text{Card}(\phi^{-1}\{q\})$  est fini, puisque le degré sortant est quant à lui toujours borné par  $k$ .

**Remarque.** La fonction de transition d'un *k-automate dénombrable*  $\mathcal{A} = (Q, \phi, q_0, F)$  est *a priori* définie sur tout l'ensemble  $\llbracket 0, k-1 \rrbracket \times Q$ . Cependant, la fonction  $\phi$  pourra être définie partiellement, c'est-à-dire seulement sur une partie  $P$  de  $\llbracket 0, k-1 \rrbracket \times Q$  car on peut toujours compléter l'automate tout en conservant la finitude de son degré en lui adjoignant, au pire, une quantité dénombrable d'*états poubelle* vers lesquels pointent les transitions non définies.

Le *k-automate complété*  $\mathcal{A}' = (Q \cup X, \psi, q_0, F)$  où  $X = \{x_q \mid q \in Q\}$  et la fonction de transition  $\psi$  est définie sur  $\llbracket 0, k-1 \rrbracket \times (Q \cup X)$  par :

- $\psi(i, q) = \phi(i, q)$  si  $(i, q) \in P$ ,
- $\psi(i, q) = x_q$  si  $(i, q) \in \llbracket 0, k-1 \rrbracket \times Q \setminus P$ ,
- $\psi(i, x) = x$  pour tout  $(i, x) \in \llbracket 0, k-1 \rrbracket \times X$ ,

reconnaît alors le même langage que l'automate  $\mathcal{A}$ .

La fonction de transition  $\phi$  peut être étendue à  $\llbracket 0, k-1 \rrbracket^* \times Q$  de deux manières différentes :

- en posant, pour tout mot  $w = w_0 w_1 \cdots w_n$  de  $\llbracket 0, k-1 \rrbracket^*$  et tout  $q \in Q$ ,

$$\overrightarrow{\phi}(w, q) = \phi_{w_n} \circ \cdots \circ \phi_{w_1} \circ \phi_{w_0}(q),$$

ce qui correspond à une lecture de gauche à droite du mot  $w$  et que l'on notera GD.

- ou en posant, pour tout mot  $w = w_0 w_1 \cdots w_n$  de  $\llbracket 0, k-1 \rrbracket^*$  et tout  $q \in Q$ ,

$$\overleftarrow{\phi}(w, q) = \phi_{w_0} \circ \phi_{w_1} \circ \cdots \circ \phi_{w_n}(q),$$

ce qui correspond à une lecture de droite à gauche du mot  $w$  et que l'on notera DG.

**Notation.** Pour un *k-automate dénombrable*  $\mathcal{A} = (Q, \phi, q_0, F)$  fixé, on note  $\overrightarrow{L}(\mathcal{A})$  le langage reconnu par  $\mathcal{A}$  lorsque  $\phi$  est étendue à  $\llbracket 0, k-1 \rrbracket^* \times Q$  par  $\overrightarrow{\phi}$ , c'est-à-dire l'ensemble des mots  $w$  de  $\llbracket 0, k-1 \rrbracket^*$  tels que  $\overrightarrow{\phi}(w, q_0)$  appartient à l'ensemble d'acceptation  $F$  :

$$\vec{L}(\mathcal{A}) = \left\{ w \in \llbracket 0, k-1 \rrbracket^* \mid \vec{\phi}(w, q_0) \in F \right\}.$$

On note également  $\overleftarrow{L}(\mathcal{A})$  le langage reconnu par  $\mathcal{A}$  lorsque  $\phi$  est étendue à  $\llbracket 0, k-1 \rrbracket^* \times Q$  par  $\overleftarrow{\phi}$ , c'est-à-dire l'ensemble des mots  $w$  de  $\llbracket 0, k-1 \rrbracket^*$  tels que  $\overleftarrow{\phi}(w, q_0)$  appartient à l'ensemble d'acceptation  $F$  :

$$\overleftarrow{L}(\mathcal{A}) = \left\{ w \in \llbracket 0, k-1 \rrbracket^* \mid \overleftarrow{\phi}(w, q_0) \in F \right\}.$$

**Définition.** Un ensemble  $E$  d'entiers est dit  $k^\infty$ -reconnaissable GD (resp. DG) s'il existe un  $k$ -automate fini ou dénombrable  $\mathcal{A}_E$ , tel que

$$\vec{L}(\mathcal{A}_E) = \rho_k(E) \quad \left( \text{resp. } \overleftarrow{L}(\mathcal{A}_E) = \rho_k(E) \right).$$

Lorsque l'ensemble  $E$  est  $k^\infty$ -reconnaissable à la fois GD et DG, on dira simplement que  $E$  est  $k^\infty$ -reconnaissable.

Un ensemble d'entiers dont le complémentaire dans  $\mathbb{N}$  est  $k^\infty$ -reconnaissable GD (resp. DG) est dit  $k^\infty$ -co-reconnaissable GD (resp. DG).

**Remarque.** La précision du sens de lecture est indispensable car les notions de  $k^\infty$ -reconnaissabilité GD et DG ne sont pas équivalentes (voir Paragraphe 4). Le sens de lecture GD correspond à une lecture des représentations en base  $k$  du chiffre le plus significatif vers le moins significatif et le sens de lecture DG correspond à une lecture des représentations en base  $k$  du chiffre le moins significatif vers le plus significatif.

**Remarque.** Dans la définition des ensembles  $k^\infty$ -reconnaissables GD, on aurait pu choisir de considérer que  $\vec{L}(\mathcal{A}_E) = 0^* \rho_k(E)$  en rajoutant une boucle indexée par 0 sur l'état initial.

En revanche, une telle transformation de la définition n'est pas toujours possible en conservant la finitude de l'ensemble final dans le cas des ensembles  $k^\infty$ -reconnaissables DG. C'est d'ailleurs pour cette raison que nous avons choisi de définir les ensembles  $k^\infty$ -co-reconnaissables comme les complémentaires des ensembles  $k^\infty$ -reconnaissables et non comme les ensembles d'entiers reconnus par des  $k$ -automates finis ou dénombrables dont l'ensemble des états d'acceptation est co-fini. Cela permet notamment d'éviter les problèmes posés par les écritures impropres des entiers dans le cas de la lecture DG des entiers. En effet, si un langage du type  $0^* \rho_k(E)$  est reconnu par un automate dénombrable dont l'ensemble d'acceptation est co-fini, il se peut que le langage  $\rho_k(E)$  ne le soit pas : la transformation de l'automate nécessaire ne préservant pas forcément à la fois la co-finitude de l'ensemble d'acceptation et la finitude du degré.

**Remarque.** Les automates considérés ici ont un ensemble d'états d'acceptation fini. Cette restriction est motivée par le fait que si l'on n'impose

aucune restriction sur la forme de cet ensemble, alors tout ensemble d'entiers  $E$  est  $k^\infty$ -reconnaisable (GD et DG) pour tout entier  $k \geq 2$ . En effet,  $\mathbb{N}$  est un ensemble dénombrable donc, quels que soient le sens de lecture et la base  $k$  choisis, on peut construire un automate dénombrable de sorte que la lecture de chaque représentation propre d'entier aboutisse sur un état différent. Sans imposer de condition sur l'ensemble des états d'acceptation, tout sous-ensemble  $E$  de  $\mathbb{N}$  peut alors être reconnu par cet automate, en choisissant comme ensemble d'acceptation les états sur lesquels aboutissent les lectures des éléments de  $\rho_k(E)$ .

**2.2. Quelques exemples pour fixer les choses.** Un premier exemple simple d'ensemble  $2^\infty$ -reconnaisable est constitué par l'ensemble des entiers non nuls dont l'écriture en base 2 contient autant de zéros que de uns :  $E_1 = \{n \in \mathbb{N}^* \mid |\rho_2(n)|_0 = |\rho_2(n)|_1\}$ . Dans ce cas, un 2-automate GD reconnaissant  $E_1$ , présenté Figure 2.1 ; son état initial est  $q_0$  et son unique état final est  $f$ . Les propriétés combinatoires du mot indicateur de cet ensemble d'entiers ont été étudiées dans [19]. Un 2-automate DG reconnaissant  $E_1$  sera explicité au paragraphe 5.1 (Figure 5.2).

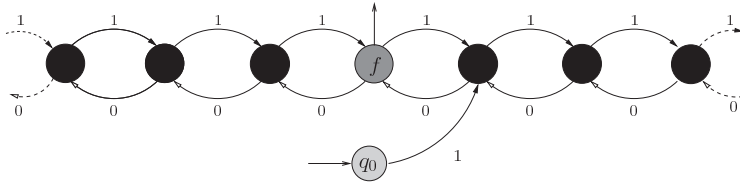


FIGURE 2.1. Graphe d'un 2-automate GD reconnaissant l'ensemble  $E_1$ .

L'ensemble d'entiers  $E_2 = \{n \in \mathbb{N} \mid \rho_3(n) = u0\tilde{u}, u \in \{1, 2\}^+\}$  est, pour sa part, un ensemble  $3^\infty$ -reconnaisable. Les 3-automates GD et DG reconnaissant  $E_2$  sont les mêmes et cet automate est présenté Figure 2.2.

**Remarque.** Ces exemples sont un peu artificiels car les vrais objets sous-jacents ne sont pas les ensembles d'entiers mais les langages des représentations en bases 2 et 3. D'autre part, ces deux exemples entrent dans le cadre d'une proposition plus générale, présentée ci-dessous. En effet, une famille de langages algébriques est reconnue par ce type d'automates. Cela tient à la correspondance entre certains  $k$ -automates dénombrables de degré borné et les automates à pile déterministes temps réel, via leurs graphes des transitions.

Avant de formuler clairement ce résultat (Proposition 2.1), rappelons qu'un  $k$ -automate à pile déterministe temps réel est la donnée d'un quintuplet  $\mathcal{M} = (Q, \Sigma, q_0w^{(0)}, K, T)$  sur lequel agit  $\llbracket 0, k - 1 \rrbracket^*$ , où :

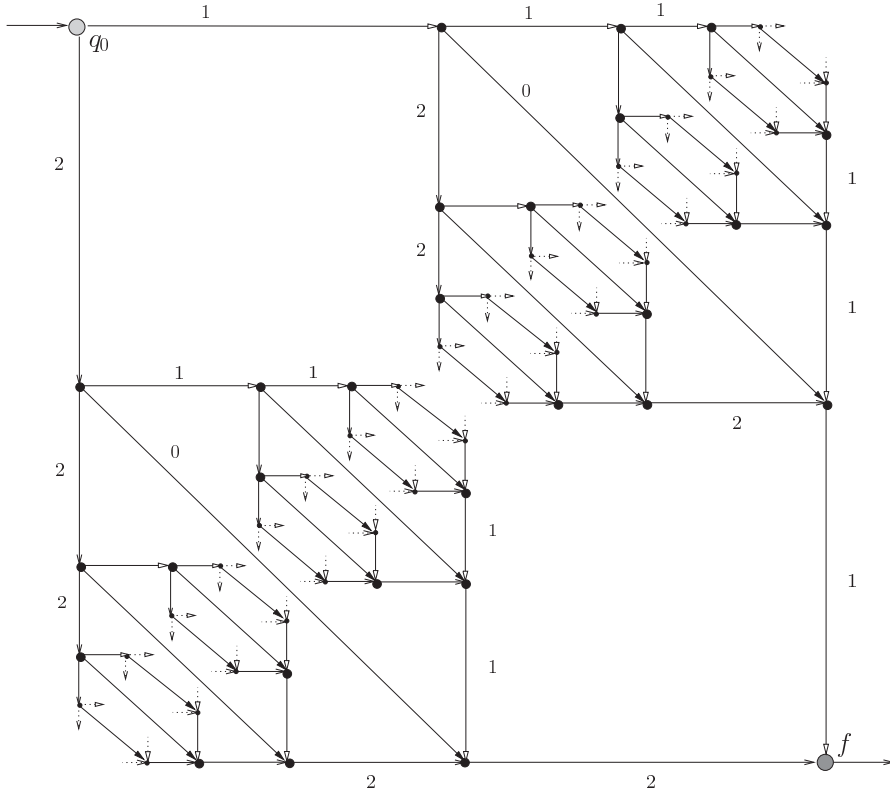


FIGURE 2.2. Graphe du 3-automate reconnaissant l'ensemble  $E_2$ .

- $Q$  est un ensemble fini, appelé *ensemble des états*,
- $\Sigma$  est un ensemble fini, appelé *ensemble des symboles de pile*,
- $q_0 w^{(0)} \in Q \times \Sigma^*$  est la *configuration initiale*,
- $K \subset \llbracket 0, k-1 \rrbracket \times Q \times (\Sigma \cup \{\varepsilon\}) \times \Sigma^* \times Q$  est l'*ensemble des règles de transition*, vérifiant l'assertion suivante : pour tout triplet  $(i, q, \sigma)$  de  $\llbracket 0, k-1 \rrbracket \times Q \times (\Sigma \cup \{\varepsilon\})$ , il existe au plus un couple  $(w, q')$  de  $\Sigma^* \times Q$  tel que  $(i, q, \sigma, w, q') \in K$ .
- $T \subset Q \times \Sigma^*$  est l'*ensemble des configurations acceptées*.

Un mot  $v = v_0 v_1 \cdots v_{n-1}$  sur l'alphabet  $\llbracket 0, k-1 \rrbracket$  est reconnu par l'automate à pile  $\mathcal{M}$  par lecture GD s'il existe une suite  $(q_0, w^{(0)}), (q_1, w^{(1)}), \dots, (q_n, w^{(n)})$  d'éléments de  $Q \times \Sigma^*$  où :

- $(q_0, w^{(0)})$  est la configuration initiale de l'automate,
- $(q_n, w^{(n)}) \in T$ ,
- pour tout  $i$  de  $\llbracket 0, n-1 \rrbracket$ , il existe deux mots  $w$  et  $w'$  de  $\Sigma^*$  tels que :

- si  $w^{(i)} \neq \varepsilon$ ,  $w^{(i)} = w_0^{(i)}w$ ,  $w^{(i+1)} = w'w$  et  $(v_i, q_i, w_0^{(i)}, w', q_{i+1}) \in K$ ,
- si  $w^{(i)} = \varepsilon$ ,  $w = \varepsilon$ ,  $w^{(i+1)} = w'$  et  $(v_i, q_i, \varepsilon, w', q_{i+1}) \in K$ .

De la même manière, un mot  $v = v_0v_1 \cdots v_{n-1}$  sur l'alphabet  $\llbracket 0, k-1 \rrbracket$  est reconnu par l'automate à pile  $\mathcal{M}$  par lecture DG s'il existe une suite  $(q_0, w^{(0)}), (q_1, w^{(1)}), \dots, (q_n, w^{(n)})$  d'éléments de  $Q \times \Sigma^*$  avec  $(q_n, w^{(n)}) \in T$  où :

- $(q_0, w^{(0)})$  est la configuration initiale de l'automate,
- $(q_n, w^{(n)}) \in T$ ,
- pour tout  $i$  de  $\llbracket 0, n-1 \rrbracket$ , il existe deux mots  $w$  et  $w'$  de  $\Sigma^*$  tels que :
  - si  $w^{(i)} \neq \varepsilon$ ,  $w^{(i)} = w_0^{(i)}w$ ,  $w^{(i+1)} = w'w$  et  $(v_{n-1-i}, q_i, w_0^{(i)}, w', q_{i+1}) \in K$ ,
  - si  $w^{(i)} = \varepsilon$ ,  $w = \varepsilon$ ,  $w^{(i+1)} = w'$  et  $(v_{n-1-i}, q_i, \varepsilon, w', q_{i+1}) \in K$ .

Par extension, on définit le langage  $\overrightarrow{L}(\mathcal{M})$  comme l'ensemble des mots de  $\llbracket 0, k-1 \rrbracket^*$  reconnus par  $\mathcal{M}$  par lecture GD et le langage  $\overleftarrow{L}(\mathcal{M})$  comme l'ensemble des mots de  $\llbracket 0, k-1 \rrbracket^*$  reconnus par  $\mathcal{M}$  par lecture DG.

Si l'ensemble des configurations acceptées par l'automate est de la forme  $T \subset Q \times \{\varepsilon\}$ , on dit que l'automate reconnaît par vidage de pile et états d'acceptation. Il existe d'autres modes d'acceptation par les automates à pile, selon la forme de  $T$ , et les automates à pile présentés ici sont des automates bien spécifiques. Nous renvoyons à [5] ou [6] pour une définition plus générale des automates à pile ainsi que pour une étude approfondie des liens entre langages algébriques et automates à pile.

**Proposition 2.1.** *Soit un entier  $k \geq 2$  et soit  $E$  un ensemble d'entiers tel que  $\rho_k(E)$  est un langage algébrique.*

*Si  $\rho_k(E)$  est reconnu par vidage de pile et états d'acceptation par un  $k$ -automate à pile déterministe temps réel par lecture GD (resp. DG), alors  $E$  est  $k^\infty$ -reconnaisable GD (resp. DG).*

*Démonstration.* Soit  $E$  un ensemble d'entiers reconnu par vidage de pile et états d'acceptation par le  $k$ -automate à pile déterministe et temps réel  $\mathcal{M} = (Q, \Sigma, q_0w^{(0)}, K, T)$ .

Un  $k$ -automate dénombrable  $\mathcal{A} = (Q', \phi, q'_0, F)$  qui reconnaît  $\rho_k(E)$  est l'automate porté par le graphe des transitions de l'automate à pile  $\mathcal{P}$  (voir par exemple [11, 12]), c'est-à-dire le graphe défini par :

- $Q' = Q \times \Sigma^*$ ,
- pour tout  $(i, q, q', \sigma, v, w) \in \llbracket 0, k-1 \rrbracket \times Q \times Q \times \Sigma \times \Sigma^* \times \Sigma^*$ ,

$$\phi(i, q\sigma w) = q'vw \iff (i, q, \sigma, v, q') \in K,$$

$$\phi(i, q\varepsilon) = q'v \iff (i, q, \varepsilon, v, q') \in K.$$

- $q'_0 = q_0w^{(0)}$ ,
- $F = T$ .



Ce  $k$ -automate est en particulier dénombrable et de degré borné (voir [22] ou [11, 12]) et il reconnaît exactement  $\rho_k(E)$ . D'autre part, puisque  $F$  est fini,  $E$  est bien  $k^\infty$ -reconnaisable.  $\square$

**Remarque.** Un autre résultat concernant les numérations en base non entière est à rapprocher de ce résultat. Dans [16] C. Frougny et B. Solomyak ont montré que, sous l'hypothèse que les expansions en base non entière  $\theta > 1$  de tous les entiers sont finies,  $\rho_\theta(\mathbb{N})$  est un langage algébrique si et seulement si  $\theta$  est un nombre de Pisot quadratique unitaire.

**Remarque.** Il existe des ensembles  $k^\infty$ -reconnaisables dont les langages des représentations en base  $k$  ne sont pas algébriques. C'est le cas, par exemple, de l'ensemble  $E_3 = \{n \in \mathbb{N} \mid \rho_3(n) = 1^i 0^j 2^i, i \in \mathbb{N}\}$  qui est  $3^\infty$ -reconnaisable GD et DG. Un 3-automate dénombrable GD reconnaissant  $E_3$  est représenté Figure 2.3 et, par analogie, on peut construire facilement un 3-automate dénombrable DG qui reconnaît  $E_3$ .

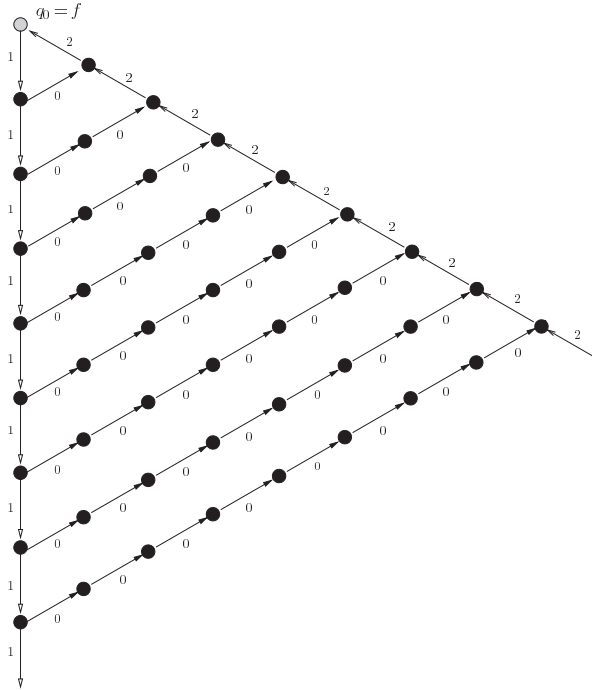


FIGURE 2.3. Graphe du 3-automate GD reconnaissant l'ensemble  $E_3$ .

### 3. Un lemme de la pompe pour les automates dénombrables

Les deux lemmes présentés ci-dessous étendent le lemme de la pompe usuel que l'on peut appliquer aux automates finis.

**Définition.** Soit  $N$  un entier positif et  $k$  un entier supérieur ou égal à 2.

On appelle *famille préfixe en base  $k$*  un ensemble fini et ordonné d'entiers distincts  $\{n_i \mid i \in \llbracket 0, N \rrbracket\}$  vérifiant, pour tout entier  $i$  de  $\llbracket 0, N - 1 \rrbracket$ ,  $\rho_k(n_i)$  est un préfixe de  $\rho_k(n_{i+1})$ , c'est à dire que, pour tout  $i$  de  $\llbracket 0, N - 1 \rrbracket$ ,  $n_i < n_{i+1}$  et

$$0 \leq n_{i+1} - n_i k^{|\rho_k(n_{i+1})| - |\rho_k(n_i)|} < k^{|\rho_k(n_{i+1})| - |\rho_k(n_i)|}.$$

On appelle *famille suffixe en base  $k$*  un ensemble fini et ordonné d'entiers distincts  $\{n_i, i \in \llbracket 0, N \rrbracket\}$  vérifiant, pour tout entier  $i$  de  $\llbracket 0, N - 1 \rrbracket$ ,  $\rho_k(n_i)$  est un suffixe de  $\rho_k(n_{i+1})$ , c'est à dire que pour tout  $i$  de  $\llbracket 0, N - 1 \rrbracket$ ,  $n_i < n_{i+1}$  et

$$n_{i+1} = n_i \pmod{k^{|\rho_k(n_i)|}}.$$

Par extension, on parlera également de *suites préfixes* (resp. *suffixes*) en base  $k$  pour des suites d'entiers  $(n_i)_{i \geq 0}$  pour lesquelles, quel que soit l'entier  $N > 0$ , la famille  $\{n_i, i \in \llbracket 0, N \rrbracket\}$  est une famille préfixe (resp. suffixe) en base  $k$ .

**Lemme 3.1.** Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaisable GD, reconnu par le  $k$ -automate  $\mathcal{A} = (Q, \phi, q_0, F)$ .

Si un sous-ensemble  $E_{\text{Pref}} = \{n_i \mid i \in \llbracket 0, |F| \rrbracket\}$  d'éléments de  $E$  forme une famille préfixe en base  $k$ , alors il existe deux entiers  $i_0 < i_1$  de  $\llbracket 0, |F| \rrbracket$  tels que la suite  $(a_n)_{n \geq 0}$  définie par :

$$a_n = n_{i_0} k^{nl} + (n_{i_1} - n_{i_0} k^l) \frac{k^{nl} - 1}{k^l - 1},$$

où  $l = |\rho_k(n_{i_1})| - |\rho_k(n_{i_0})|$ , est une suite d'éléments de  $E$ .

Autrement dit, il existe deux mots  $u, v \in \llbracket 0, k - 1 \rrbracket^*$ , avec  $v \neq \varepsilon$ , définis par  $u = \rho_k(n_{i_0})$  et  $uv = \rho_k(n_{i_1})$ , tels que pour tout entier  $n \geq 0$ , l'entier  $a_n = \pi_k(uv^n)$  appartient à  $E$ .

*Démonstration.* Sous les hypothèses du lemme, on introduit la famille d'états de  $F$  suivante :

$$G = \left\{ \overrightarrow{\phi}(\rho_k(n_i), q_0) \mid n_i \in E_{\text{Pref}} \right\}.$$

L'ensemble  $G$  étant inclus dans  $F$ , il existe un état  $f \in F$  et deux entiers  $n_{i_0} < n_{i_1}$  de  $E_{\text{Pref}}$  tels que :

$$\overrightarrow{\phi}(\rho_k(n_{i_0}), q_0) = \overrightarrow{\phi}(\rho_k(n_{i_1}), q_0) = f.$$

On pose  $u = \rho_k(n_{i_0})$ . La famille  $E_{\text{Pref}}$  étant une famille préfixe de  $E$ , il existe un mot  $v$  sur  $\llbracket 0, k - 1 \rrbracket$  avec  $v \neq \varepsilon$ , tel que  $\rho_k(n_{i_1}) = uv$ . On peut déduire

de ce qui précède que  $\overrightarrow{\phi}(v, f) = f$  puisque l'automate  $\mathcal{A}$  est déterministe. Ainsi, pour tout entier  $n$  positif,  $\overrightarrow{\phi}(v^n, f) = f$ .

Soit  $(a_n)_{n \geq 0}$  la suite définie par  $a_n = \pi_k(uv^n)$ . Ces entiers sont reconnus par  $\mathcal{A}$  puisque  $\overrightarrow{\phi}(uv^n, q_0) = \overrightarrow{\phi}(u, q_0) = f$ .

On a  $a_0 = n_{i_0}$  et  $a_1 = n_{i_1}$ . De plus, on a les deux égalités suivantes :

$$\pi_k(u) = n_{i_0} \quad \text{et} \quad \pi_k(v) = n_{i_1} - n_{i_0} k^{|\rho_k(n_{i_1})| - |\rho_k(n_{i_0})|}.$$

Par suite, il vient, pour tout  $n \geq 0$  :

$$a_n = \pi_k(uv^n) = \pi_k(u0^{n|v|}) + \pi_k(v^n) = \pi_k(u)k^{n|v|} + \pi_k(v) \sum_{i=0}^{n-1} k^{i|v|},$$

d'où :

$$a_n = n_{i_0} k^{n|v|} + (n_{i_1} - n_{i_0} k^{|v|}) \frac{k^{n|v|} - 1}{k^{|v|} - 1},$$

ce qui conclut la démonstration du lemme.  $\square$

Le Lemme 3.1 a bien sûr son équivalent pour les ensembles  $k^\infty$ -reconnaisables DG, qui utilise l'existence de familles suffixes.

**Lemme 3.2.** *Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaisable DG, reconnu par le  $k$ -automate  $\mathcal{A} = (Q, \phi, q_0, F)$ .*

*Si un sous-ensemble  $E_{\text{Suff}} = \{n_i, i \in \llbracket 0, |F| \rrbracket\}$  d'éléments de  $E$  forme une famille suffixe en base  $k$ , alors il existe deux entiers  $i_0 < i_1$  de  $\llbracket 0, |F| \rrbracket$  tels que la suite  $(a_n)_{n \geq 0}$  définie par :*

$$a_n = \frac{k^{nl} - 1}{k^l - 1} (n_{i_1} - n_{i_0}) + n_{i_0},$$

où  $l = |\rho_k(n_{i_1})| - |\rho_k(n_{i_0})|$ , est une suite d'éléments de  $E$ .

*Autrement dit, il existe deux mots  $v, w \in \llbracket 0, k-1 \rrbracket^*$ , avec  $v \neq \varepsilon$ , définis par  $w = \rho_k(n_0)$  et  $vw = \rho_k(n_1)$ , tels que pour tout entier  $n \geq 0$ , l'entier défini par  $a_n = \pi_k(v^n w)$  appartient à  $E$ .*

*Démonstration.* Sous les hypothèses du lemme, on introduit la famille d'états de  $F$  suivante :

$$G = \left\{ \overleftarrow{\phi}(\rho_k(n_i), q_0) \mid n_i \in E_{\text{Suff}} \right\}.$$

L'ensemble  $G$  étant inclus dans  $F$ , il existe un état  $f \in F$  et deux entiers  $n_{i_0} < n_{i_1}$  de  $E_{\text{Suff}}$  tels que :

$$\overleftarrow{\phi}(\rho_k(n_{i_0}), q_0) = \overleftarrow{\phi}(\rho_k(n_{i_1}), q_0) = f.$$

On pose  $w = \rho_k(n_{i_0})$ . La famille  $E_{\text{Suff}}$  étant une famille suffixe d'éléments de  $E$ , il existe un mot  $v$  sur  $\llbracket 0, k-1 \rrbracket$ , avec  $v \neq \varepsilon$ , tel qu'on peut écrire  $\rho_k(n_{i_1}) = vw$ . On peut déduire de ce qui précède que  $\overleftarrow{\phi}(v, f) = f$  puisque

l'automate  $\mathcal{A}$  est déterministe. Ainsi, pour tout entier  $n$  positif,  $\overleftarrow{\phi}(v^n, f) = f$ .

Soit  $(a_n)_{n \geq 0}$  la suite définie par  $a_n = \pi_k(v^n w)$ . Cette famille d'entiers est reconnue par  $\mathcal{A}$  puisque  $\overleftarrow{\phi}(v^n w, q_0) = \overleftarrow{\phi}(w, q_0) = f$ .

On a  $a_0 = n_{i_0}$ ,  $a_1 = n_{i_1}$ . De plus, on a les deux égalités suivantes :

$$\pi_k(w) = n_{i_0} \quad \text{et} \quad \pi_k(v0^{|\rho_k(n_{i_0})|}) = n_{i_1} - n_{i_0}.$$

Par suite, il vient, pour tout  $n \geq 0$  :

$$a_n = \pi_k(v^n w) = \pi_k(v^n 0^{|w|}) + \pi_k(w) = \pi_k(v0^{|w|}) \sum_{i=0}^{n-1} k^{i|v|} + \pi_k(w),$$

d'où :

$$a_n = (n_{i_1} - n_{i_0}) \frac{k^{n|v|} - 1}{k^{|v|} - 1} + n_{i_0},$$

ce qui conclut la démonstration du lemme. □

Ces deux lemmes de pompage permettent en particulier de montrer le caractère non-reconnaisable de certains ensembles d'entiers en les mettant en défaut ; c'est le cas par exemple pour l'ensemble des nombres premiers.

### 3.1. Application des lemmes de la pompe à l'ensemble des nombres premiers.

**Théorème 3.1.** *Quel que soit l'entier  $k \geq 2$ , l'ensemble  $\mathbb{P}$  des nombres premiers n'est ni  $k^\infty$ -reconnaisable GD ni  $k^\infty$ -reconnaisable DG.*

Ce théorème répond au problème ouvert [20, Problème 1, Question 2]. Sa démonstration est scindée en deux parties : dans un premier temps, on va montrer que l'ensemble des nombres premiers n'est pas  $k^\infty$ -reconnaisable GD ; viendra ensuite la deuxième partie où on montrera que l'ensemble des nombres premiers n'est pas  $k^\infty$ -reconnaisable DG.

Dans la suite du paragraphe,  $k$  désigne un entier supérieur ou égal à 2 et  $\mathbb{P}$  désigne l'ensemble des nombres premiers.

**Lemme 3.3.** *Il existe une suite préfixe en base  $k$  de nombres premiers.*

*Démonstration.* On cherche à montrer qu'on peut extraire de l'ensemble  $\mathbb{P}$  des nombres premiers une suite  $\mathbf{p} = (p_n)_{n \geq 0}$  telle que :

$$0 < p_{n+1} - p_n k^{[\log_k p_{n+1}] - [\log_k p_n]} < k^{[\log_k p_{n+1}] - [\log_k p_n]},$$

c'est-à-dire telle que :

$$\frac{p_n}{k^{[\log_k(p_n)]}} < \frac{p_{n+1}}{k^{[\log_k(p_{n+1})]}} < \frac{p_n}{k^{[\log_k(p_n)]}} + \frac{1}{k^{[\log_k(p_n)]}},$$

condition que l'on peut écrire ainsi :

$$k^{\{\log_k(p_n)\}} < k^{\{\log_k(p_{n+1})\}} < k^{\{\log_k(p_n)\}} + \frac{1}{k^{[\log_k(p_n)]}},$$

où  $\{x\}$  désigne la partie fractionnaire de  $x$ .

L'existence d'une telle suite est assurée par la densité de la famille  $\{\{\log_k p\}, p \in \mathbb{P}\}$  dans  $[0, 1]$ . En effet, si  $\{\{\log_k p\} \mid p \in \mathbb{P}\}$  est dense dans  $[0, 1]$ , alors la famille  $\{k^{\{\log_k p\}} \mid p \in \mathbb{P}\}$  l'est dans  $[1, k]$  et donc on construit la suite  $\mathbf{p}$  en fixant  $p_0 \in \mathbb{P}$  et pour tout  $n \in \mathbb{N}$ , on choisit  $p_{n+1}$  de sorte que  $k^{\{\log_k(p_{n+1})\}}$  appartienne à l'intervalle  $]k^{\{\log_k(p_n)\}}, k^{\{\log_k(p_n)\}} + \frac{1}{k^{\lfloor \log_k(p_n) \rfloor}}[$ .

Il reste donc à prouver que la famille  $\{\{\log_k p\} \mid p \in \mathbb{P}\}$  est effectivement dense dans  $[0, 1]$ . Pour cela, supposons par l'absurde qu'il existe un intervalle non vide  $]a, b[ \subset [0, 1]$  ne contenant aucun élément de  $\{\{\log_k p\} \mid p \in \mathbb{P}\}$ . Pour tout  $n \geq 0$ , on a donc  $\mathbb{P} \cap ]k^{n+a}, k^{n+b}] = \emptyset$  et cela implique

$$\pi(k^{n+a}) = \#(\mathbb{P} \cap [0, k^{n+a})) = \#(\mathbb{P} \cap [0, k^{n+b})) = \pi(k^{n+b}).$$

Ainsi, pour tout  $n \geq 1$ ,  $\frac{\pi(k^{n+b})}{\pi(k^{n+a})} = 1$ . Cependant, le théorème des nombres premiers assure que  $\pi(k^{n+a})$  est équivalent à  $\frac{k^{n+a}}{(n+a) \ln(k)}$  lorsque  $n$  tend vers l'infini, et donc  $\frac{\pi(k^{n+b})}{\pi(k^{n+a})}$  est équivalent à  $k^{b-a}$  lorsque  $n$  tend vers l'infini. Ces deux résultats sont incompatibles puisque on a supposé que  $b - a \neq 0$ . Ainsi, tout intervalle non vide de  $[0, 1]$  contient au moins un élément de  $\{\{\log_k p\}, p \in \mathbb{P}\}$ .  $\square$

*Démonstration. Première partie du Théorème 3.1.* Procédons par l'absurde en supposant que  $\mathbb{P}$  soit reconnaissable par un automate fini ou dénombrable  $\mathcal{A} = (Q, \phi, q_0, F)$  avec ensemble des états terminaux  $F$  fini. Nous allons montrer que cela est exclu en mettant en défaut le lemme de pompage 3.1. Le Lemme 3.3 assure l'existence d'une suite préfixe  $\mathbf{p} = (p_n)_{n \geq 0}$  de nombres premiers et donc d'une famille préfixe à  $|F| + 1$  éléments de nombres premiers quelle que soit la valeur de  $|F|$ . On peut donc appliquer le lemme de pompage 3.1 à la famille  $\{p_n \mid n \in \llbracket 0, |F| \rrbracket\}$ .

Il existe deux nombres premiers  $p_{i_0} < p_{i_1}$  issus de cette famille tels que la suite  $\mathbf{a} = (a_n)_{n \geq 0}$ , définie pour tout  $n \geq 0$  par :

$$a_n = p_{i_0} k^{nl} + \frac{k^{nl} - 1}{k^l - 1} (p_{i_1} - p_{i_0} k^l),$$

soit une suite de nombres premiers, avec  $l = \lfloor \log_k p_{i_1} \rfloor - \lfloor \log_k p_{i_0} \rfloor > 0$ . On remarquera en particulier que  $p_{i_1} - p_{i_0} k^l > 0$ , sinon  $p_{i_1}$  ne serait pas premier et on a également  $a_1 = p_{i_1} > k^l - 1$ .

De plus, puisque  $p_{i_0} = a_0$  et  $p_{i_1} = a_1$  on peut, pour tout entier  $n$ , exprimer  $a_{n+1}$  comme suit :

$$a_{n+1} = a_1 k^{nl} + \frac{k^{nl} - 1}{k^l - 1} (a_1 - a_0 k^l).$$

En particulier, on a

$$(k^l - 1)a_{n+1} = -k^l(k^{nl} - 1)a_0 \pmod{a_1}.$$

Pour  $n = a_1 - 1$ , on obtient, grâce au petit théorème de Fermat,  $k^{nl} = 1 \pmod{a_1}$ , et donc :

$$(k^l - 1)a_{a_1} = 0 \pmod{a_1}.$$

Puisque  $a_1 > k^l - 1$ ,  $a_1$  divise  $a_{a_1}$  et donc  $a_{a_1}$  n'est pas premier et cela contredit le fait que  $\mathbf{a}$  est une suite de nombres premiers.

Ainsi, il n'existe pas de  $k$ -automate dénombrable  $\mathcal{A} = (Q, \phi, q_0, F)$  avec  $F$  fini qui reconnaît  $\mathbb{P}$  par lecture GD.  $\square$

Pour montrer que l'ensemble des nombres premiers n'est également pas  $k^\infty$ -reconnaissable GD, nous allons utiliser le lemme qui suit.

**Lemme 3.4.** *Il existe une suite  $\mathbf{x} = (x_n)_{n \geq 0}$  suffixe en base  $k$  et telle que, pour tout entier  $n \geq 0$ ,  $x_n$  est un nombre premier.*

On construit une telle suite  $\mathbf{x} = (x_n)_{n \geq 0}$  par récurrence. On fixe un nombre premier  $x_0 > k$  puis, pour tout entier  $n \geq 0$ , le théorème de la progression arithmétique de Dirichlet assure qu'il existe une infinité de nombres premiers congrus à  $x_n$  modulo  $k^{|\rho_k(x_n)|}$ . On choisit alors  $x_{n+1} > k^{|\rho_k(x_n)|}$  parmi ces nombres premiers.

*Démonstration. Deuxième partie du Théorème 3.1*

On raisonne une fois de plus par l'absurde. Supposons que l'ensemble des nombres premiers  $\mathbb{P}$  soit  $k^\infty$ -reconnaissable DG. Il existe donc un  $k$ -automate dénombrable  $\mathcal{A} = (Q, \phi, q_0, F)$  avec  $F$  fini qui reconnaît  $\mathbb{P}$ .

Si  $F$  est fini, le Lemme 3.4 assure l'existence d'une suite suffixe  $\mathbf{x} = (x_n)_{n \geq 0}$  de nombres premiers et donc d'une famille suffixe à  $|F| + 1$  éléments de nombres premiers. On peut donc appliquer le lemme de pompage 3.2 à la famille  $\{x_n \mid n \in \llbracket 0, |F| \rrbracket\}$ .

Il existe deux nombres premiers  $x_{i_0} < x_{i_1}$  issus de la suite  $\mathbf{x}$  tels que la suite  $\mathbf{a}' = (a'_n)_{n \geq 0}$ , définie pour tout  $n \geq 0$  par :

$$a'_n = (x_{i_1} - x_{i_0}) \frac{k^{nl} - 1}{k^l - 1} + x_{i_0},$$

où  $l = \lfloor \log_k x_{i_1} \rfloor - \lfloor \log_k x_{i_0} \rfloor$ , est une suite de nombres premiers. On a également pour tout entier  $n \geq 1$ ,

$$a'_{n+1} = (x_{i_1} - x_{i_0}) k^l \frac{k^{nl} - 1}{k^l - 1} + a'_1,$$

De cette égalité, on ne retiendra que l'égalité modulo  $a'_1$  suivante :

$$a'_{n+1}(k^l - 1) = (x_{i_1} - x_{i_0}) k^l (k^{nl} - 1) \pmod{a'_1}.$$

Pour  $n = a'_1 - 1$ , cette égalité modulo  $a'_1$  devient, grâce au petit théorème de Fermat :

$$a'_{a'_1}(k^l - 1) = 0 \pmod{a'_1}.$$

Comme  $a'_1$  est premier et satisfait  $a'_1 > k^l - 1$ , cela implique que  $a'_{a'_1} = 0 \pmod{a'_1}$ , ce qui contredit le fait que la suite  $\mathbf{a}'$  est une suite de nombres premiers.  $\square$

**Remarque.** L'ensemble des nombres premiers n'est pas non plus  $k^\infty$ -reconnaissable pour les deux sens de lecture.

Supposons que l'ensemble des nombres composés soit  $k^\infty$ -reconnaissable GD. Il existe un  $k$ -automate dénombrable  $\mathcal{A} = (Q, \phi, q_0, F)$  avec  $F$  fini qui reconnaît  $\mathbb{N} \setminus \mathbb{P}$  par lecture GD. Par commodité, on suppose aussi que  $Q$  est restreint aux états accessibles à partir de  $q_0$ .

Dans le graphe de  $\mathcal{A}$ , tous les éléments de  $Q \setminus F$  sont à distance 1 de  $F$ . En effet, pour tout nombre premier  $p$ , on a  $\rho_k(p)0 = \rho_k(kp)$  et  $kp$  est un nombre composé. Ainsi, tous les états de  $Q$  accessibles à partir de  $q_0$  sont à distance finie de  $F$ . Comme  $F$  est fini et que l'automate  $\mathcal{A}$  est de degré fini, l'ensemble  $Q$  est donc fini. Par conséquent, cela implique que l'ensemble des nombres composés et celui des nombres premiers sont reconnus par des automates finis, ce qui est faux. Voir par exemple [13] (chap. 5).

Dans le cas de la lecture DG, on procède également par l'absurde. Supposons que l'ensemble des nombres composés soit  $k^\infty$ -reconnaissable DG, c'est-à-dire qu'il existe un  $k$ -automate dénombrable  $\mathcal{A} = (Q, \phi, q_0, F)$  avec  $F$  fini qui reconnaît  $\mathbb{N} \setminus \mathbb{P}$  par lecture DG. Dans le graphe de  $\mathcal{A}$ , tous les états sont à distance au plus 2 de  $F$ . En effet, soit  $Q$  un état, que l'on suppose différent de  $q_0$  puisque  $q_0$  est dans  $F$ , et soit  $w$  un mot non vide de  $\llbracket 0, k-1 \rrbracket^*$  conduisant à  $q$  depuis  $q_0$  :  $\overleftarrow{\phi}(w, q_0) = q$ . On note  $n = \pi_k(w)$ .

Considérons l'ensemble des entiers  $E = \left\{ n + j \cdot k^{|w|} \mid j \in \llbracket 1, k+1 \rrbracket \right\}$ , dont les représentations en base  $k$  forment l'ensemble suivant :

$$\rho_k(E) = \{jw \mid j \in \llbracket 1, k-1 \rrbracket\} \cup \{10w, 11w\}.$$

Comme  $k^l$  vaut 1 ou  $-1$  modulo  $k+1$ , l'un des éléments de  $E$  est divisible par  $k+1$ , et est donc composé. Plus précisément, l'entier  $n + ik^{|w|}$ , où  $i = n \cdot (-1)^{|w|+1} \pmod{k+1}$  est divisible par  $k+1$ . Le mot correspondant conduit donc à un état de  $F$ , et ce en une ou deux étapes depuis  $q$ . Comme  $F$  est fini et que le degré de  $\mathcal{A}$  borné,  $Q$  est donc fini. Comme l'ensemble des nombres premiers n'est pas  $k$ -reconnaissable, on aboutit à la même contradiction que dans le cas de la lecture GD.

### 3.2. Un critère de non- $k^\infty$ -reconnaissabilité GD.

**Lemme 3.5.** Soient  $k \geq 2$  et  $E = \{e_n \mid n \in \mathbb{N}\}$  un ensemble d'entiers satisfaisant les conditions suivantes :

- (1) L'ensemble des parties fractionnaires  $\{\{\log_k(e_n)\} \mid n \in \mathbb{N}\}$  est dense dans  $[0, 1[$ ,
- (2) L'ordre de croissance de la suite  $e_n$  est plus grand que  $\alpha^n$ , pour tout  $\alpha > 1$ .

Alors l'ensemble  $E$  n'est pas  $k^\infty$ -reconnaisable GD.

*Démonstration.* Le fait que l'ensemble  $\{\{\log_k(e_n)\} \mid n \in \mathbb{N}\}$  des parties fractionnaires des éléments de  $E$  soit dense dans  $[0, 1[$  assure l'existence d'une suite préfixe d'éléments de  $E$ . En effet, sous cette condition de densité, on peut extraire une sous-suite  $\mathbf{a} = (a_n)_{n \geq 0}$  d'entiers de  $E$  vérifiant

$$k^{\{\log_k(a_n)\}} < k^{\{\log_k(a_{n+1})\}} < k^{\{\log_k(a_n)\}} + \frac{1}{k^{\lfloor \log_k(a_n) \rfloor}},$$

c'est-à-dire telle que :

$$\frac{a_n}{k^{\lfloor \log_k(a_n) \rfloor}} < \frac{a_{n+1}}{k^{\lfloor \log_k(a_{n+1}) \rfloor}} < \frac{a_n}{k^{\lfloor \log_k(a_n) \rfloor}} + \frac{1}{k^{\lfloor \log_k(a_n) \rfloor}}.$$

Cette condition est en fait équivalente à la condition vérifiée par une suite préfixe d'entiers :

$$0 < a_{n+1} - a_n k^{\lfloor \log_k a_{n+1} \rfloor - \lfloor \log_k a_n \rfloor} < k^{\lfloor \log_k a_{n+1} \rfloor - \lfloor \log_k a_n \rfloor}.$$

Supposons, sous les conditions du lemme, que  $E$  soit  $k^\infty$ -reconnaisable. On peut appliquer le Lemme 3.1 à la suite  $\mathbf{a}$ . Il existe deux entiers  $i_0$  et  $i_1$  et une suite  $\mathbf{b} = (b_n)_{n \geq 0}$  d'éléments de  $E$  définie par

$$b_n = a_{i_0} k^{nl} + (a_{i_1} - a_{i_0} k^l) \frac{k^{nl} - 1}{k^l - 1},$$

avec  $l = \lfloor \log_k a_{i_1} \rfloor - \lfloor \log_k a_{i_0} \rfloor$ . Cela contredit l'hypothèse que l'ordre de croissance de la suite  $e_n$  est plus grand que  $\alpha^n$  pour tout  $\alpha > 1$ , puisque  $b_n < (a_{i_0} + 1)k^{nl}$  pour tout  $n \geq 0$ . □

#### 4. Non-équivalence entre $k^\infty$ -reconnaisabilités GD et DG

**Notation.** Soit un entier  $k \geq 2$ .

On note  $\overleftarrow{\text{Rec}}_\infty(k)$  (resp.  $\overrightarrow{\text{Rec}}_\infty(k)$ ) la famille des ensembles d'entiers  $k^\infty$ -reconnaisables DG (resp. GD).

De même,  $\overleftarrow{\text{coRec}}_\infty(k)$  (resp.  $\overrightarrow{\text{coRec}}_\infty(k)$ ) désigne la famille d'ensembles d'entiers  $k^\infty$ -co-reconnaisables DG (resp. GD).

La famille des ensembles  $k$ -reconnaisables au sens usuel est quant à elle notée  $\text{Rec}(k)$ .

Les notions de reconnaissabilité par automate dénombrable par lecture GD et lecture DG ne sont pas équivalentes. Comme nous venons de le voir pour les lemmes de pompage, les propriétés des ensembles d'entiers utilisées dans le sens GD et dans le sens DG sont assez éloignées.



Nous donnons dans ce paragraphe des exemples d'ensembles d'entiers reconnus dans un sens et pas dans l'autre.

**Théorème 4.1.** *Soit un entier  $k \geq 2$ .*

*Les notions de  $k^\infty$ -reconnaissabilité GD et de  $k^\infty$ -reconnaissabilité DG ne sont pas équivalentes :*

$$\overrightarrow{\text{Rec}}_\infty(k) \setminus \overleftarrow{\text{Rec}}_\infty(k) \neq \emptyset \quad \text{et} \quad \overleftarrow{\text{Rec}}_\infty(k) \setminus \overrightarrow{\text{Rec}}_\infty(k) \neq \emptyset.$$

Pour montrer ce théorème, nous allons exhiber, via les Propositions 4.1 et 4.2, deux familles d'ensembles d'entiers qui sont reconnaissables par des automates dénombrables uniquement par un sens de lecture.

**Proposition 4.1.** *Soit un entier  $k \geq 2$ .*

*L'ensemble d'entiers  $X = \{k^n + n \mid n \geq k\}$  est  $k^\infty$ -reconnaissable GD mais n'est pas  $k^\infty$ -reconnaissable DG. Il n'est pas non plus  $k^\infty$ -co-reconnaissable DG.*

*Démonstration.* Remarquons dans un premier temps qu'on a l'égalité

$$\rho_k(X) = \{10^{n-|\rho_k(n)|} \rho_k(n) \mid n \geq k\}.$$

Afin de construire un  $k$ -automate GD qui reconnaît  $X$ , on utilise le résultat préliminaire suivant : Pour tous entiers  $n \neq n'$  supérieurs ou égaux à  $k$ , tels que  $\rho_k(n) = n_l \cdots n_1 n_0$  et  $\rho_k(n') = n'_l \cdots n'_1 n'_0$ , les mots  $10^{n-l-1} n_l n_{l-1}$  et  $10^{n'-l'-1} n'_l n'_{l-1}$  sont différents.

En effet, si pour deux entiers  $n < n'$ , on a  $10^{n-l-1} n_l n_{l-1} = 10^{n'-l'-1} n'_l n'_{l-1}$  alors, en particulier  $n - l - 1 = n' - l' - 1$ . En observant le comportement de la fonction  $n \mapsto n - \lfloor \log_k n \rfloor$ , cela n'est possible que si  $n = k^{l+1} - 1$  et  $n' = k^{l'+1} - 1$ . Or dans ce cas,  $n_l n_{l-1} = (k-1)(k-1)$  et  $n'_l n'_{l-1} = 10$  et donc  $10^{n-l-1} n_l n_{l-1} \neq 10^{n'-l'-1} n'_l n'_{l-1}$ , ce qui aboutit à une contradiction.

Ce résultat n'est, en fait, vraiment utile que dans le cas où  $k = 2$  puisque dans ce cas  $k - 1 = 1$  et de ce fait  $n_l = n'_l$  donc on doit aller regarder les deuxièmes chiffres des mots  $\rho_k(n)$  et  $\rho_k(n')$  pour pouvoir être sûr de différencier  $10^{n-l-1} \rho_k(n)$  et  $10^{n'-l'-1} \rho_k(n')$ . Cependant, cela permet d'obtenir une construction d'un  $k$ -automate GD qui reconnaît  $X$  valable quel que soit  $k \geq 2$ .

Construisons le  $k$ -automate  $\mathcal{A} = (Q, \phi, q_{(0,0)}, F)$  qui reconnaît  $X = \{k^n + n \mid n \geq k\}$  par lecture GD. L'ensemble  $Q$  est composé de  $k$  lignes d'états et d'un arbre complet :

$$Q = \{q_{(n,i)} \mid n \in \mathbb{N}, i \in \llbracket 0, k-1 \rrbracket\} \cup \{s_w, w \in \llbracket 0, k-1 \rrbracket^*\}.$$

L'ensemble d'acceptation est  $F = \{s_\varepsilon\}$  et la fonction de transition  $\phi$  est définie par :

$$- \phi(1, q_{(0,0)}) = q_{(1,0)},$$

- pour tout  $n \geq 1$ ,  $\phi(0, q_{(n,0)}) = q_{(n+1,0)}$ ,
- pour tout  $n \geq k$  tel que  $\rho_k(n) = n_l \cdots n_1 n_0$ ,

$$\phi(n_l, q_{(n+1-|\rho_k(n)|,0)}) = q_{(n+1-|\rho_k(n)|,n_l)},$$

- pour tout mot  $w = w_0 w_1 \cdots w_l$  de  $\llbracket 0, k-1 \rrbracket^+$ ,  $\phi(w_0, s_w) = s_{w_1 \cdots w_l}$
- pour tout  $n \geq k$  tel que  $\rho_k(n) = n_l \cdots n_1 n_0$ ,

$$\phi(n_{l-1}, q_{(n+1-|\rho_k(n)|,n_l)}) = s_{n_{l-2} \cdots n_1 n_0}.$$

Cet automate est présenté, dans le cas  $k = 2$ , à la figure 4.1. Il est facile de montrer que l'automate  $\mathcal{A}$  est de degré fini et reconnaît  $X$ .

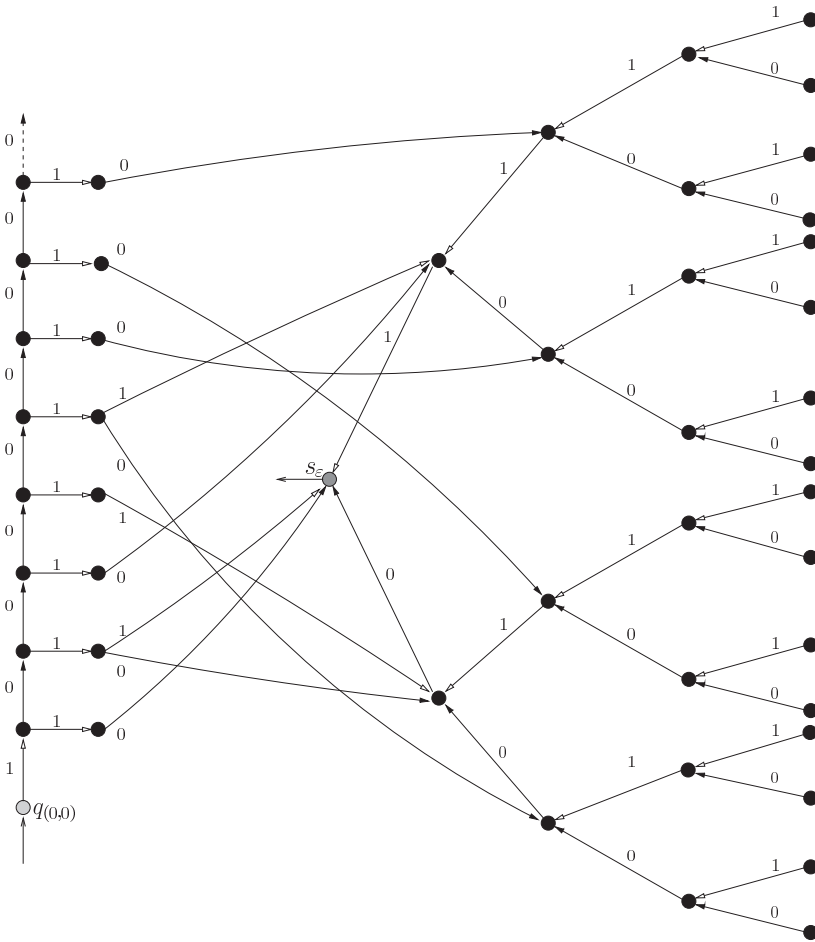


FIGURE 4.1. 2-automate GD reconnaissant  $\{2^n + n \mid n \geq 2\}$ .

Lorsque  $k > 2$ , on peut en fait simplifier cet automate en supprimant les états  $\{q_{(n,i)} \mid n \in \mathbb{N}, i \in \llbracket 1, k-1 \rrbracket\}$  et en modifiant la définition de  $\phi$  sur les états  $\{q_{(n,0)} \mid n \in \mathbb{N}\}$  en posant, pour tout  $n \geq k$  tel que  $\rho_k(n) = n_l \cdots n_1 n_0$ ,

$$\phi(n_l, q_{(n+1-|\rho_k(n)|,0)}) = s_{n_{l-1} \cdots n_1 n_0}.$$

On procède par l'absurde pour montrer que  $X$  n'est ni  $k^\infty$ -reconnaissable DG ni  $k^\infty$ -co-reconnaissable DG.

Supposons que  $X$  soit reconnu par un  $k$ -automate  $\mathcal{A}' = (Q', \psi, q'_0, F')$  fini ou dénombrable par lecture DG avec  $F'$  fini.

Il existe une suite d'entiers  $\mathbf{u} = (u_n)_{n \geq 0}$  extraite de  $X$  telle que  $\rho_k(u_n)$  soit suffixe de  $\rho_k(u_{n+1})$  : il suffit de choisir  $u_0 \in X$  et de poser  $u_{n+1} = k^{u_n} + u_n$ .

D'après le Lemme 3.2, il existe deux mots  $v$  et  $w$ , avec  $v \neq \varepsilon$ , tels que, pour tout  $i \geq 0$ , l'entier  $a_i = \pi_k(v^i w)$  appartient à  $X$ . Il existe donc, pour tout  $i \geq 0$ , un entier  $n_i$  tel que  $a_i = k^{n_i} + n_i$ , c'est-à-dire que  $v^i w = 10^{n_i - |\rho_k(n_i)|} \rho_k(n_i)$ . Puisque, par construction, la suite  $(n_i)_{i \geq 0}$  est strictement croissante, pour  $i$  suffisamment grand, on a  $n_i - |\rho_k(n_i)| \geq 2|v| - 1$ . Ainsi  $v^2$  est préfixe de  $10^{n_i - |\rho_k(n_i)|}$ , ce qui est absurde. L'ensemble  $X$  n'est donc pas  $k^\infty$ -reconnaissable DG.

Supposons maintenant que  $X$  soit  $k^\infty$ -co-reconnaissable DG. L'ensemble  $\mathbb{N} \setminus X$  est donc reconnu par un  $k$ -automate  $\mathcal{A}'' = (Q'', \delta, q''_0, F'')$  avec  $F''$  fini par lecture DG. Or cela est également absurde. En effet, comme pour tout entier  $n \geq 1$  de  $X$  l'entier  $kn + 1$  n'appartient pas à  $X$ , cela implique que pour tout état  $q$  de  $Q'' \setminus F''$ ,  $\delta(1, q)$  appartient à  $F''$ . L'automate  $\mathcal{A}''$  étant de degré entrant fini, tous les états de  $Q'' \setminus F''$  accessibles à partir de  $q''_0$  sont en nombre fini, et on peut donc supposer que  $Q'' \setminus F''$  et  $Q''$  sont finis. On doit pouvoir alors appliquer le même lemme de la pompe à la suite  $\mathbf{u} = (u_n)_{n \geq 0}$  présentée ci-dessus et on a vu que cela aboutit à une contradiction.  $\square$

**Proposition 4.2.** *L'ensemble d'entiers  $Y = \{(kn + 1)k^{|\rho_k(n)|}, n \in \mathbb{N}\}$  est  $k^\infty$ -reconnaissable DG mais n'est pas  $k^\infty$ -reconnaissable GD. Il n'est pas non plus  $k^\infty$ -co-reconnaissable GD.*

*Démonstration.* On a  $\rho_k(Y) = \{\rho_k(n)10^{|\rho_k(n)|} \mid n \in \mathbb{N}\}$ . Dans un premier temps, on construit le  $k$ -automate  $\mathcal{A} = (Q, \phi, q_0, F)$  suivant qui reconnaît  $Y$  par lecture DG où  $Q$  est formé de deux lignes d'états :  $Q = \{q_n \mid n \in \mathbb{N}\} \cup \{q'_n \mid n \in \mathbb{N}\}$ , l'ensemble terminal est  $F = \{q'_0\}$  et où la fonction de transition  $\phi$  est définie par :

- pour tout  $n \geq 0$ ,  $\phi(0, q_n) = q_{n+1}$ ,
- pour tout  $n \geq 0$ ,  $\phi(1, q_n) = q'_n$ ,
- pour tout  $n \geq 1$  et tout  $i \in \llbracket 0, k-1 \rrbracket$ ,  $\phi(i, q'_{n+1}) = q'_n$ ,
- pour tout  $i \in \llbracket 1, k-1 \rrbracket$ ,  $\phi(i, q'_1) = q'_0$ .

Cet automate est présenté à la figure 4.2.

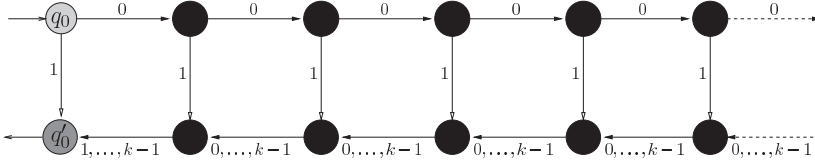


FIGURE 4.2.  $k$ -automate DG reconnaissant  $\{(kn + 1)k^{|\rho_k(n)|} \mid n \in \mathbb{N}\}$ .

Pour montrer que  $Y$  n'est pas  $k^\infty$ -reconnaissable GD, on procède par l'absurde comme dans la preuve de la Proposition 4.1 :

Pour montrer que  $Y$  n'est pas reconnu par un  $k$ -automate fini ou dénombrable par lecture GD ayant un ensemble terminal fini, on met en défaut le Lemme 3.1, en utilisant la suite préfixe  $\mathbf{u}' = (u'_n)_{n \geq 0}$  définie par  $u'_0 \in Y$  et  $u'_{n+1} = (ku'_n + 1)k^{|\rho_k(u'_n)|}$ .

On montre par l'absurde que  $Y$  n'est pas  $k^\infty$ -co-reconnaissable GD. Supposons que  $\mathbb{N} \setminus Y$  soit reconnu par un  $k$ -automate  $\mathcal{A}' = (Q', \psi, q'_0, F')$  fini ou dénombrable par lecture GD avec  $F'$  fini. Pour tout entier  $n$  de  $Y$ ,  $kn + 1$  n'appartient pas à  $Y$  puisque le seul élément de  $Y$  non divisible par  $k$  est 1, de sorte que tout état de  $Q' \setminus F'$  accessible à partir de  $q'_0$  est à distance 1 de  $F'$ . On peut alors restreindre  $\mathcal{A}'$  à un automate fini et donc  $Y$  est reconnaissable. Cela est absurde puisque  $Y$  n'est pas  $k^\infty$ -reconnaissable GD. □

**Remarque.** Les supports des lettres des mots  $k^\infty$ -automatiques, introduits et étudiés dans [18, 20] sont des ensembles  $k^\infty$ -reconnaissables GD ou  $k^\infty$ -co-reconnaissables GD. On peut donc engendrer ces ensembles en utilisant un procédé de substitution-projection, où l'alphabet de la substitution est dénombrable. C'est d'ailleurs ce qui a, dans un premier temps, motivé le choix de la reconnaissabilité GD pour les ensembles d'entiers dans [20].

Ce type de substitution a été étudié d'un point de vue dynamique par S. Ferenczi dans [14].

D'autre part, une étude de la répartition modulo 1 des suites  $(n\alpha)_{n \in E}$  pour  $\alpha$  irrationnel et  $E$  un ensemble  $k^\infty$ -reconnaissable GD a été menée par C. Mauduit et E. Fouvry [15] dans le cas où  $E$  est l'ensemble  $E_1$  défini au paragraphe 2.2 et par C. Mauduit dans [20], dans le cas où le  $k$ -automate reconnaissant  $E$  est porté par un réseau de  $\mathbb{R}^d$ .

## 5. Propriétés des familles d'ensembles d'entiers $k^\infty$ -automatiques

**5.1. Automates minimaux DG et GD.** La notion d'automate minimal [13] peut être définie pour des langages quelconques, y compris donc les

langages des représentations en base  $k$  des ensembles  $k^\infty$ -reconnaissables, et constitue un outil puissant. La minimisation du nombre d'états via la congruence syntaxique (congruence de Nérode) permet d'obtenir un critère de  $k^\infty$ -reconnaissabilité mais également de mettre en évidence des cas de dégénérescence, c'est-à-dire, des cas de  $k$ -reconnaissabilité simple. Elle sera également utile pour montrer des résultats de stabilité exposés au paragraphe 5.2.

Nous aurons besoin par la suite des notations suivantes :

**Notation.** Soient  $L$  un langage de  $\llbracket 0, k-1 \rrbracket^*$  et  $w$  un mot de  $\llbracket 0, k-1 \rrbracket^*$ . On note :

$$\begin{aligned}\overrightarrow{L_w} &= w^{-1}L = \{v \in \llbracket 0, k-1 \rrbracket^* \mid wv \in L\}, \\ \overleftarrow{L_w} &= Lw^{-1} = \{v \in \llbracket 0, k-1 \rrbracket^* \mid v w \in L\}.\end{aligned}$$

La famille de langages  $\{\overrightarrow{L_w} \mid w \in \llbracket 0, k-1 \rrbracket^*\}$  (resp.  $\{\overleftarrow{L_w} \mid w \in \llbracket 0, k-1 \rrbracket^*\}$ ) définit une relation de congruence à droite (resp. à gauche) sur  $\llbracket 0, k-1 \rrbracket^*$ , notée  $\overrightarrow{\sim}_L$  (resp.  $\overleftarrow{\sim}_L$ ) de la manière suivante : pour tous mots  $w, v$  de  $\llbracket 0, k-1 \rrbracket^*$ ,

$$w \overrightarrow{\sim}_L v \iff \overrightarrow{L_w} = \overrightarrow{L_v} \quad (\text{resp. } w \overleftarrow{\sim}_L v \iff \overleftarrow{L_w} = \overleftarrow{L_v}).$$

**Remarque.** Si  $w$  et  $u$  sont des mots de  $\llbracket 0, k-1 \rrbracket^*$ , on a :

$$\overrightarrow{L_{wu}} = \overrightarrow{(\overrightarrow{L_w})_u} \quad \text{et} \quad \overleftarrow{L_{wu}} = \overleftarrow{(\overleftarrow{L_u})_w}$$

**Notation.** Dans le cas d'un  $k$ -automate fini ou dénombrable quelconque  $\mathcal{A} = (Q, \phi, q_0, F)$  on utilisera, par analogie avec la notation  $\overrightarrow{L_w}$  (resp.  $\overleftarrow{L_w}$ ), la notation  $\overrightarrow{X_q}$  (resp.  $\overleftarrow{X_q}$ ) pour tout état  $q \in Q$  et tout sous-ensemble d'états  $X \subset Q$  désignant l'ensemble des mots de  $\llbracket 0, k-1 \rrbracket^*$  qui sont les étiquettes des chemins débutant en  $q$  et aboutissant dans un état de  $X$  par lecture GD (resp. DG) :

$$\begin{aligned}\overrightarrow{X_q} &= \{v \in \llbracket 0, k-1 \rrbracket^* \mid \overrightarrow{\phi}(v, q) \in X\} \\ (\text{resp. } \overleftarrow{X_q} &= \{v \in \llbracket 0, k-1 \rrbracket^* \mid \overleftarrow{\phi}(v, q) \in X\}),\end{aligned}$$

de sorte que, pour tout mot  $w$  de  $\llbracket 0, k-1 \rrbracket^*$  tel que  $\overrightarrow{\phi}(w, q_0) = q$  (resp.  $\overleftarrow{\phi}(w, q_0) = q$ ), on ait  $\overrightarrow{F_q} = \overrightarrow{L(\mathcal{A})_w}$  (resp.  $\overleftarrow{F_q} = \overleftarrow{L(\mathcal{A})_w}$ ).

**Définition.** Soit  $E$  un ensemble d'entiers quelconque. On définit le  $k$ -automate minimal GD associé à  $E$ , noté  $\overrightarrow{\mathcal{M}(E)} = (\overrightarrow{Q(E)}, \overrightarrow{\phi_E}, \overrightarrow{q_{0,E}}, \overrightarrow{F(E)})$  comme suit :

$$- \overrightarrow{Q(E)} = \{\overrightarrow{(\rho_k(E))_w} \mid w \in \llbracket 0, k-1 \rrbracket^*\},$$

- $\overrightarrow{q}_{0,E} = \overrightarrow{(\rho_k(E))_\varepsilon} = \rho_k(E)$ ,
- $\overrightarrow{F}(E) = \left\{ \overrightarrow{(\rho_k(E))_w} \mid w \in \rho_k(E) \right\}$ ,
- et la fonction  $\overrightarrow{\phi}_E$  est définie pour tout  $L \in \overrightarrow{Q}(E)$  et tout  $i \in \llbracket 0, k-1 \rrbracket$  par  $\overrightarrow{\phi}_E(i, L) = \overrightarrow{L}_i$ . Cette fonction est étendue à  $\llbracket 0, k-1 \rrbracket^*$  en posant  $\overrightarrow{\phi}_E(w, L) = \overrightarrow{L}_w$ .

On définit de manière similaire, pour un ensemble d'entiers  $E$ , le  $k$ -automate minimal DG associé à  $E$ , noté  $\overleftarrow{\mathcal{M}}(E) = (\overleftarrow{Q}(E), \overleftarrow{\phi}_E, \overleftarrow{q}_{0,E}, \overleftarrow{F}(E))$ , en inversant le sens des flèches dans les notations.

Par exemple, les automates minimaux GD et DG associés à l'ensemble  $E_1$ ,  $\overrightarrow{\mathcal{M}}_{E_1}$  et  $\overleftarrow{\mathcal{M}}_{E_1}$ , sont représentés dans les Figures 5.1 et 5.2, avec la convention  $L = \rho_2(E_1)$ .

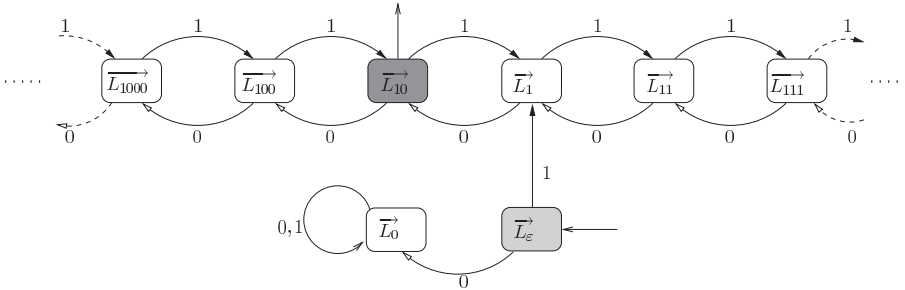


FIGURE 5.1. Graphe du  $k$ -automate minimal GD associé à  $E_1$ .

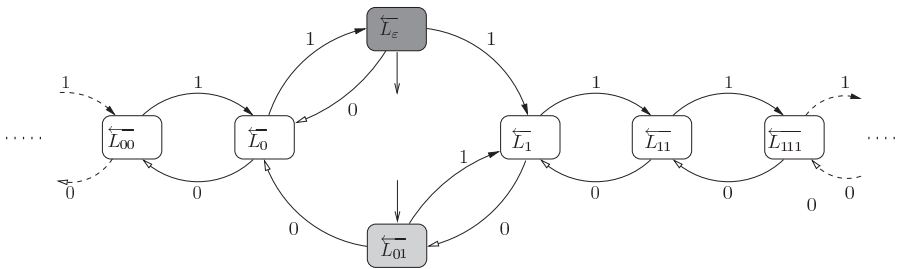


FIGURE 5.2. Graphe du  $k$ -automate minimal DG associé à  $E_1$ .

**Remarque.** Il est facile de montrer qu'un ensemble d'entiers est reconnu par le  $k$ -automate minimal GD (*resp.* DG) qui lui est associé. D'autre part, les automates minimaux sont des automates accessibles, c'est-à-dire que tout état est accessible à partir de l'état initial mais ces automates sont

aussi réduits, c'est à dire que si deux états  $s$  et  $s'$  d'un automate minimal d'ensemble final  $F$  vérifient  $\overrightarrow{F}_q = \overrightarrow{F}_{q'}$  (ou  $\overleftarrow{F}_q = \overleftarrow{F}_{q'}$  selon le sens de lecture), alors  $q = q'$ . Voir par exemple [25, Chapitre 1, section 3].

**Proposition 5.1.** *Soient  $E$  un ensemble  $k^\infty$ -reconnaisable GD et  $\overrightarrow{\mathcal{M}(E)} = (\overrightarrow{Q(E)}, \overrightarrow{\phi_E}, \overrightarrow{q}_{0,E}, \overrightarrow{F(E)})$  son  $k$ -automate minimal GD.*

*L'ensemble  $E$  est  $k^\infty$ -reconnaisable GD si et seulement si :*

- (1) *les états de la composante co-accessible de  $\overrightarrow{\mathcal{M}(E)}$ , c'est-à-dire les états à partir desquels au moins un état final est accessible, sont de degré fini,*
- (2)  *$\overrightarrow{F(E)} = \{\overrightarrow{L}_w \mid w \in \rho_k(E)\}$  est fini. De plus, si  $\overrightarrow{Q(E)}$  est fini, alors  $E$  est  $k$ -reconnaisable.*

*Le même résultat tient bien sûr dans le cas de la lecture DG. L'énoncé s'obtient en inversant le sens des flèches dans les notations.*

Ce critère de  $k^\infty$ -reconnaisabilité est une conséquence directe de la proposition suivante.

**Proposition 5.2.** *Soient  $E$  un ensemble  $k^\infty$ -reconnaisable GD et  $\overrightarrow{\mathcal{M}(E)} = (\overrightarrow{Q(E)}, \overrightarrow{\phi_E}, \overrightarrow{q}_{0,E}, \overrightarrow{F(E)})$  son  $k$ -automate minimal GD.*

*Si un  $k$ -automate accessible  $\mathcal{A} = (Q, \phi, q_0, F)$  reconnaît  $E$  par lecture GD, alors il existe une application surjective  $\overrightarrow{\Phi} : Q \rightarrow \overrightarrow{Q(E)}$  telle que :*

- $\overrightarrow{\Phi}(q_0) = \overrightarrow{q}_{0,E}$ ,
- pour tout  $i$  de  $\llbracket 0, k-1 \rrbracket$  et tout état  $q$  de  $Q$ ,  $\overrightarrow{\Phi}(\phi(i, q)) = \overrightarrow{\phi_E}(i, \overrightarrow{\Phi}(q))$ ,
- $\overrightarrow{\Phi}(F) = \overrightarrow{F(E)}$ .

*D'autre part, si  $\mathcal{A}$  est de degré fini, alors les états de  $\overrightarrow{Q(E)}$  à partir desquels au moins un état final est accessible (états de la composante co-accessible de  $\overrightarrow{\mathcal{M}(E)}$ ) sont de degré fini.*

*Le même résultat tient bien sûr dans le cas de la lecture DG. L'énoncé s'obtient en inversant le sens des flèches dans les notations.*

*Démonstration.* Pour une démonstration de la première partie de cette proposition, on se réfère également à [25]. La conservation du degré fini se montre par contraposée :

Notons  $L = \rho_k(E)$ . Si un état  $\overrightarrow{L}_w \neq \emptyset$  de  $\overrightarrow{Q(E)}$  est de degré infini, montrons que tout automate  $\mathcal{A} = (Q, \phi, q_0, F)$  se projetant par  $\overrightarrow{\Phi}$  sur  $\overrightarrow{\mathcal{M}(E)}$  a forcément un état de degré infini.

Supposons qu'il existe une infinité de mots  $v \in \llbracket 0, k-1 \rrbracket^*$  pour lesquels les langages  $\overrightarrow{L}_v$  sont distincts, non vides et tels que  $\overrightarrow{L}_w = \overrightarrow{\phi_E}(i, \overrightarrow{L}_v)$  pour une certaine lettre  $i \in \llbracket 0, k-1 \rrbracket$ . En utilisant la définition de la fonction

$\overrightarrow{\Phi}$ , on obtient qu'il existe un état  $q$  de  $Q$  tel que  $\overrightarrow{F}_q = \overrightarrow{L}_w$  et une infinité d'états distincts  $q_v$  de  $Q$  tels que  $\overrightarrow{F}_{q_v} = \overrightarrow{L}_v$  vérifiant  $\overrightarrow{F}_{\phi(i,q_v)} = \overrightarrow{F}_q$ . Puisque  $\overrightarrow{F}_q \neq \emptyset$ , il existe un mot  $u$  tel que  $\overrightarrow{\phi}(u, q) \in F$  et ainsi,  $\overrightarrow{\phi}(iu, q_v) \in F$  pour une infinité d'états distincts  $q_v$  de  $Q$ . Ces états  $q_v$  sont tous à distance au plus  $|u| + 1$  de l'ensemble terminal  $F$ . Puisque  $F$  est fini et que l'automate  $\mathcal{A}$  est de degré sortant borné par  $k$ , un de ces états est forcément de degré entrant infini.  $\square$

**5.2. Propriétés de stabilité.** Seules certaines propriétés de stabilité de la famille des ensembles d'entiers  $k$ -reconnaissables s'étendent au cadre des ensembles  $k^\infty$ -reconnaissables. Dans ce paragraphe, nous décrivons les comportements de deux familles d'ensembles d'entiers  $k^\infty$ -reconnaissables sous les opérations usuelles d'intersection, de passage au complémentaires, d'union, ainsi que les liens entre  $\overrightarrow{\text{Rec}}_\infty(k)$  et  $\overrightarrow{\text{Rec}}_\infty(k')$  (resp.  $\overleftarrow{\text{Rec}}_\infty(k)$  et  $\overleftarrow{\text{Rec}}_\infty(k')$ ).

**Proposition 5.3.** *Soit un entier  $k \geq 2$ .*

$$\begin{aligned} \text{Rec}(k) &= \overrightarrow{\text{Rec}}_\infty(k) \cap \overrightarrow{\text{coRec}}_\infty(k), \\ \text{Rec}(k) &= \overleftarrow{\text{Rec}}_\infty(k) \cap \overleftarrow{\text{coRec}}_\infty(k). \end{aligned}$$

*Démonstration.* Les preuves diffèrent quelque peu selon le sens de lecture. Cela tient à la gestion des écritures impropres d'entiers dans le sens de lecture DG.

Commençons par la preuve dans le sens GD.

Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaissable GD pour lequel  $\overline{E} = \mathbb{N} \setminus E$  est aussi  $k^\infty$ -reconnaissable GD. On note  $L = \rho_k(E)$  et  $L' = \rho_k(\overline{E})$ . Pour montrer que  $\overline{E}$  est  $k$ -reconnaissable, on va montrer que son automate minimal GD  $\overrightarrow{\mathcal{M}}(E) = (\overrightarrow{Q}(E), \overrightarrow{\phi}_E, \overrightarrow{q}_{0,E}, \overrightarrow{F}(E))$  est fini, c'est-à-dire que  $\overrightarrow{Q}(E) = \{\overrightarrow{L}_w \mid w \in \llbracket 0, k-1 \rrbracket^*\}$  est fini.

Puisque  $E$  et  $\overline{E}$  sont  $k^\infty$ -reconnaissables GD, alors  $\{\overrightarrow{L}_w \mid w \in L\}$  et  $\{\overrightarrow{L}'_w \mid w \in L'\}$  sont finis. D'autre part, puisque  $\overrightarrow{L}'_w = \llbracket 0, k-1 \rrbracket^* \setminus \overrightarrow{L}_w$  si  $w \in \llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^*$ , les ensembles  $\{\overrightarrow{L}_w \mid w \in L'\}$  sont en bijection avec les ensembles  $\{\overrightarrow{L}'_w \mid w \in L'\}$ , de sorte que  $\{\overrightarrow{L}_w \mid w \in L \cup L'\}$  est aussi fini. Comme  $L \cup L' = \llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^* \cup \{\varepsilon\}$ , il reste à montrer la finitude de l'ensemble  $\{\overrightarrow{L}_w \mid w \in 0\llbracket 0, k-1 \rrbracket^*\}$ .

Puisque  $\overrightarrow{\mathcal{M}}(E)$  reconnaît un langage propre, aucun mot commençant par la lettre 0 n'est accepté par l'automate. Par conséquent, pour tout mot  $w \in 0\llbracket 0, k-1 \rrbracket^*$ , on a  $\overrightarrow{L}_w = \emptyset$ . Ainsi,  $\{\overrightarrow{L}_w \mid w \in 0\llbracket 0, k-1 \rrbracket^*\} = \{\emptyset\}$  et donc  $\{\overrightarrow{L}_w \mid w \in \llbracket 0, k-1 \rrbracket^*\}$  est fini. L'automate  $\overrightarrow{\mathcal{M}}(E)$  est alors un automate fini et  $\overline{E}$  est reconnaissable.



Passons maintenant à la preuve dans le sens DG.

Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaissable DG tel que  $\overline{E} = \mathbb{N} \setminus E$  est aussi  $k^\infty$ -reconnaissable DG. On note  $L = \rho_k(E)$  et  $L' = \rho_k(\overline{E})$ . Pour montrer que  $E$  est  $k$ -reconnaissable, on va montrer que son automate minimal DG  $\overleftarrow{\mathcal{M}}(E) = (\overleftarrow{Q}(E), \overleftarrow{\phi}_E, \overleftarrow{q}_{0,E}, \overleftarrow{F}(E))$  est fini, c'est-à-dire que  $\overleftarrow{Q}(E) = \{\overleftarrow{L}_w \mid w \in \llbracket 0, k-1 \rrbracket^*\}$  est fini. L'automate minimal de  $\overline{E}$  est noté  $\overleftarrow{\mathcal{M}}(\overline{E}) = (\overleftarrow{Q}(\overline{E}), \overleftarrow{\phi}_{\overline{E}}, \overleftarrow{q}_{0,\overline{E}}, \overleftarrow{F}(\overline{E}))$ .

Puisque  $E$  et  $\overline{E}$  sont  $k^\infty$ -reconnaissables DG, alors  $\{\overleftarrow{L}_w \mid w \in L\}$  et  $\{\overleftarrow{L}'_w \mid w \in L'\}$  sont finis. D'autre part, les ensembles  $\{\overleftarrow{L}_w \mid w \in L'\}$  sont en bijection avec les ensembles  $\{\overleftarrow{L}'_w \mid w \in L'\}$ ; plus précisément, le lien entre  $\overleftarrow{L}_w$  et  $\overleftarrow{L}'_w$  est donné par l'égalité suivante :

$$(5.1) \quad \overleftarrow{L}'_w = \begin{cases} (\llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^* \cup \{\varepsilon\}) \setminus \overleftarrow{L}_w & \text{si } w \in \llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^* \\ \cup \{\varepsilon\} & \\ (\llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^*) \setminus \overleftarrow{L}_w & \text{si } w \in 0 \llbracket 0, k-1 \rrbracket^* \end{cases}$$

Ainsi,  $\{\overleftarrow{L}_w \mid w \in L \cup L'\}$  est aussi fini. Comme  $L \cup L' = \llbracket 1, k-1 \rrbracket \llbracket 0, k-1 \rrbracket^* \cup \{\varepsilon\}$ , il reste à montrer la finitude de l'ensemble  $\{\overleftarrow{L}_w \mid w \in 0 \llbracket 0, k-1 \rrbracket^*\}$ . C'est là où la preuve diffère du cas de la lecture GD.

On sépare l'ensemble des mots  $w$  de  $0 \llbracket 0, k-1 \rrbracket^*$  en deux parties selon que  $1w$  appartient à  $L$  ou non.

On considère dans un premier temps l'ensemble  $\{v \in 0 \llbracket 0, k-1 \rrbracket^* \mid 1v \in L\}$ . Pour tout mot  $w$  de cet ensemble, le langage  $\overleftarrow{L}_{1w}$  est un élément de  $\overleftarrow{F}(E)$  et donc  $\overleftarrow{L}_w$  est à distance 1 de  $\overleftarrow{F}(E)$  dans l'automate  $\overleftarrow{\mathcal{M}}(E)$ . Comme  $\overleftarrow{\mathcal{M}}(E)$  est de degré fini, l'ensemble des langages  $\overleftarrow{L}_w$  où  $w$  parcourt  $\{v \in 0 \llbracket 0, k-1 \rrbracket^* \mid 1v \in L\}$  est donc fini.

Supposons maintenant que  $w$  soit un mot de l'ensemble  $\{v \in 0 \llbracket 0, k-1 \rrbracket^* \mid 1v \notin L\}$ . Puisque  $1w \notin L$ , on a  $1w \in L'$  et donc  $\overleftarrow{L}'_{1w}$  est un élément de  $\overleftarrow{F}(\overline{E})$ . Ainsi  $\overleftarrow{L}'_w$  est à distance 1 de  $\overleftarrow{F}(\overline{E})$  dans l'automate  $\overleftarrow{\mathcal{M}}(\overline{E})$ . Puisque  $\overleftarrow{\mathcal{M}}(\overline{E})$  est de degré fini, alors l'ensemble des langages  $\{\overleftarrow{L}'_v \mid \overleftarrow{L}'_{1v} \in \overleftarrow{F}(\overline{E})\}$  est fini. Ainsi, l'ensemble  $\{\overleftarrow{L}'_v \mid v \in 0 \llbracket 0, k-1 \rrbracket^*, 1v \notin L\}$  est fini et via l'égalité (5.1), l'ensemble  $\{\overleftarrow{L}_v \mid v \in 0 \llbracket 0, k-1 \rrbracket^*, 1v \notin L\}$  l'est aussi.

Ce dernier résultat permet de conclure que  $\overleftarrow{Q}(E) = \{\overleftarrow{L}_w \mid w \in \llbracket 0, k-1 \rrbracket^*\}$  est fini et par conséquent,  $\overleftarrow{\mathcal{M}}(E)$  est un automate fini. L'ensemble  $E$  est donc un ensemble  $k$ -reconnaissable.  $\square$

**Remarque.** Voici un exemple d'application de la proposition 5.3 : On considère l'ensemble  $E_1 = \{n \mid |\rho_2(n)|_1 = |\rho_2(n)|_0\}$ , ensemble des entiers

possédant autant de zéros que de uns dans leurs écritures binaires. L'ensemble  $E_1$  est  $2^\infty$ -reconnaisable mais pas  $2^\infty$ co-reconnaisable.

En effet, si cela était le cas,  $E_1$  serait en fait 2-reconnaisable d'après la proposition 5.3, or cela est exclu puisque  $\rho_2(E)$  est un langage algébrique non rationnel.

**Proposition 5.4.** *Soit un entier  $k \geq 2$ .*

*Les familles  $\overrightarrow{\text{Rec}}_\infty(k)$  et  $\overleftarrow{\text{Rec}}_\infty(k)$  sont stables par intersection.*

*En particulier, l'intersection d'un ensemble  $k$ -reconnaisable avec un ensemble  $k^\infty$ -reconnaisable  $GD$  (resp.  $DG$ ) est  $k^\infty$ -reconnaisable  $GD$  (resp.  $DG$ ).*

*Démonstration.* Soient deux ensembles d'entiers  $k^\infty$ -reconnaisables  $GD$  (resp.  $DG$ )  $N$  et  $N'$  d'automates minimaux  $\mathcal{M}(N) = (Q(N), \phi, q_0, F(N))$  et  $\mathcal{M}(N') = (Q(N'), \psi, q'_0, F(N'))$ .

Il suffit de construire, à partir des automates  $\mathcal{M}(N)$  et  $\mathcal{M}(N')$ , le  $k$ -automate produit  $\mathcal{A} = (Q, \delta, q''_0, F)$  qui reconnaît l'ensemble  $N \cap N'$  en posant :

- $Q = Q(N) \times Q(N')$ ,
- $\delta(i, (a, b)) = (\phi(i, a), \psi(i, b))$  pour tout  $(a, b) \in Q(N) \times Q(N')$  et tout  $i \in \llbracket 0, k - 1 \rrbracket$ ,
- $q''_0 = (q_0, q'_0)$ ,
- $F = F(N) \times F(N')$ .

La composante co-accessible depuis  $F$  dans l'automate  $\mathcal{A}$  est de degré fini car  $\mathcal{M}(N)$  et  $\mathcal{M}(N')$  possèdent cette propriété et  $F$  est fini puisque  $F(N)$  et  $F(N')$  le sont tout les deux. □

**Proposition 5.5.** *Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaisable  $GD$  (resp.  $DG$ ) et  $E'$  un ensemble fini d'entiers. Les ensembles d'entiers  $E \cup E'$  et  $E \setminus E'$  sont  $k^\infty$ -reconnaisables  $GD$  (resp.  $DG$ ).*

*Démonstration.* Soit  $E$  un ensemble d'entiers  $k^\infty$ -reconnaisable  $GD$  (resp.  $DG$ ) et  $E'$  un ensemble fini d'entiers.

Le fait que  $E \setminus E'$  soit  $k^\infty$ -reconnaisable  $GD$  (resp.  $DG$ ) découle directement de la proposition précédente puisque l'ensemble  $\mathbb{N} \setminus E'$  est  $k$ -reconnaisable donc en particulier  $k^\infty$ -reconnaisable.

Dans le sens  $GD$ , pour montrer que l'union de ces deux ensembles est  $k^\infty$ -reconnaisable  $GD$ , on va modifier un  $k$ -automate  $GD$   $\mathcal{A} = (Q, \phi, q_0, F)$  qui reconnaît  $E$  pour qu'il reconnaisse  $E \cup E'$ . Soit  $L = \max\{|\rho_k(n)|, n \in E'\}$ . On pose  $Q' = \{\overrightarrow{\phi}(w, q_0) \mid w \in \llbracket 0, k - 1 \rrbracket^*, |w| > L\}$  et  $Q'' = \{q'_w \mid w \in \llbracket 0, k - 1 \rrbracket^*, |w| \leq L\}$ .

On construit le  $k$ -automate  $GD$   $\mathcal{A}' = (Q' \cup Q'', \psi, q''_0, F')$  en posant

- $q''_0 = q'_\varepsilon$ ,
- $\psi$  est définie ainsi :
  - pour tout  $w \in \llbracket 0, k - 1 \rrbracket^*$  avec  $|w| < L$ ,  $\psi(i, q'_w) = q'_{wi}$ ,

- pour tout  $w \in \llbracket 0, k-1 \rrbracket^L$ ,  $\psi(i, q'_w) = \overrightarrow{\phi}(wi, q_0)$ ,
- si  $q \in Q'$ ,  $\psi(i, q) = \phi(i, q)$ ,
- $F' = (F \cap Q') \cup \{q'_{\rho_k(n)}, n \in E \cup E', |\rho_k(n)| \leq L\}$ ,

de sorte que  $\mathcal{A}'$  est un  $k$ -automate de degré fini qui reconnaît  $E \cup E'$  par lecture GD. L'ensemble  $F'$  est fini puisque  $F$  l'est.

De la même manière dans le sens de lecture DG, si  $\mathcal{A} = (Q, \phi, q_0, F)$  est un  $k$ -automate DG qui reconnaît  $E$ , on pose  $L = \max\{|\rho_k(n)|, n \in E'\}$ ,  $Q' = \{\overleftarrow{\phi}(w, q_0) \mid w \in \llbracket 0, k-1 \rrbracket^*, |w| > L\}$  et  $Q'' = \{q'_w \mid w \in \llbracket 0, k-1 \rrbracket^*, |w| \leq L\}$ . On construit le  $k$ -automate DG  $\mathcal{B} = (Q' \cup Q'', \delta, q''_0, F')$ , avec  $q''_0$  et  $F'$  définis comme ci-dessus et  $\delta$  définie par :

- si  $q \in Q'$ ,  $\delta(i, q) = \phi(i, q)$ ,
- pour tout  $w \in \llbracket 0, k-1 \rrbracket^*$  avec  $|w| < L$ ,  $\delta(i, q'_w) = q'_{iw}$ ,
- pour tout  $w \in \llbracket 0, k-1 \rrbracket^L$ ,  $\delta(i, q'_w) = \overleftarrow{\phi}(iw, q_0)$ ,

de sorte que  $\mathcal{B}$  est un  $k$ -automate de degré fini qui reconnaît  $E \cup E'$  par lecture DG. L'ensemble  $F'$  est fini puisque  $F$  l'est. □

**Proposition 5.6.** *Soit un entier  $k \geq 2$ . Les familles  $\overleftarrow{\text{Rec}}_\infty(k)$  et  $\overrightarrow{\text{Rec}}_\infty(k)$  ne sont pas stables par union.*

*Démonstration.* Pour un entier  $k \geq 2$  fixé, nous allons expliciter deux ensembles  $k^\infty$ -reconnaissables dont l'union ne l'est pas.

Soit  $N = \{k^{2i} + k^i + 1 \mid i \geq 1\}$  et  $N' = \{k^i + 1 \mid i \geq 1\}$ . L'ensemble  $N$  est dans  $\overrightarrow{\text{Rec}}_\infty(k) \cap \overleftarrow{\text{Rec}}_\infty(k)$  et l'ensemble  $N'$  est  $k$ -reconnaissable. Cependant,  $N'' = N \cup N'$  n'est ni dans  $\overrightarrow{\text{Rec}}_\infty(k) \cup \overleftarrow{\text{Rec}}_\infty(k)$  ni dans  $\text{coRec}_\infty(k) \cup \overleftarrow{\text{coRec}}_\infty(k)$ .

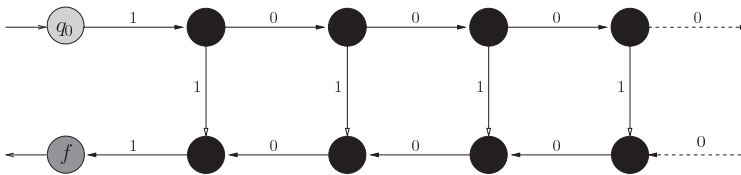


FIGURE 5.3. Graphe du  $k$ -automate reconnaissant  $N$ .

Un 2-automate  $\mathcal{A} = (Q_{\mathcal{A}}, \phi, q_0, \{f\})$  qui reconnaît  $N$  dans les deux sens de lecture est représenté Figure 5.3 (seuls les états co-accessibles y sont représentés). Le  $k$ -automate fini  $\mathcal{B} = (Q_{\mathcal{B}}, \psi, q'_0, \{f'\})$  qui reconnaît  $N'$  est quant à lui présenté Figure 5.4 (seuls les états co-accessibles y sont représentés).

Soit  $\overline{\mathcal{M}}(N'') = (\overline{Q}(N''), \overline{\phi}_{N''}, \overline{q}_{0, N''}, \overline{F}(N''))$  l'automate minimal GD associé à  $N''$  et soit  $L = \rho_2(N'')$ . On a  $\overline{F}(N'') = \{\overrightarrow{L}_w \mid w \in L\}$ . Or, pour tout

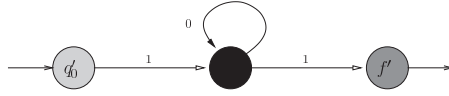


FIGURE 5.4. Graphe du  $k$ -automate reconnaissant  $N'$ .

$i \geq 0$ ,  $\overrightarrow{L_{10^{i+1}}} = \{\varepsilon, 0^i 1\}$ , donc  $\overrightarrow{F(N'')}$  est infini. La proposition 5.1 permet donc de conclure que  $N''$  n'est pas  $k^\infty$ -reconnaissable GD.

Dans le sens DG, on utilise exactement le même raisonnement : si on appelle  $\overleftarrow{\mathcal{M}(N'')} = (\overleftarrow{Q(N'')}, \overleftarrow{\phi_{N''}}, \overleftarrow{q_{0,N''}}, \overleftarrow{F(N'')})$  l'automate minimal DG associé à  $N''$ , on a  $\overleftarrow{F(N'')} = \{\overleftarrow{L_w} \mid w \in L\}$ . Or, pour tout  $i \geq 0$ ,  $\overleftarrow{L_{10^{i+1}}} = \{\varepsilon, 10^i\}$ , donc  $\overleftarrow{F(N'')}$  est infini. On conclut de la même manière que  $N''$  n'est pas  $k^\infty$ -reconnaissable.

Remarquons aussi que l'ensemble  $N''$  n'est pas non plus dans l'ensemble  $\overrightarrow{\text{coRec}}_\infty(k) \cup \overleftarrow{\text{coRec}}_\infty(k)$ .

En effet, par l'absurde, supposons que  $\mathbb{N} \setminus N''$  soit  $k^\infty$ -reconnaissable DG et notons  $\mathcal{M} = (Q, \delta, q''_0, F)$  son  $k$ -automate minimal DG. Puisqu'aucune écriture en base  $k$  des éléments de  $N''$  ne contient le motif 1111, alors, pour tout état  $q$  de  $Q$ , l'état  $\overrightarrow{\delta}(1111, q)$  est un état de  $F$ . Ainsi, tout état est à distance au plus 4 de l'ensemble terminal. Puisque  $F$  est fini et que  $\mathcal{M}$  est de degré fini, alors  $S$  est en fait fini et  $\mathcal{M}$  est un automate fini. Ainsi, les ensembles  $\mathbb{N} \setminus N''$  et  $N''$  sont  $k$ -reconnaissables, ce qui est absurde puisque  $N''$  n'est même pas  $k^\infty$ -reconnaissable.

On utilise le même argument pour le sens de lecture GD. □

**Proposition 5.7.** *Soient deux entiers  $k \geq 2$  et  $l \geq 2$ , on a les relations suivantes :*

$$\overleftarrow{\text{Rec}}_\infty(k) = \overleftarrow{\text{Rec}}_\infty(k^l) \quad \text{et} \quad \overrightarrow{\text{Rec}}_\infty(k) \subsetneq \overrightarrow{\text{Rec}}_\infty(k^l),$$

*Démonstration. Première partie :  $\overleftarrow{\text{Rec}}_\infty(k) = \overleftarrow{\text{Rec}}_\infty(k^l)$ .*

Montrons dans un premier temps que  $\overleftarrow{\text{Rec}}_\infty(k) \subset \overleftarrow{\text{Rec}}_\infty(k^l)$ . Soit  $E \in \overleftarrow{\text{Rec}}_\infty(k)$  et  $\mathcal{A} = (Q, \phi, q_0, F)$  un  $k$ -automate fini ou dénombrable DG qui reconnaît  $E$ . Intuitivement, la première idée pour construire un  $k^l$ -automate DG  $\mathcal{A}' = (Q', \psi, q'_0, F')$  qui reconnaît  $E$  à partir de  $\mathcal{A}$  est de poser  $Q' = Q$  et  $\psi(i, q) = \overleftarrow{\phi}(0^{l-|\rho_k(i)|} \rho_k(i), q)$ . Cependant, cela implique que l'ensemble final doit être composé des  $\overleftarrow{\phi}(0^j, f)$  pour tout  $0 \leq j \leq l-1$  et  $f \in F$ . Cette construction peut poser des problèmes si par exemple, il existe un entier  $n$  n'étant pas dans  $E$  pour lequel  $\overleftarrow{\phi}(\rho_k(n), q_0) = \overleftarrow{\phi}(0^i, f)$ .

Pour pallier à cela, on rajoute dans l'ensemble  $Q'$ , en plus de  $Q$ ,  $l-1$  copies de  $F$  formant avec  $F$  lui-même l'ensemble terminal  $F'$ , dans lesquelles

on envoie les transitions  $\psi(i, s)$  lorsque  $\overleftarrow{\phi}(\rho_k(i), s) \in F$  selon le nombre  $l - |\rho_k(i)|$  de zéros à lire avant la nouvelle plage de  $l$  chiffres à lire. Cela permet de garder en mémoire le nombre de zéros à rajouter dans l'écriture en base  $k$ , si une autre lettre en base  $k^l$  doit être lue.

Plus formellement, on construit un  $k^l$ -automate DG  $\mathcal{A}' = (Q', \psi, q'_0, F')$  qui reconnaît  $E$  comme suit :

- $F' = \{f_{(q,j)} \mid q \in F, j \in \llbracket 0, l-1 \rrbracket\}$  où  $f_{(q,0)} = q$  et  $f_{(q,j)} \notin Q$  pour tout  $j \neq 0$ ,
- $Q' = Q \cup F'$ ,
- $q'_0 = q_0$ ,
- si on pose  $q' = \overleftarrow{\phi}(\rho_k(i), q)$  pour tout  $i \in \llbracket 0, k^l - 1 \rrbracket$  et tout  $q \in Q$ , l'application  $\psi$  est définie sur  $\llbracket 0, k^l - 1 \rrbracket \times Q$  par :
  - si  $i \neq 0$  et  $q' \notin F$ ,  $\psi(i, q) = \overleftarrow{\phi}(0^{l-|\rho_k(i)|}\rho_k(i), q)$ ,
  - si  $i \neq 0$  et  $q' \in F$ ,  $\psi(i, q) = f_{(q', l-|\rho_k(i)|)}$ ,
  - $\psi(0, q) = \overleftarrow{\phi}(0^l, q)$ ,
  - si  $i \in \llbracket 0, k^l - 1 \rrbracket$  et  $(q, j) \in F \times \llbracket 1, l-1 \rrbracket$ ,  $\psi(i, f_{(q,j)}) = \psi(i, \overleftarrow{\phi}(0^j, q))$ .

Il est facile de voir que l'automate  $\mathcal{A}'$  reconnaît bien  $\rho_k(E)$ , que l'ensemble final  $F'$  de  $\mathcal{A}'$  est fini si  $F$  l'est et que le  $k^l$ -automate  $\mathcal{A}'$  est de degré fini puisque  $\mathcal{A}$  est de degré fini, ainsi :  $\overrightarrow{\text{Rec}}_\infty(k) \subset \overrightarrow{\text{Rec}}_\infty(k^l)$ .

Pour montrer l'inclusion réciproque, on considère un ensemble  $E$  de  $\overleftarrow{\text{Rec}}_\infty(k^l)$  reconnu par un  $k^l$ -automate DG  $\mathcal{A} = (Q, \phi, q_0, F)$ . On construit le  $k$ -automate DG  $\mathcal{A}' = (Q', \psi, q'_0, F')$  défini à partir de  $\mathcal{A}$  comme suit :

- $Q' = Q \times L$ , où  $L = \{w \in \llbracket 0, k-1 \rrbracket^* \mid |w| \leq l\}$ ,
- $q'_0 = (q_0, \varepsilon)$ ,
- l'application  $\psi$  est définie pour tout triplet  $(i, q, w)$  de  $\llbracket 0, k-1 \rrbracket \times Q \times L$  par :
  - $\psi(i, (q, w)) = (q, iw)$  si  $|w| \leq l-1$ ,
  - $\psi(i, (q, w)) = (\phi(\pi_{k^l}(w), q), i)$  si  $|w| = l$ .
- l'ensemble d'acceptation est donné par :

$$F' = (F \times \{\varepsilon\}) \cup \{(q, w) \mid q \in Q, w \in L \setminus 0\llbracket 0, k-1 \rrbracket^*, \phi(\pi_{k^l}(w), q) \in F\}.$$

On vérifie aisément que  $\mathcal{A}'$  reconnaît  $E$  et que  $F'$  est fini puisque  $\mathcal{A}$  est de degré fini et  $F$  est fini. De plus, l'automate  $\mathcal{A}'$  est lui aussi de degré fini car le degré entrant d'un état  $(q, w)$  de  $Q'$  est égal à 0 si  $w = \varepsilon$ , au degré entrant de  $q$  dans  $\mathcal{A}$  si  $|w| = 1$  et à 1 si  $|w| \geq 2$ . Par conséquent, on a également l'inclusion  $\overleftarrow{\text{Rec}}_\infty(k^l) \subset \overleftarrow{\text{Rec}}_\infty(k)$ , ce qui termine la première partie de la preuve.  $\square$

*Démonstration. Deuxième partie :*  $\overrightarrow{\text{Rec}}_\infty(k) \subsetneq \overrightarrow{\text{Rec}}_\infty(k^l)$ .

Montrons que  $\overrightarrow{\text{Rec}}_\infty(k) \subset \overrightarrow{\text{Rec}}_\infty(k^l)$ . Soit  $E \in \overrightarrow{\text{Rec}}_\infty(k)$  et  $\mathcal{A} = (Q, \phi, q_0, F)$  un  $k$ -automate dénombrable GD qui reconnaît  $E$ . On construit le  $k^l$ -automate GD  $\mathcal{A}' = (Q', \psi, q'_0, F')$  défini comme suit :

- $Q' = Q \cup \{q'_0\}$ , où  $q'_0$  est un état n'appartenant pas à  $Q$ ,
- l'application  $\psi$  est définie pour tout  $i \in \llbracket 0, k^l - 1 \rrbracket$  et  $q \in Q$ , par :
  - $\psi(i, q'_0) = \overrightarrow{\phi}(\rho_k(i), q_0)$  si  $i \neq 0$ ,
  - $\psi(0, q'_0) = \overrightarrow{\phi}(0^l, q_0)$ ,
  - $\psi(i, q) = \overrightarrow{\phi}(0^{l-|\rho_k(i)|} \rho_k(i), q)$  pour  $q \neq q'_0$ .
- $F' = F \cup \{q'_0\}$  si  $q_0 \in F$  ou  $F' = F$  sinon.

Cet automate reconnaît bien  $E$ . Il suffit pour cela de remarquer que si  $\rho_{k^l}(n) = w_0 w_1 \cdots w_p$ , alors

$$\rho_k(n) = \rho_k(w_0) 0^{l-|\rho_k(w_1)|} \rho_k(w_1) \cdots 0^{l-|\rho_k(w_p)|} \rho_k(w_p).$$

Ainsi, si l'entier  $n$  tel que  $\rho_{k^l}(n) = w_0 w_1 \cdots w_p$  est reconnu par  $\mathcal{A}'$  en base  $k^l$ , alors  $\rho_k(n)$  est reconnu par  $\mathcal{A}$  car :

$$\overrightarrow{\phi}(\rho_k(w_0) 0^{l-|\rho_k(w_1)|} \rho_k(w_1) \cdots 0^{l-|\rho_k(w_p)|} \rho_k(w_p), q_0) = \overrightarrow{\psi}(w_0 w_1 \cdots w_p, q'_0).$$

De même, si  $\rho_k(n) = w_0 w_1 \cdots w_{pl+j}$ , alors

$$\rho_{k^l}(n) = \pi_k(w_0 \cdots w_j) \pi_k(w_{j+1} \cdots w_{l+j}) \cdots \pi_k(w_{(p-1)l+j+1} \cdots w_{pl+j}),$$

donc si l'entier  $n$  tel que  $\rho_{k^l}(n) = w_0 w_1 \cdots w_{pl+j}$  est reconnu par  $\mathcal{A}$  en base  $k$ , alors  $\rho_{k^l}(n)$  est reconnu par  $\mathcal{A}'$  car :

$$\begin{aligned} \overrightarrow{\psi}(\pi_k(w_0 \cdots w_j) \pi_k(w_{j+1} \cdots w_{l+j}) \cdots \pi_k(w_{(p-1)l+j+1} \cdots w_{pl+j}), q'_0) \\ = \overrightarrow{\phi}(w_0 w_1 \cdots w_{pl+j}, q_0). \end{aligned}$$

D'autre part, puisque le degré de  $\mathcal{A}$  est fini, celui de  $\mathcal{A}'$  l'est aussi : pour chaque état de  $Q' = Q \cup \{q'_0\}$ , le degré entrant de  $\mathcal{A}'$  en  $q' \in Q'$  est fini. En effet,  $q'_0$  est de degré entrant nul dans l'automate  $\mathcal{A}'$  et pour tout  $q$  de  $Q$ , le degré entrant dans l'automate  $\mathcal{A}'$  est borné par le cardinal de l'ensemble  $\{q'' \in Q \mid \exists w \in \llbracket 0, k - 1 \rrbracket^l, \overrightarrow{\phi}(w, q'') = q\}$  (auquel il faut éventuellement rajouter le nombre de flèches venant de  $q'_0$ ). L'ensemble  $F'$  étant fini puisque  $F$  l'est, on obtient donc  $\overrightarrow{\text{Rec}}_\infty(k) \subset \overrightarrow{\text{Rec}}_\infty(k^l)$ .

Montrons maintenant que cette inclusion est stricte. Pour cela, nous exhibons un exemple d'ensemble  $(k^l)^\infty$ -reconnaissable qui n'est ni  $k^\infty$ -reconnaissable ni  $k^\infty$ -co-reconnaissable.

Soit l'ensemble  $E = \{k^{2li} + k^{li} \mid i \geq 1\} \cup \{k^{li+1} \mid i \geq 1\}$ . Cet ensemble est dans  $\overrightarrow{\text{Rec}}_\infty(k^l)$ , reconnu par l'automate présenté à la figure 5.5 (les états poubelles sont omis) et son complémentaire  $\mathbb{N} \setminus E$  est un élément de  $\overrightarrow{\text{coRec}}_\infty(k^l)$ . Supposons que  $E$  soit également dans  $\overrightarrow{\text{Rec}}_\infty(k)$ , reconnu par un  $k$ -automate  $\mathcal{A} = (Q, \phi, q_0, F)$ . Il existe alors deux entiers  $i \neq j$  tels que la

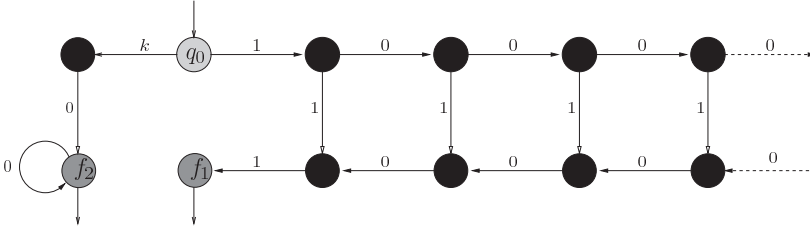


FIGURE 5.5. Graphe du  $k^l$ -automate reconnaissant  $E$ .

lecture GD de  $\rho_k(k^{li+1}) = 10^{li+1}$  et celle de  $\rho_k(k^{lj+1}) = 10^{lj+1}$  conduisent l'automate  $\mathcal{A}$  au même état final :  $\vec{\phi}(\rho_k(k^{li+1}), q_0) = \vec{\phi}(\rho_k(k^{lj+1}), q_0)$ . Ainsi,

$$\vec{\phi}(\rho_k(k^{li+1})0^{l-2}10^{l(i+1)}, q_0) = \vec{\phi}(\rho_k(k^{lj+1})0^{l-2}10^{l(i+1)}, q_0),$$

ce qui se traduit par le fait que la lecture GD de  $10^{l(i+1)-1}10^{l(i+1)}$  et celle de  $10^{l(j+1)-1}10^{l(i+1)}$  conduisent l'automate  $\mathcal{A}$  au même état. Or  $x = \pi_k(10^{l(i+1)-1}10^{l(i+1)}) = k^{2l(i+1)} + k^{l(i+1)}$  est un élément de  $E$  et  $y = \pi_k(10^{l(j+1)-1}10^{l(i+1)}) = k^{l(i+j+2)} + k^{l(i+1)}$  n'est pas un élément de  $E$ , ce qui est contradictoire.

Si  $E \in \overrightarrow{\text{coRec}}_\infty(k)$ , son complémentaire est reconnu par un  $k$ -automate  $\mathcal{A}' = (Q', \psi, q'_0, F')$  avec  $F'$  fini. On utilise le même raisonnement pour aboutir à une contradiction. Puisque pour tout entier  $i \geq 1$ ,  $k^{li}$  n'appartient pas à  $E$ , il existe deux entiers  $i \neq j$  tels que la lecture GD de  $\rho_k(k^{li}) = 10^{li}$  et celle de  $\rho_k(k^{lj}) = 10^{lj}$  conduisent l'automate  $\mathcal{A}'$  au même état final :  $\vec{\psi}(\rho_k(k^{li}), q'_0) = \vec{\psi}(\rho_k(k^{lj}), q'_0)$ . De ce fait,

$$\vec{\psi}(\rho_k(k^{li})0^{l-1}10^{l(i+1)}, q'_0) = \vec{\psi}(\rho_k(k^{lj})0^{l-1}10^{l(i+1)}, q'_0),$$

et on conclut à une contradiction comme ci-dessus. Nous avons donc construit un ensemble d'entiers  $E \in \overrightarrow{\text{Rec}}_\infty(k^l) \setminus (\overrightarrow{\text{Rec}}_\infty(k) \cup \overrightarrow{\text{coRec}}_\infty(k))$  ce qui termine la preuve.  $\square$

### 6. Ouvertures

A l'issue de cette étude, on peut se poser la question de la reconnaissabilité d'autres familles d'entiers usuelles comme par exemple la famille des carrés parfaits, question déjà posée dans [20, Problème 1, Question 1] et qui fait suite aux travaux concernant la reconnaissabilité de l'ensemble des carrés dans le cadre des automates finis [24, 21]. Ce type de question est également à rapprocher de la question d'une éventuelle extension du théorème de Cobham sur la dépendance à la base [9] : un ensemble d'entiers

$k^\infty$ -reconnaissable et  $l^\infty$ -reconnaissable pour deux bases  $k$  et  $l$  multiplicativement indépendantes est-il  $p^\infty$ -reconnaissable dans toute base  $p \geq 2$ ? Cette question précise la question posée dans [20, Problème 2]. Un autre angle d'attaque de cette question consisterait à se demander s'il existe des ensembles d'entiers autres que les unions finies de progressions arithmétiques qui sont  $p^\infty$ -reconnaissables pour toute base  $p \geq 2$ .

On peut également envisager l'étude des propriétés de stabilité de ces familles d'ensembles d'entiers par d'autres opérations simples, comme par exemple l'addition d'un entier ou la multiplication par une constante.

Une autre piste serait d'étudier les propriétés diophantiennes des réels engendrés par automates dénombrables, dans le but d'étendre les résultats de B. Adamczewski et J. Cassaigne [1]. Cela nécessite sans doute une restriction du cadre de travail, notamment en ce qui concerne la géométrie des automates dénombrables. Il semble raisonnable de se restreindre pour cette question à des automates "structurés", comme par exemple les automates dont les graphes sont engendrés par des grammaires finies. Cela permettrait non seulement de toucher une large famille d'ensembles d'entiers liés à des langages algébriques, étudiés particulièrement dans [7, 8], mais aussi de fournir un cadre suffisamment rigide pour espérer des résultats conséquents dans cette voie.

## Bibliographie

- [1] B. ADAMCZEWSKI ET J. CASSAIGNE, *Diophantine properties of real numbers generated by finite automata*. *Compositio Math.* **142** (2006), 1351–1372.
- [2] J.-P. ALLOUCHE ET J. SHALLIT, *Automatic sequences. Theory, applications, generalizations*. Cambridge University Press, 2003.
- [3] J.-P. ALLOUCHE ET J. SHALLIT, *The ring of  $k$ -regular sequences*. *Theor. Comp. Sci.* **98** (1992), n°2, 163–197.
- [4] J.-P. ALLOUCHE ET J. SHALLIT, *The ring of  $k$ -regular sequences, II*. *Theor. Comp. Sci.* **307** (2003), 3–29.
- [5] J.-M. AUTEBERT, *Langages algébriques*. Etudes et recherche en informatique, Masson, 1987.
- [6] J.-M. AUTEBERT, J. BERSTEL ET L. BOASSON, *Context-free languages and pushdown automata*. In *Handbook of formal languages*, Vol. 1, Springer (1997), 111–174.
- [7] J. BERSTEL, *On Sets of Numbers Recognized by Push-Down Automata*. IEEE Conference record of 13th Annual Symposium of Foundations of Computer Sciences (FOCS'72) (1972), 200–206.
- [8] J. BERSTEL, *Sur la densité asymptotique de langages formels*. *Proceedings of International Colloquium on Automata, Languages and Programming (ICALP'72)* (1973), 345–358.
- [9] A. COBHAM, *On the base-dependence of sets of numbers recognizable by finite automata*. *Math. Syst. Theory* **3** (1969), 186–192.
- [10] A. COBHAM, *Uniform-tag sequences*. *Math. Syst. Theory* **6** (1972), 164–192.
- [11] D. CAUCAL, *On the regular structure of prefix rewriting*. *Theor. Comp. Sci.* **106** (1992), 61–86.
- [12] D. CAUCAL, *On infinite transition graphs having a decidable monadic theory*. *Theor. Comp. Sci.* **290** (2003), 79–115.
- [13] S. EILENBERG, *Automata, languages and machines, Vol. A*. Academic Press, 1974.



- [14] S. FERENCZI, *Substitutions on an infinite alphabet*. Ann. Inst. Fourier **56** n°7 (2006), 2315–2343.
- [15] C. FOUVRY ET C. MAUDUIT, *Sur les entiers dont la somme des chiffres est moyenne*. J. Number Theory **114** n°1 (2005), 135–152.
- [16] C. FROUGNY ET B. SOLOMYAK, *On the context-freeness of the  $\theta$ -expansions of the integers*. Internat. J. Algebra Comput. **9** n°3-4 (1999), 347–350.
- [17] J. HARTMANIS ET H. SHANK, *On the recognition of primes by automata*. J. Assoc. Comput. Mach. **15** (1968), 382–389.
- [18] M. LE GONIDEC, *Sur la complexité des mots infinis engendrés par des  $q$ -automates dénombrables*. Ann. Inst. Fourier **56** n°7 (2006), 2463–2491.
- [19] M. LE GONIDEC, *Drunken man infinite words complexity*. RAIRO-Theor. Inf. Appl. **42** (2008), 599–613.
- [20] C. MAUDUIT, *Propriétés arithmétiques des substitutions et automates infinis*. Ann. Inst. Fourier **56** n°7 (2006), 2525–2549.
- [21] M. MINSKY ET S. PAPERT, *Unrecognizable sets of numbers*. J. Assoc. Comput. Mach. **13** (1966), 281–286.
- [22] D. MULLER ET P. SHUPP, *The theory of ends, pushdown automata and second-order logic*. Theor. Comp. Sci. **37** (1985), 51–75.
- [23] M. RIGO, *Generalization of automatic sequences for numeration systems on a regular language*. Theor. Comp. Sci. **244** (2000), 271–281.
- [24] R. W. RITCHIE, *Finite automata and the set of squares*. J. Assoc. Comput. Mach. **10** (1963), 528–531.
- [25] J. SAKAROVITCH, *Éléments de théorie des automates*. Vuibert, 2003.
- [26] M.-P. SCHÜTZENBERGER, *A remark on acceptable sets of numbers*. J. Assoc. Comput. Mach. **15** (1968), 300–303.

Julien CASSAIGNE  
Institut de mathématiques de Luminy  
case 907  
13288 Marseille Cedex 9, France  
*E-mail*: [cassaigne@iml.univ-mrs.fr](mailto:cassaigne@iml.univ-mrs.fr)  
*URL*: <http://iml.univ-mrs.fr/~cassaign/>

Marion LE GONIDEC  
Université du Sud - Toulon - Var  
Avenue de l'Université - BP20132  
83957 La Garde Cedex, France  
*E-mail*: [marion.le-gonidec@univ-tln.fr](mailto:marion.le-gonidec@univ-tln.fr)  
*URL*: <http://legonidec.univ-tln.fr/>