Académie Universitaire Wallonie-Europe

Université de Liège

Faculté des Sciences

Département de Mathématiques

Problème de décision pour les ensembles ultimement périodiques dans un système de numération non-standard

Mémoire présenté par
Elise VANDOMME
en vue de l'obtention du grade
de Master en Sciences Mathématiques

Année académique 2009-2010

Je remercie sincèrement mon promoteur, Michel Rigo,
pour l'attention dont il a fait preuve.

Je tiens aussi à remercier Émilie Charlier pour son aide
et Marie Ernst pour ses conseils avisés.

Enfin, merci à ma famille qui m'a soutenue tout au long de ce travail .

Table des matières

Introduction

A l'heure actuelle, l'homme utilise constamment les nombres entiers. Pour les manipuler facilement, il les représente sous forme de mots de la manière suivante. Dans un système de numération défini par la suite $(U_i)_{i\geq 0}$, l'entier n est représenté par la suite de lettre $w_1 \dots w_\ell$ si $n = \sum_{i=1}^{\ell} U_i w_i$. Par exemple, le système de numération en base 2 est défini par la suite $(2^i)_{i\geq 0}$. Nous pouvons voir un système de numération comme une bijection $\text{rep} : \mathbb{N} \to L$ entre $\mathbb{N}$ et un langage L. Chaque partie X de $\mathbb{N}$ est alors envoyé sur un sous-ensemble de $\text{rep}(L)$.

Il est naturel de s'intéresser aux parties de $\mathbb{N}$ qui correspondent à des sous-langages simples, i.e., les langages réguliers. Une partie de $\mathbb{N}$ qui envoyé sur un langage régulier est dit reconnaissable. Pour tout système de numération en base entière, il est bien connu que toute union finie de progressions arithmétiques est reconnaissable. En 1969, A. Cobham a prouvé la réciproque du résultat précédent [8], i.e., dans tout système de numération en base entière, les seuls ensembles d'entier reconnaissables sont les unions finies de progressions arithmétiques. Ce résultat est connu sous le nom de théorème de Cobham. Rappelons, avant de l'énoncer, que deux entiers sont multiplicativement indépendants si, pour tous entiers m et n, nous avons $p^m \neq q^n$.

Théorème de Cobham. Soit deux entiers $p, q \geq 2$ multiplicativement indépendants. Un ensemble X d'entiers positifs est à la fois U_p-reconnaissable et U_q-reconnaissable si et seulement si il est ultimement périodique.

Une preuve de ce résultat réalisée par A. Muchnik peut être trouvée dans [6]. Donc, les ensembles ultimement périodiques, i.e., les unions finies de progressions arithmétiques, jouent un rôle important. De plus, ces ensembles infinis sont codés par un nombre fini d'informations. Vu le théorème de

Cobham, les ensembles ultimement périodiques sont appelés reconnaissables dans la littérature. Cela nous conduit à nous intéresser au problème de décision suivant.

Problème. Étant donné un système de numération linéaire U tel que $\mathrm{rep}_U(\mathbb{N})$ est U-reconnaissable et un ensemble X d'entiers positifs donné par un automate acceptant $\mathrm{rep}_U(X)$, pouvons-nous décider si X est ou non ultimement périodique ?

Nous allons montrer que le problème est décidable pour une classe de systèmes de numération linéaires. De même l'extension du problème à une classe de systèmes de numération abstraits est décidable. Le sujet est développé selon la même structure que l'article *A decision problem for ultimately periodic sets in non-standard numeration systems* de J. Bell, E. Charlier, A. S. Fraenkel et M. Rigo [3], qui a servi de base à ce travail.

Le premier chapitre est consacré au rappel des notions de base concernant les langages, automates et systèmes de numération.

Le but du deuxième chapitre est de donner une réponse au problème de décision sous certaines hypothèses supplémentaires. Pour ce faire, nous donnons tout d'abord une borne supérieure des périodes admissibles pour un ensemble X U-reconnaissable lorsque nous supposons qu'il est ultimement périodique. Ensuite, nous donnons une borne supérieure des prépériodes admissibles. Remarquons que ces bornes dépendent du nombre d'états de l'automate minimal reconnaissant le langage $\mathrm{rep}_U(X)$. Puisqu'il y a un nombre fini de périodes et prépériodes admissibles, nous pouvons construire pour chaque paire de période et prépériode admissibles un automate acceptant l'ensemble ultimement périodique correspondant. Il ne reste plus qu'à comparer ces langages obtenus avec $\mathrm{rep}_U(X)$.

Cette méthode se base sur la quantité $N_U(m)$ qui est le nombre de valeurs distinctes apparaissant infiniment souvent dans la suite $(U_i \bmod m)_{i\geq 0}$. Le résultat principal de ce chapitre 2 est le théorème suivant.

Théorème. Soit $U = (U_i)_{i\geq 0}$ un système de numération linéaire tel que $\mathbb{N}$ est U-reconnaissable satisfaisant la condition $\lim_{i\to+\infty} U_{i+1} - U_i = +\infty$. Supposons que $\lim_{m\to+\infty} N_U(m) = +\infty$. Alors, nous pouvons décider si un ensemble U-reconnaissable est ou non ultimement périodique.

Dans le troisième chapitre, nous nous intéressons aux systèmes qui vérifient l'hypothèse $\lim_{m\to+\infty} N_U(m) = +\infty$. Pour caractériser ces systèmes, nous avons recours à des méthodes d'analyse p-adiques, d'algèbre et quelques résultats sur les suites linéaires récurrentes. Cela nous conduit à étudier la suite $(U_i \bmod p^v)_{i\geq 0}$ pour tout $v \geq 1$, où p est un nombre premier divisant a_k.

Dans le quatrième et dernier chapitre, nous étendons le problème de décision aux systèmes de numération abstraits. Dans ce cadre, nous appliquons le même type de méthodes à une classe de systèmes de numération abstraits. La procédure de décision correspondante est donnée par le théorème 4.4.1.

Enfin, nous observons que, pour tous les systèmes de numération abstraits, le problème de décision est équivalent au problème de périodicité HD0L. Il s'énonce comme suit : "Étant donné un morphisme f prolongeable sur une lettre a et un codage g, pouvons-nous décider si le mot infini $g(f\omega(a))$ est ou non ultimement périodique. Le théorème 4.4.1 donne une procédure de décision pour des exemples spécifiques du problème de périodicité.

Chapitre 1

Notions de base

1.1 Langages et automates

Rappelons quelques notions de la théorie des automates et des langages, cf. [24] pour une vue plus complète de ce sujet.

Définition 1.1.1. Un *alphabet* Σ est un ensemble fini de symboles, généralement appelés *lettres*. Un *mot* sur Σ est une suite finie de lettres de Σ. La *longueur* d'un mot w, notée $|w|$, est le nombre de lettres constituant w. L'unique mot de longueur 0 est le mot correspondant à la suite vide. Nous l'appelons le *mot vide* et il est noté ε. L'ensemble des mots finis sur Σ est noté Σ^*. Un *langage* sur Σ est une partie de Σ^*.

Un *mot infini* x sur Σ est une suite infinie de lettres $x_0x_1x_2\cdots$ ayant comme indices les entiers positifs. L'ensemble des mots infinis sur Σ est noté $\Sigma^{\mathbb{N}}$.

Définition 1.1.2. Pour un mot w sur Σ, nous définissons le *miroir* de w, noté w^R, par récurrence sur la longueur de w. Si $|w| = 0$, alors $w = \varepsilon$ et $w^R = \varepsilon$. Sinon $|w| > 0$ et nous pouvons écrire $w = \sigma u$ avec $\sigma \in \Sigma$ et $u \in \Sigma^*$. Dans ce cas, $w^R = u^R\sigma$. Si L est un langage sur Σ, alors le *miroir* de L est le langage $L^R = \{w^R \mid w \in L\}$.

Définition 1.1.3. Un *automate fini déterministe* est la donnée du quintuple $\mathcal{A} = (Q, q_0, F, \Sigma, \delta)$ où Q est un ensemble fini dont les éléments sont les états de $\mathcal{A}$, $q_0 \in Q$ est l'état initial, $F \subseteq Q$ est l'ensemble des états finals, Σ est l'alphabet de l'automate et $\delta : Q \times \Sigma \to Q$ est la fonction de transition de $\mathcal{A}$. Nous étendons la fonction de transition à $Q \times \Sigma^*$ de la manière suivante : pour tout $q \in Q$,

$$\delta(q, \varepsilon) = q$$

et

$$\delta(q, \sigma w) = \delta(\delta(q, \sigma), w), \quad \forall \sigma \in \Sigma, w \in \Sigma^*.$$

Nous représentons les états de l'automate par des cercles. L'état initial est désigné par une flèche entrante sans label et les états finaux sont repérés grâce à un double cercle. Une transition $\delta(q, \sigma) = q'$, avec $\sigma \in \Sigma$ et $q, q' \in Q$, est représentée par un arc orienté de q vers q' et de label σ.

Le langage $L \subseteq \Sigma^*$ est *accepté* par l'automate $\mathcal{A}$ si

$$L = \{w \in \Sigma^* \mid \delta(q_0, w) \in F\}.$$

Dans ce cas, nous disons que le langage L est *régulier*.

Le résultat suivant est souvent utile pour vérifier si un langage n'est pas régulier.

Lemme 1.1.4 (Lemme de la pompe)**.** *Soit $L \subseteq \Sigma^*$ un langage régulier. Il existe un entier ℓ tel que pour tout mot satisfaisant $|w| \ell$, il existe $x, y, z \in \Sigma^*$ tels que $w = xyz$, avec $|xy| \leq \ell$, $y \neq \varepsilon$, et $xy^* \subset L$.*

La version suivante du lemme de la pompe fournit une condition nécessaire et suffisante.

Lemme 1.1.5 (Lemme de la pompe,version forte)**.** *Un langage $L \subseteq \Sigma^*$ est régulier si et seulement si il existe un entier ℓ tel que pour tout mot $w \in L$ satisfaisant $|w| \geq \ell$, il existe $x, y, z \in \Sigma^*$ tels que $w = xyz$, $y \neq \varepsilon$ et pour tout $v \in \Sigma^*$ et tout $i \geq 0$, nous avons*

$$wv \in L \Leftrightarrow xy^i zv \in L.$$

Définition 1.1.6. Soient $L \subseteq \Sigma^*$ un langage sur l'alphabet fini Σ et x un mot fini sur Σ. Posons

$$x^{-1}.L = \{z \in \Sigma^* \mid xz \in L\}.$$

Nous pouvons maintenant définir la *congruence de Myhill-Nérode* $\sim_L$ comme suit. Soient $x, y \in \Sigma^*$. Nous avons

$$x \sim_L y \Leftrightarrow x^{-1}.L = y^{-1}.L.$$

Proposition 1.1.7. *Soit $L \subseteq \Sigma^*$ un langage. La relation $\sim_L$ est une relation d'équivalence. Il s'agit même d'une congruence à droite, i.e.,*

$$\forall z \in \Sigma^*, x \sim_L \Rightarrow xz \sim_L yz.$$

Lemme 1.1.8. *Soient $L \subseteq \Sigma^*$ et u, v deux mots sur Σ. Nous avons*

$$(uv)^{-1}.L = v^{-1}.(u^{-1}.L).$$

Définition 1.1.9. Nous définissons l'*automate minimal*

$$\mathcal{M}_L = (Q_L, q_{0,L}, F_L, \Sigma, \delta_L)$$

d'un langage $L \subseteq \Sigma^*$ comme suit :
- $Q_L = \{w^{-1}.L \mid w \in \Sigma^*\}$,
- $q_{0,L} = \varepsilon^{-1}.L = L$,
- $F_L = \{w^{-1}.L \mid w \in L\} = \{q \in Q_L \mid \varepsilon \in q\}$,
- $\delta_L(q, \sigma) = \sigma^{-1}.q$, pour tous $q \in Q_L$, $\sigma \in \Sigma$.

La fonction de transition de l'automate s'étend à $Q_L \times \Sigma^*$ par

$$\delta_L(q, w) = w^{-1}.q, \forall q \in Q_L, w \in \Sigma^*.$$

Cette définition a du sens car la fonction de transition ne dépend pas du représentant choisi. En effet, si un état de $\mathcal{M}_L$ est de la forme $x^{-1}.L = y^{-1}.L$, avec $x, y \in \Sigma^*$, alors $x \sim_L y$. Puisque $\sim_L$ est une congruence à droite, nous avons pour tout $\sigma \in \Sigma$, $x\sigma \sim_L y\sigma$ et donc $(x\sigma)^{-1}.L = (y\sigma)^{-1}.L$. Vu le lemme 1.1.8, cela implique $\sigma^{-1}.(x^{-1}.L) = \sigma^{-1}.(y^{-1}.L)$. Par conséquent, nous obtenons $\delta_L(x^{-1}.L, \sigma) = \delta_L(y^{-1}.L, \sigma)$.

Théorème 1.1.10 (Théorème de Myhill-Nérode)**.** *Un langage $L \in \Sigma^*$ est régulier si et seulement si la congruence $\sim_L$ est d'indice fini (i.e., si $\sim_L$ a un nombre fini de classes d'équivalence).*

1.2 Ensembles ultimement périodiques

Définition 1.2.1. Une partie $X \subseteq \mathbb{N}$ est *ultimement périodique* s'il existe $N \geq 0$ et $p > 0$ tels que pour tout $n \geq N$,

$$n \in X \Leftrightarrow n + p \in X.$$

Le plus petit entier p satisfaisant une telle propriété est appelé la *période* de X et le plus petit entier N correspondant est appelé la *prépériode*.

Remarque 1.2.2. Observons qu'une partie $X \subseteq \mathbb{N}$ est ultimement périodique si et seulement si X est une union finie de progressions arithmétiques. En effet, supposons qu'il existe $N \geq 0$ et $p \geq 0$ tels que pour tout $n \geq N$,

$$n \in X \Leftrightarrow n + p \in X.$$

Dès lors, X s'écrit comme une union finie de progressions arithmétiques,

$$X = \left(\bigcup_{\substack{x \in X \\ x < N}} (x + \mathbb{N} \cdot 0) \right) \cup \left(\bigcup_{\substack{x \in X \\ N \leq x < N+p}} (x + \mathbb{N} \cdot p) \right).$$

Réciproquement, si

$$X = \bigcup_{i=1}^{t} (q_i + \mathbb{N} \cdot p_i),$$

alors en prenant $p = \mathrm{ppcm}_{i=1,\ldots,t} p_i$ et $N = max_{i=1,\ldots,t} q_i$, nous avons

$$\forall n \geq N, \quad n \in X \Leftrightarrow n + p \in X.$$

1.3 Systèmes de numération

Définition 1.3.1. Un *système de numération positionnel* est donné par une suite (strictement) croissante $U = (U_i)_{(i \geq 0)}$ d'entiers tels que

- $U_0 = 1$
- $C_U := \sup_{i \geq 0} \lceil U_{i+1}/U_i \rceil$ est fini.

La première condition assure que tout entier a au moins une représentation. La deuxième condition implique que l'alphabet $A_U = \{0, \ldots, C_U - 1\}$ est fini.

Définition 1.3.2. La *U-représentation gloutonne* d'un entier positif n est l'unique mot $\mathrm{rep}_U(n) = w_\ell \cdots w_0$ sur A_U satisfaisant

- $n = \sum_{i=0}^{\ell} w_i U_i$,
- $w_\ell \neq 0$,
- $\sum_{i=0}^{t} w_i U_i < U_{t+1}, \forall t = 0, \ldots, \ell$.

Nous posons $\mathrm{rep}_U(0) = \varepsilon$ où ε désigne le mot vide.

Un ensemble d'entiers $X \subseteq \mathbb{N}$ est *U-reconnaissable* si le langage

$$\mathrm{rep}_U(X) = \{\mathrm{rep}_U(x) \mid x \in X\}$$

sur l'alphabet A_U est régulier.

Exemple 1.3.3. Considérons le système de numération en base 2. La suite U est donnée par $U_i = 2^i$ pour tout $i \geq 0$. Il s'agit bien d'un système de numération positionnel puisque $U_0 = 1$ et $C_U := \sup_{i \geq 0} \lceil 2^{i+1}/2^i \rceil = 2$ est fini. Nous avons par exemple

$$\mathrm{rep}_U(17) = 10001 \text{ et } \mathrm{val}_U(101101) = 45.$$

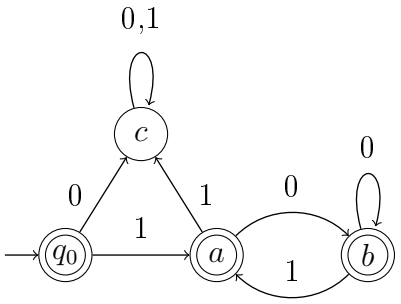

FIGURE 1.1 – Automate acceptant le langage $\mathrm{rep}_F(\mathbb{N})$.

Exemple 1.3.4. Soit $(F_i)_{i\geq 0}$ la suite des nombres de Fibonacci définie par $F_0 = 1$, $F_1 = 2$ et $F_{i+2} = F_{i=1} + F_i$ pour tout $i \geq 0$. Le *système de numération de Fibonacci* est donné par $F = (F_i)_{i\geq 0} = (1, 2, 3, 5, 8, 13, 21, 34, \ldots)$. Nous avons,

$$\mathrm{rep}_F(17) = 100101 \text{ et } \mathrm{val}_F(101001) = 19.$$

Nous remarquons que le langage $\mathrm{rep}_F(\mathbb{N})$ est l'ensemble contenant le mot vide et les mots commençant par 1 qui ne contiennent pas le facteur 11, i.e.,

$$\mathrm{rep}_F(\mathbb{N}) = \{\varepsilon\} \cup 1\{0,1\}^* \setminus \{0,1\}^*11\{1,0\}^*.$$

Ce langage est régulier et il est accepté par l'automate représenté dans la Figure (1.1).

Définition 1.3.5. Si $x = x_\ell \cdots x_0$ est un mot sur un alphabet fini d'entiers, alors la *valeur U-numérique* de x est $\mathrm{val}_U(x) = \sum_{i=0}^{\ell} x_i U_i$.

Le résultat suivant est une conséquence du caractère glouton de la représentation.

Proposition 1.3.6. *Soient un système de numération positionnel $U = (U_i)_{i\geq 0}$ et x, y deux mots sur A_U. Si xy est une représentation gloutonne et si la lettre la plus à gauche de y n'est pas 0, alors y est aussi une représentation gloutonne.*

Démonstration. Soient deux entiers $j = |y|$ et $\ell = |x| - j$. Notons $x = z_\ell \cdots z_{j+1}$, $y = z_j \cdots z_0$ avec $z_\ell, \ldots, z_0 \in A_U$. Puisque le mot $xy = z_\ell \cdots z_0$ est une représentation gloutonne, il vient, par définition,

$$\sum_{i=0}^{t} z_i U_i < U_{t+1} \quad \forall t \in \{0, \ldots, \ell\}.$$

En particulier, nous avons $\sum_{i=0}^{t} z_i U_i < U_{t+1}$ pour tout $t \in \{0, \ldots, j\}$. De plus, z_j n'est pas 0 par hypothèse. De ces deux observations, nous déduisons que le mot y est une représentation gloutonne. □

Définition 1.3.7. Nous ordonnons l'alphabet A_U selon l'ordre naturel des chiffres et nous définissons l'*ordre lexicographique* comme suit : pour deux mots finis $x, y \in A_U^*$ de même longueur, x est lexicographiquement plus petit que y s'il existe $w, x', y' \in A_U^*$ et $a, b \in A_U$ tels que $x = wax'$, $y = wby'$ et $a < b$. Dans ce cas, nous notons $x <_{lex} y$.

Nous étendons l'ordre lexicographique aux mots infinis de la manière suivante. Pour deux mots infinis $x, y \in A_U^{\mathbb{N}}$, nous avons $x <_{lex} y$ si et seulement s'il existe $w \in A_U^*$, $a, b \in A_U$ et $x', y' \in A_U^{\mathbb{N}}$ tels que $x = wax'$, $y = wby'$ et $a < b$.

Nous munissons alors A_U^* d'une relation d'ordre : l'*ordre généalogique*. Pour deux mots $x, y \in A_U^*$, x est généalogiquement plus petit que y si $|x| < |y|$ ou $x <_{lex} y$. Nous notons alors $x <_{gén} y$.

Remarque 1.3.8. Pour $m, n \in \mathbb{N}$, nous avons

$$m < n \Leftrightarrow \mathrm{rep}_U(m) <_{gén} \mathrm{rep}_U(n).$$

En effet, soient $v = v_k \cdots v_0 = \mathrm{rep}_U(m)$ et $w = w_j \cdots w_0 = \mathrm{rep}_U(n)$ avec $u_k \leq m < u_{k+1}$ et $u_j \leq n < u_{j+1}$. Supposons, $v <_{gén} w$. Alors $k \leq j$. Si $k < j$, $u_{k+1} \leq u_j$ et $m < n$. Si $k = j$, il existe i tel que $v_i <_{gén} w_i$ et $v_k \cdots v_{i+1} = w_k \cdots w_{i+1}$. D'où, il vient

$$\begin{aligned} m &= v_k U_k + \cdots + v_0 U_0 \\ &= w_k U_k + \cdots + w_{i+1} U_{i+1} + v_i U_i + \cdots + v_0 U_0 \\ &\leq w_k U_k + \cdots + w_{i+1} U_{i+1} + (w_i - 1) U_i + v_{i-1} U_{i-1} + \cdots + v_0 U_0 \\ &< w_k U_k + \cdots + w_{i+1} U_{i+1} + w_i U_i \\ &\leq n \end{aligned}$$

car $v_{i-1} U_{i-1} + \cdots + v_0 U_0 < U_i$ vu le caractère glouton de la représentation. Donc, nous obtenons $m < n$.

Définition 1.3.9. Un système de numération positionnel $U = (U_i)_{i \geq 0}$ est dit *linéaire* si la suite U satisfait une relation de récurrence linéaire homogène à coefficients entiers, i.e., s'il existe un naturel $k \geq 1$ et des coefficients constants $a_1, \ldots, a_k$ tels que pour tout $i \geq 0$, nous avons

$$U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \qquad \text{avec } a_1, \ldots, a_k \in \mathbb{Z}\,. \tag{1.1}$$

Nous appelons k l'*ordre* de la relation de récurrence.

Exemple 1.3.10. Considérons à nouveau le système de numération de Fibonacci $F = (F_i)_{i\geq 0}$. Il est linéaire puisque la suite F satisfait la relation de récurrence linéaire $F_{i+2} = F_{i+1} + F_i$ pour tout $i \geq 2$.

Définition 1.3.11. Soit un système de numération positionnel linéaire $U = (U_i)_{i\geq 0}$ satisfaisant (1.1). Le *polynôme caractéristique* de U est

$$\chi_U(x) = x^k - a_1 x^{k-1} - \cdots - a^k.$$

Exemple 1.3.12. Considérons le système de numération de Fibonacci $F = (F_i)_{i\geq 0}$. Son polynôme caractéristique est

$$\chi_F(x) = x^2 - x - 1.$$

Dans ce travail, nous nous intéressons au problème de décision suivant.

Problème 1. Soient un système de numération linéaire U tel que $\mathbb{N}$ est U-reconnaissable et un ensemble $X \subseteq \mathbb{N}$ U-reconnaissable donné par un automate acceptant $\mathrm{rep}_U(X)$. Pouvons-nous décider si X est ou non ultimement périodique ?

Remarquons que la régularité de $\mathrm{rep}_U(\mathbb{N})$ implique qu'il existe un ensemble $X \subseteq \mathbb{N}$ tel que $\mathrm{rep}_U(X)$ est régulier. En effet, vu la proposition 2.3.2, si $\mathbb{N}$ est U-reconnaissable, alors tout ensemble ultimement périodique est U-reconnaissable.

Remarque 1.3.13. Si nous nous restreignons aux systèmes de numération en base entière $b \geq 2$, plusieurs résultats sont connus. Par exemple, J. Honkala a montré dans [14] que le problème de décision décrit ci-dessus est décidable. Toutefois, nous verrons que nous ne pouvons pas appliquer notre procédure dans ce cas, puisque $N_U(m) \not\to +\infty$ lorsque $m \to +\infty$ où U est un système de numération en base entière ≥ 2.

Remarque 1.3.14. A. Muchnik a montré dans [22] que le problème 1 était décidable pour tout système de numération U linéaire pour lequel $\mathrm{rep}_U(\mathbb{N})$ et l'addition sont reconnaissables par automates. Rappelons que, pour un système de numération U, l'addition est calculable par un automate fini si le langage

$$\left\{ \begin{pmatrix} 0^{m-|\mathrm{rep}_U(x)|}\mathrm{rep}_U(x) \\ 0^{m-|\mathrm{rep}_U(y)|}\mathrm{rep}_U(y) \\ 0^{m-|\mathrm{rep}_U(z)|}\mathrm{rep}_U(z) \end{pmatrix} \mid x, y, z \in \mathbb{N}, x + y = z, m \max_{t\in\{x,y,z\}} |rep_U(t)| \right\}$$

est régulier. Notons que nous avons ajouté des zéros en tête des deux composantes les plus courtes pour que les trois mots aient la même longueur.

Cependant, il n'est pas facile de caractériser les systèmes de numération U pour lesquels l'addition est calculable par automate fini. Quelques exemples sont donnés dans [11]. Par exemple, le système de numération linéaire U défini par

$$U_{i+4} = 3U_{i+3} + 2U_{i+2} - 3U_{i-4} \forall i \geq 0$$

et $1 = U_0 < U_1 < U_2 < U_3$ est tel que l'addition n'est pas calculable par automate fini. Cependant, comme nous le verrons dans l'exemple 3.4.6, notre procédure de décision pourra être appliquée à ce système.

Chapitre 2

Une procédure de décision pour une classe de systèmes de numération linéaires

Dans les sections suivantes, nous considérons souvent un système de numération positionnel $U = (U_i)_{i\geq 0}$ satisfaisant la condition :

$$\lim_{i\to+\infty} U_{i+1} - U_i = +\infty \tag{2.1}$$

Remarquons que ce n'est pas une condition très contraignante. D'habitude, la suite U a une croissance exponentielle, $U_i \simeq \beta^i$ pour un $\beta > 1$ et donc (2.1) est satisfait.

2.1 Préliminaires

Lemme 2.1.1. *Soit $U = (U_i)_{i\geq 0}$ un système de numération positionnel satisfaisant* (2.1)*. Alors, pour tout j, il existe un naturel L tel que pour tout $\ell \geq L$, les mots*

$$10^{\ell-|\mathrm{rep}_U(t)|}\mathrm{rep}_U(t),\ \textit{avec } t = 0,\ldots,U_j - 1$$

sont des U-représentations gloutonnes. Autrement dit, si w est une U-représentation gloutonne, alors pour tout r suffisamment grand, $10^r w$ est aussi une U-représentation gloutonne.

Démonstration. Puisque $\mathrm{rep}_U(U_j) = 10^j$, $\mathrm{rep}_U(U_j - 1)$ est le plus grand mot de longueur j dans le langage $\mathrm{rep}_U(\mathbb{N})$. Par hypothèse, nous avons $\lim_{i\to+\infty} U_{i+1} - U_i = +\infty$. Donc il existe L tel que pour tout $\ell \geq L$,

$U_{\ell+1} - U_\ell > U_j - 1$. Par conséquent, pour tout $\ell \geq L$, $10^{\ell-j}\mathrm{rep}_U(U_j - 1)$ est la U-représentation gloutonne de $U_\ell + U_j - 1 < U_{\ell+1}$. La conclusion en découle car, pour tout $t \in \{0, \ldots, U_j - 1\}$, nous avons $U_\ell + t < U_{\ell+1}$ et donc $10^{\ell-j}\mathrm{rep}_U(U_j - 1)$ est une U-représentation gloutonne. □

Exemple 2.1.2. Considérons le système de numération $U = (U_i)_{i\geq 0}$ défini par $U_i = i + 1$ pour tout$0 \leq i < 5$ et $U_{5i+r} = 5^{i+1} + r$ pour tout $i \geq 1$ avec $r \in \{0, \ldots, 4\}$. Si l'indice i est de la forme $5j + r$ avec un entier $j \geq 1$ et $r \in \{0, \ldots, 3\}$, nous obtenons $U_{i+1} - U_i = 5^{j+1} + r + 1 - 5^{j+1} - r = 1$. Donc, $U_{i+1} - U_i = 1$ pour un nombre infini de i et le système U ne satisfait pas la condition (2.1). Le lemme précédent n'est pas applicable dans ce cas. Nous avons $\mathrm{rep}_U(4) = 1000$ mais le mot $10^{5i+1}1000$ n'est pas une U-représentation gloutonne. En effet, nous avons

$$\begin{aligned}\mathrm{val}_U(10^{5i+1}1000) &= \mathrm{val}_U(10^{5i+5}) + 4 \\ &= U_{5(i+1)} + 4 = 5^{i+2} + 4 = U_{5(i+2)+4} = U_{5i+9}.\end{aligned}$$

Or, la U-représentation gloutonne de U_{5i+9} est $\mathrm{rep}_U(U_{5i+9}) = 10^{5i+9}$.

Remarque 2.1.3. Dans le lemme 2.1.1, nous ne pouvons pas échanger l'ordre des quantificateurs de j et L car L dépend de j. Par exemple, prenons la suite $(U_i)_{i\geq 0}$ définie par

$$U_i = \frac{(i+1)(i+2)}{2} \qquad \forall i \geq 0.$$

Elle satisfait la relation de récurrence linéaire

$$U_{i+3} = 3U_{i+2} - 3U_{i+1} + U_i \qquad \forall i \geq 3.$$

De plus, nous avons

$$\lim_{i\to+\infty} U_{i+1} - U_i = \lim_{i\to+\infty} i + 2 = +\infty.$$

Pour tout entier $i \geq 1$, $k = U_i - 1$ est l'unique valeur telle que $U_i = U_k - U_{k-1}$ puisque pour tout entier $j \geq 1$, $U_j + U_{j-1} = j + 1$. Nous remarquons que pour tout $i \geq 1$, le mot 10^i est une U-représentation gloutonne et les U-représentations gloutonnes de la forme 10^n10^i sont exactement celles pour lesquelles $n \geq U_i - i - 1$.

Intéressons-nous maintenant à une classe particulière de systèmes de numération positionnels, voir aussi [9, 17, 20].

Définition 2.1.4. Les *systèmes de numération de Bertrand* associés à un réel $\beta > 1$ sont définis de la manière suivante. Soit $A_\beta = \{0, \ldots, \lfloor \beta \rfloor\}$. Tout $x \in [0, 1]$ peut s'écrire

$$x = \sum_{i=1}^{+\infty} c_i \beta^{-i}, \text{avec } c_i \in A_\beta.$$

Alors la suite $(c_i)_{i \geq 1}$ est appelée la *β-représentation* de x. Une β-représentation de x, qui est maximale pour l'ordre lexicographique, est notée $d_\beta(x)$ et est appelée *β-développement*. Un β-développement $(c_i)_{i \geq 0}$ est *fini* s'il existe N tel que $c_i = 0$ pour tout $i \geq N$. Si le β-développement de 1 est fini, alors il existe $m \geq 1$ tel que $d_\beta(1) = t_1 \cdots t_m$, avec $t_m \neq 0$ et nous posons

$$d^*_\beta(1) := (t_1 \cdots t_{m-1}(t_m - 1))^\omega.$$

Sinon, $d_\beta(1)$ est infini et nous posons $d^*_\beta(1) := d_\beta(1)$.

Notons D_β l'ensemble des β-développements des nombres appartenant à $[0, 1[$ et $F(D_\beta)$ l'ensemble des facteurs de D_β.

Définition 2.1.5. Un langage L est dit *extensible à droite* si

$$v \in L \Rightarrow v0 \in L.$$

Exemple 2.1.6. Considérons la suite $U = (U_i)_{i \geq 0}$ commençant par $U_0 = 1$, $U_1 = 3$ et $U_2 = 5$. Le langage $\mathrm{rep}_U(\mathbb{N})$ n'est pas extensible à droite. En effet, le mot $2 = \mathrm{rep}_U(2)$ est dans $\mathrm{rep}_U(\mathbb{N})$. Par contre, le mot 20 n'est pas une représentation valide puisque sa valeur U-numérique est $\mathrm{val}_U(20) = 6$, mais la U-représentation gloutonne de 6 est $rep_U(6) = 101$.

Théorème 2.1.7 (Théorème de Bertrand)**.** *Soit $U = (U_i)_{i \geq 0}$ un système de numération positionnel. Soit A_U l'alphabet canonique. Il existe un nombre réel β tel que $\mathrm{rep}_U(\mathbb{N}) = F(D_\beta)$ si et seulement si $\mathrm{rep}_U(\mathbb{N})$ est extensible à droite. Dans ce cas, si $d^*_\beta(1) = (d_i)_{i \geq 1}$, la suite U est déterminée par*

$$U_i = d_1 U_{i-1} + \cdots + d_i U_0 + 1. \tag{2.2}$$

Définition 2.1.8. Les systèmes de numération qui satisfont le théorème précédent sont dits *associés à β* et notés U_β.

Exemple 2.1.9. Considérons le système de numération de Fibonacci $F = (F_i)_{i \geq 0}$. Le langage $\mathrm{rep}_F(\mathbb{N})$, qui est l'ensemble contenant le mot vide et les mots commençant par 1 qui ne contiennent pas le facteur 11, est clairement extensible à droite. Donc, il satisfait les hypothèses du théorème de Bertrand

2.1.7. Le système de numération de Fibonacci est associé au nombre d'or $\beta = \frac{1+\sqrt{5}}{2}$. En effet, nous avons

$$d_\beta(1) = 11 \text{ car } 1 = \frac{2}{1+\sqrt{5}} + (\frac{2}{1+\sqrt{5}})^2.$$

Par conséquent, $d^*_\beta(1) = (10)^\omega$ et la suite $F = (F_i)_{i\geq 0}$ est bel et bien déterminée par

$$\begin{aligned} F_i &= 1F_{i-1} + 0F_{i-2} + \underbrace{1F_{i-3} + 0F_{i-4} + \cdots + \alpha F_0 + 1}_{=F_{i-2}} \\ &= F_{i-1} + F_{i-2}, \end{aligned}$$

où α vaut 0 si i est pair et 1 sinon.

Proposition 2.1.10 (Condition de Parry)**.** *Soient un réel $\beta > 1$ et une suite $(x_n)_{n\geq 0}$ d'entiers positifs. La suite $(x_n)_{n\geq 0}$ appartient à D_β si et seulement si, pour tout $k \geq 0$, $(x_n)_{n\geq k} <_{lex} d^*_\beta(1)$.*

Remarque 2.1.11. La condition de Parry signifie que tous les facteurs des mots de D_β sont strictement inférieurs au mot infini $d^*_\beta(1)$. Par conséquent, si β est un *nombre de Parry*, i.e., tel que $d_\beta(1)$ est fini ou ultimement périodique, alors nous pouvons construire canoniquement un automate $\mathcal{A}$ acceptant exactement le langage $F(D_\beta) = \mathrm{rep}_{U_\beta}(\mathbb{N})$. Construisons d'abord l'automate $\mathcal{A}_\beta$ acceptant le langage $0^*F(D_\beta) = 0^*\mathrm{rep}_{U_\beta}(\mathbb{N})$.

Si $d_\beta(1)$ est fini, il s'écrit $d_\beta(1) = t_1 \cdots t_m$, avec $t_i \in A_\beta$ pour tout $i \in \{1, \ldots, m\}$. L'ensemble des états de $\mathcal{A}_\beta$ est $Q = \{q_1, \ldots, q_m\}$. L'état initial est q_1. Tous les états sont finals. Les transitions sont tels que, pour tout $1 \leq j < m$,

$$q_j \xrightarrow{t_j} q_{j+1}$$

et, pour tout $r \in A_\beta$ tel que $r <_{lex} t_j$ et tout $1 \leq j \leq m$,

$$q_j \xrightarrow{r} q_1.$$

Les autres transitions ne sont pas définies (ou nous pouvons dire qu'elles vont dans un "puits"). L'automate $\mathcal{A}_\beta$ est représenté à la figure 2.1 où les états finals sont représentés par un simple cercle dans un souci de simplicité.

Si $d_\beta(1)$ est ultimement périodique, alors il existe des entiers N et P tels que, nous avons,

$$d_\beta(1) = t_1 \cdots t_N (t_{N+1} \cdots t_{N+p})^\omega$$

avec $t_i \in A_\beta$ pour tout $i \in \{1, \ldots, N+P\}$. Les états de l'automate $\mathcal{A}_\beta$ sont $q_1, \ldots, q_{N+P}$, tous sont finals. L'état initial est q_1. L'automate $\mathcal{A}_\beta$ est

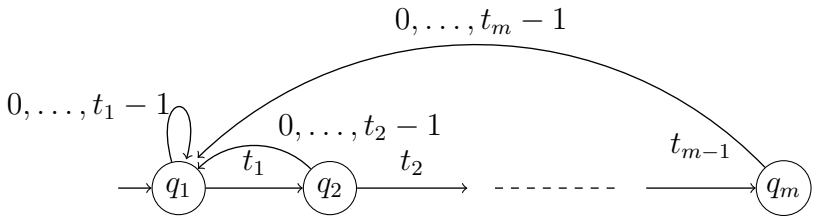

FIGURE 2.1 – Automate $\mathcal{A}_\beta$ si $d_\beta(1)$ est fini.

construit de la même manière que dans le cas fini pour les états $q_1, \ldots, q_N$. Il y a juste une partie supplémentaire qui s'occupe de la période. Une représentation de $\mathcal{A}_\beta$ est donnée à la figure 2.2, avec la même convention pour les états finals.

L'automate $\mathcal{A}_\beta$ accepte le langage $0^*\mathrm{rep}_{U_\beta}$. Pour avoir un automate acceptant exactement rep_{U_β}, modifions légèrement $\mathcal{A}_\beta$. L'état q_1 n'est plus un état initial et nous ajoutons le nouvel état initial q_0 avec les transitions

$$q_0 \xrightarrow{t_1} q_2 \text{ et } q_0 \xrightarrow{r} q_1 \text{ pour } r \in \{0, \ldots, t_1 - 1\}.$$

L'automate acceptant rep_{U_β} est représenté à la figure 2.3.

Exemple 2.1.12. Soit $\beta \simeq 3,383$ la racine réelle du polynôme $x^3 - 3x^2 - x - 1$. Considérons le système $U_\beta = (U_i)_{i \geq 0}$ associé à β. Nous avons $d_\beta(1) = 311$ et $d^*_\beta(1) = (310)^\omega$. Par le théorème 2.1.7, la suite U_β satisfait

$$U_0 = 1, U_1 = 3U_0 + 1 = 4, U_2 = 3U_1 + U_0 + 1 = 14$$

et, pour tout $i \geq 0$,

$$\begin{aligned} U_{i+3} &= 3U_{i+2} + U_{i+1} + 0U_i + \underbrace{3U_{i-1} + U_{i-2} + \cdots + 1}_{=U_i} \\ &= 3U_{i+2} + U_{i+1} + U_i. \end{aligned}$$

Donc, nous avons $U_\beta = (1, 4, 14, 47, 159, 538, \cdots)$. De plus, puisque $d_\beta(1)$ est fini, β est un nombre de Parry. L'automate $\mathcal{A}_\beta$, à la figure 2.1 accepte le langage $F(D_\beta) = \mathrm{rep}_{U_\beta}(\mathbb{N})$.

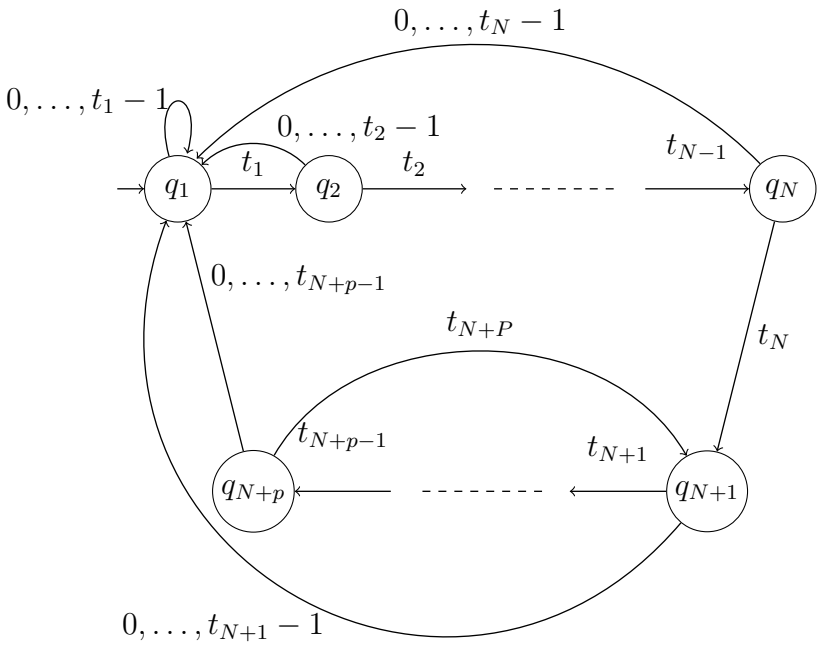

FIGURE 2.2 – Automate $\mathcal{A}_\beta$ si $d_\beta(1)$ est ultimement périodique.

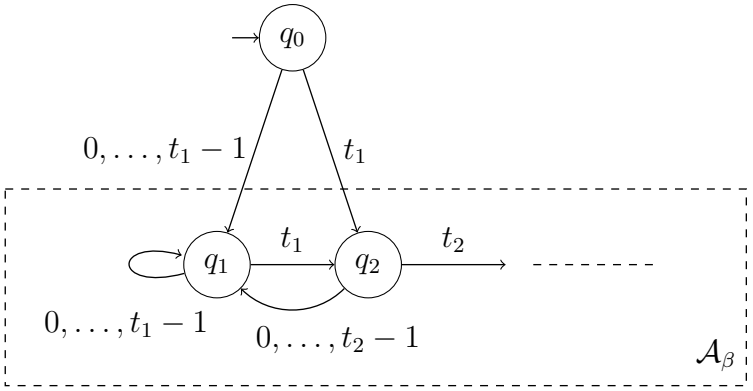

FIGURE 2.3 – Automate acceptant $\mathrm{rep}_{U_\beta}(\mathbb{N})$.

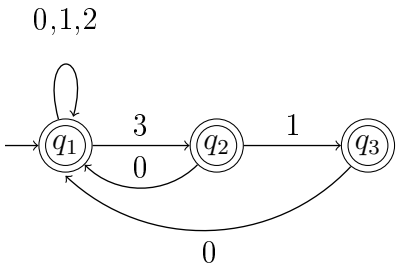

FIGURE 2.4 – Automate $\mathcal{A}_\beta$ avec $\beta \simeq 3,383$.

Remarque 2.1.13. Soit β un nombre de Parry. Vu la forme de l'automate $\mathcal{A}_\beta$, nous obtenons la propriété suivante, qui est plus forte que le lemme 2.1.1. Si x et y sont des U_β-représentations gloutonnes, alors $x0y$ est aussi une U_β-représentation gloutonne.

Exemple 2.1.14. Prenons le système de numération de Fibonacci F_β où $\beta = \frac{1+\sqrt{5}}{2}$. Puisque $d_\beta = 11$ est fini, β est un nombre de Parry. Donc, pour deux mots x et $y \in \mathrm{rep}_{F_\beta}(\mathbb{N})$, nous avons $x0y \in \mathrm{rep}_{F_\beta}(\mathbb{N})$.

Remarque 2.1.15. Lorsque β est un nombre de Parry, le système associés à β satisfait une relation de récurrence linéaire. Par le théorème de Bertrand, la suite U_β satisfait (2.2). Vu [9], il en découle que la suite U_β compte le nombre de mots de longueur n acceptés par $\mathcal{A}$, où $\mathcal{A}$ est l'automate acceptant exactement $F(D_\beta) = \mathrm{rep}_{U_\beta}(\mathbb{N})$. Vu la remarque 4.1.6, cette suite satisfait une relation de récurrence linéaire.

Définition 2.1.16. Soit $X \subseteq \mathbb{N}$ un ensemble d'entiers. Le *mot caractéristique* de X est le mot infini $x_0x_1x_2\cdots$ sur $\{0,1\}$ défini par $x_i = 1$ si et seulement si $i \in X$.

Considérons maintenant $X \subseteq \mathbb{N}$ un ensemble ultimement périodique. Le mot caractéristique de X est donc un mot infini sur $\{0,1\}$ de la forme $x_0x_1x_2\cdots = uv^\omega$ où u et v sont choisis pour qu'ils soient de longueurs minimales. Nous disons que la longueur $|u|$ de u est la *prépériode* de X et la longueur de $|v|$ de v est la *période* de X. Alors, pour tout $n \geq |u|$, $n \in X$ si et seulement si $n + |v| \in X$. Cette définition est donc cohérente avec 1.2.1.

Lemme 2.1.17. *Soit $X \subseteq \mathbb{N}$ un ensemble ultimement périodique de période p_X et de prépériode a_X. Soient $i, j \geq a_X$. Si $i \not\equiv j \bmod p_x$, alors il existe $t < p_X$ tel que nous avons, soit $i + t \in X$ et $j + t \notin X$, soit $i + t \notin X$ et $j + t \in X$.*

Démonstration. Sans perte de généralité, nous pouvons supposer $i > j$ et $p_X > 1$. Procédons par contradiction et supposons que, pour tout $t \in \{0, \ldots, p_X\}$, nous avons $i + t \in X \Leftrightarrow j + t \in X$. Soient $p \in \{1, \ldots, p_X - 1\}$ tel que $p \equiv j - i \pmod{p_X}$ et un entier $n \geq i$. Nous pouvons écrire $n \equiv i + r \pmod{p_X}$ avec $r \in \{0, \ldots, p_X - 1\}$. Il vient $n + p \equiv j + r \pmod{p_X}$. Donc, nous obtenons

$$n + p \in X \Leftrightarrow j + r \in X \Leftrightarrow i + r \in X \Leftrightarrow n \in X.$$

Ceci est en contradiction avec le fait que p_X est la période minimale.

□

2.2 Borne inférieure et supérieure sur la période

Définition 2.2.1. Pour une suite $(U_i)_{i\geq 0}$ d'entiers, nous notons $N_U(m) \in \{1, \ldots, m\}$ le nombre de valeurs qui sont prises infiniment souvent par la suite $(U_i \bmod m)_{i\geq 0}$.

Exemple 2.2.2. Considérons la suite $(U_i)_{i\geq 0}$ définie à l'exemple 2.1.2. Pour tout $i \geq 1$, nous avons $U_{5i+r} = 5^{i+1} + r \equiv r \pmod 5$, avec $r \in \{0, \ldots, 4\}$. Donc, dans la suite $(U_i \bmod 5)$, les valeurs $0, 1, 2, 3, 4$ apparaissent infiniment souvent. De plus, comme il ne peut y avoir d'autres valeurs différentes dans cette suite, nous avons $N_U(5) = 5$.

Proposition 2.2.3. *Soit $U = (U_i)_{i\geq 0}$ un système de numération satisfaisant (2.1). Si $X \subseteq \mathbb{N}$ est un ensemble ultimement périodique U-reconnaissable de période p_X, alors tout automate fini déterministe acceptant $\mathrm{rep}_U(X)$ a au moins $N_U(p_X)$ états.*

Démonstration. Soit a_X la prépériode de X. Par le lemme 2.1.1, il existe L tel que pour tout $h \geq L$, les mots

$$10^{h-|\mathrm{rep}_U(t)|}\mathrm{rep}_U(t), \text{avec } t = 0, \ldots, p_X - 1,$$

sont des U-représentations gloutonnes puisque p_X est tel que $U_{j-1} \leq p_X \leq U_j$.

La suite $(U_i \bmod p_X)_{i\geq 0}$ prend infiniment souvent $N := N_U(p_X)$ valeurs distinctes. Soient $h_1, \ldots, h_N \geq L$ tels que

$$i \neq j \Rightarrow U_{h_i} \not\equiv U_{h_j} \bmod p_X.$$

Les $h_1, \ldots, h_N$ peuvent être choisis tels que $U_{h_i} > a_X$ pour $i \in \{1, \ldots, N\}$. Par le lemme 2.1.17, pour tous $i, j \in \{1, \ldots, N\}$ tels que $i \neq j$ (et donc $U_{h_i} \not\equiv$

$U_{h_j} \bmod p_X$), il existe $t_{i,j} < p_X$ tel que soit $U_{h_i} + t_{i,j} \in X$ et $U_{h_j} + t_{i,j} \notin X$, ou soit $U_{h_i} + t_{i,j} \notin X$ et $U_{h_j} + t_{i,j} \in X$. Donc le mot

$$w_{i,j} = 0^{|\mathrm{rep}_U(p_X-1)|-|\mathrm{rep}_U(t_{i,j})|}\mathrm{rep}_U(t_{i,j})$$

est tel que soit

$$10^{h_i-|\mathrm{rep}_U(p_X-1)|}w_{i,j} \in \mathrm{rep}_U(X) \text{ et } 10^{h_j-|\mathrm{rep}_U(p_X-1)|}w_{i,j} \notin \mathrm{rep}_U(X)$$

ou

$$10^{h_i-|\mathrm{rep}_U(p_X-1)|}w_{i,j} \notin \mathrm{rep}_U(X) \text{ et } 10^{h_j-|\mathrm{rep}_U(p_X-1)|}w_{i,j} \in \mathrm{rep}_U(X)$$

car

$$\begin{aligned} 10^{h_i-|\mathrm{rep}_U(p_X-1)|}w_{i,j} &= 10^{h_i-|\mathrm{rep}_U(p_X-1)|}0^{|\mathrm{rep}_U(p_X-1)|-|\mathrm{rep}_U(t_{i,j})|}\mathrm{rep}_U(t_{i,j}) \\ &= 0^{h_i-|\mathrm{rep}_U(t_{i,j})|}\mathrm{rep}_U(t_{i,j}). \end{aligned}$$

Par conséquent, les mots

$$10^{h_1-|\mathrm{rep}_U(p_X-1)|}, \ldots, 10^{h_N-|\mathrm{rep}_U(p_X-1)|}$$

sont deux à deux non équivalents pour la relation $\sim_{\mathrm{rep}_U(X)}$. Donc l'automate minimal de $\mathrm{rep}_U(X)$ a au moins $N = N_U(p_X)$ états. □

Une conséquence immédiate de cette proposition nous donne une borne inférieure sur la période d'un ensemble périodique accepté par un automate fini déterministe.

Corollaire 2.2.4. *Soit $U = (U_i)_{i\geq 0}$ un système de numération positionnel satisfaisant* (2.1)*. Supposons que $\lim_{m\to+\infty} N_U(m) = +\infty$. Alors la période d'un ensemble ultimement périodique $X \subseteq \mathbb{N}$ tel que $\mathrm{rep}_U(X)$ est accepté par un AFD à d états est bornée par le plus petit entier s_0 tel que pour tout $m \geq s_0$, $N_U(m) > d$.*

Démonstration. Vu la proposition 2.2.3, nous savons que $d \geq N_U(p_X)$ car d est plus grand ou égal au nombre d'états de l'automate minimal acceptant $\mathrm{rep}_U(X)$. □

Le résultat suivant permet d'obtenir une borne supérieure sur la période d'un ensemble périodique accepté par un AFD.

Proposition 2.2.5. *Soient $U = (U_i)_{i\geq 0}$ un système de numération positionnel satisfaisant* (2.1) *et $X \subseteq \mathbb{N}$ un ensemble ultimement périodique U-reconnaissable de période p_X. Soit c un diviseur de p_X. Si 1 apparaît un nombre infini de fois dans la suite $(U_i \bmod c)_{i\geq 0}$, alors tout AFD acceptant $\mathrm{rep}_U(X)$a au moins c états.*

Démonstration. Soit a_X la prépériode de X. En appliquant plusieurs fois le lemme 2.1.1, nous voyons qu'il existe $n_1, \ldots, n_c$ tels que

$$10^{n_c}10^{n_{c-1}} \cdots 10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|-|\mathrm{rep}_U(t)|}\mathrm{rep}_U(t)$$

avec $t = 0, \ldots, p_X - 1$, sont des U-représentations gloutonnes.

Puisque 1 apparaît un nombre infini de fois dans la suite $(U_i \bmod c)_{i\geq 0}$, $n_1, \ldots, n_c$ peuvent être choisis tels que pour tout $j = 1, \ldots, c$,

$$\mathrm{val}_U(10^{n_j} \cdots 10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|}) \equiv j \bmod c.$$

En effet, nous pouvons choisir n_1 de sorte qu'il existe i_1 tel que

$$\mathrm{val}_U(10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|}) = U_{i_1} \text{ avec } U_{i_1} \equiv 1 \bmod c$$

ainsi que n_j, avec $j = 2, \ldots, c$, de sorte qu'il existe i_j tel que

$$\mathrm{val}_U(10^{n_j}0^{n_{j-1}+1} \cdots 0^{n_1+1}0^{|\mathrm{rep}_U(p_X-1)|}) = U_{i_j} \text{ avec } U_{i_j} \equiv 1 \bmod c.$$

Nous obtenons donc

$$\begin{aligned}
x_j &:= \mathrm{val}_U(10^{n_j} \cdots 10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|}) \\
&= \mathrm{val}_U(10^{n_j}0^{n_{j-1}+1} \cdots 0^{n_1+1}0^{|\mathrm{rep}_U(p_X-1)|}) \\
&\quad + \mathrm{val}_U(10^{n_{j-1}}0^{n_{j-2}+1} \cdots 0^{n_1+1}0^{|\mathrm{rep}_U(p_X-1)|}) \\
&\quad + \cdots + \mathrm{val}_U(10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|}) \\
&= U_{i_j} + \cdots + U_{i_1} \\
&\equiv 1 + \cdots + 1 \pmod c \\
&\equiv j \pmod c
\end{aligned}$$

De plus, nous pouvons prendre n_1 tel qu'il vérifie aussi

$$x_1 = \mathrm{val}_U(10^{n_1}0^{|\mathrm{rep}_U(p_X-1)|}) > a_X.$$

Donc, nous avons $x_i > a_X$ pour tout $i \in \{1, \ldots, c\}$. Par le lemme 2.1.17, pour tous $i, j \in \{1 \ldots, n\}$ tels que $i \neq j$, il existe $t_{i,j} \in X$ tel que soit $x_i + t_{i,j} \in X$ et $x_j + t_{i,j} \notin X$, ou soit $x_i + t_{i,j} \notin X$ et $x_j + t_{i,j} \in X$. Remarquons que le fait que c divise p_X implique $|\mathrm{rep}_U(t_{i,j})| < |\mathrm{rep}_U(p_X - 1)|$. D'où, nous avons, pour tout $1 \leq i \leq c$,

$$\mathrm{rep}_U(x_i+t_{i,j}) = 10^{n_i} \cdots 10^{n_1+|\mathrm{rep}_U(p_X-1)|-|\mathrm{rep}_U(t_{i,j})|}\mathrm{rep}_U(t_{i,j}) = 10^{n_i} \cdots 10^{n_1}w_{i,j}.$$

Donc, le mot $w_{i,j} = 0^{|\mathrm{rep}_U(p_X-1)|-|\mathrm{rep}_U(t_{i,j})|}\mathrm{rep}_U(t_{i,j})$ est tel que

$$10^{n_i} \cdots 10^{n_1}w_{i,j} \in X \text{ et } 10^{n_j} \cdots 10^{n_1}w_{i,j} \notin X$$

ou tel que

$$10^{n_i} \cdots 10^{n_1} w_{i,j} \notin X \text{ et } 10^{n_j} \cdots 10^{n_1} w_{i,j} \in X.$$

Par conséquent, les mots

$$10^{n_i} \cdots 10^{n_1} \text{ et } 10^{n_j} \cdots 10^{n_1}$$

ne sont pas équivalents pour $\sim_{\mathrm{rep}_U(X)}$. Donc, tout AFD acceptant $\mathrm{rep}_U(X)$ a au moins c états. □

Définition 2.2.6. Pour une suite $(U_i)_{i\geq 0}$ d'entiers, si $(U_i \bmod m)_{i\geq 0}$ est ultimement périodique, nous notons sa prépériode (minimale) $\iota_U(m)$ et sa période (minimale) $\pi_U(m)$.

Remarque 2.2.7. Nous observons que, pour toute suite U satisfaisant une relation de récurrence linéaire d'ordre k de la forme (1.1), la suite $(U_i \bmod m)$ est ultimement périodique. De plus, nous avons

$$N_U(m) \leq \pi_U(m) \leq (N_U(m))^k.$$

En effet, puisque les valeurs apparaissant infiniment souvent sont les valeurs se trouvant dans la période, $N_U(m) \leq \pi_U(m)$. D'un autre coté, pour trouver la période de $(U_i \bmod m)$, il suffit d'inspecter les k-uples composés de k éléments consécutifs. Dès qu'il y en a deux identiques, nous avons la période. Or, le nombre de k-uples disctincts dans la suite $(U_i \bmod m)$ est $(N_U(m))^k$. Donc $\pi_U(m) \leq (N_U(m))^k$ Par conséquent, $\lim_{m\to+\infty} N_U(m) = +\infty$ si et seulement si $\lim_{m\to+\infty} \pi_U(m) = +\infty$. Remarquons que si $m = p \cdot q$ avec $\mathrm{pgcd}(p, q) = 1$, alors $\pi_U(m) = \mathrm{ppcm}\{\pi_u(p), \pi_u(q)\}$.

Exemple 2.2.8. Considérons à nouveau la suite $(U_i)_{i\geq 0}$ définie par $U_i = \frac{(i+1)(i+2)}{2}$ pour tout $i \geq 0$. Ses premiers termes sont

$$(U_i)_{i\geq 0} = (1,3,6,10,15,21,28,36,45,55,66,78,91,105,$$
$$120,136,153,171,190,210,231,253,276,300,325,\cdots).$$

Puisque la suite $(U_i)_{i\geq 0}$ satisfait la relation de récurrence d'ordre 3,

$$U_{i+3} = U_{i+2} - U_{i+1} + U_i \text{ pour tout } i \geq 0,$$

la suite $(U_i \bmod 15)_{i\geq 0}$ est ultimement périodique et pour trouver la période, il suffit d'inspecter les premières valeurs de cette suite à la recherche de deux blocs de trois élements consécutifs qui sont égaux. Les premiers termes de $(U_i \bmod 15)_{i\geq 0}$ sont

$$(1,3,6,10,0,6,13,6,0,10,6,3,1,0,0,1,3,6,10,0,6,13,6,0,10,\cdots).$$

Nous voyons que le triplet $(U_0, U_1, U_2) = (1, 3, 6)$ est égal au triplet (U_{15}, U_{16}, U_{17}). Dans ce cas, la suite $(U_i \bmod 15)_{i\geq 0}$ est même périodique de période $\pi_U(15) = 15$. Les différentes valeurs prises dans la période sont $0, 1, 3, 6, 10$ et 13. Ce sont les seules valeurs répétées infiniment souvent. D'où, nous avons $N_U(15) = 5$.

Exemple 2.2.9. Prenons la suite de Fibonacci $F = (F_i)_{i\geq 0}$. Si nous écrivons les premiers termes de la suite, nous avons

$$(1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, \cdots).$$

Considérons maintenant la suite $(F_i \bmod 4)_{i\geq 0}$. Ses premiers termes sont

$$(1, 2, 3, 1, 0, 1, 1, 2, 3, 1, 0, 1, 1, 2, 3, 1, 0 \cdots).$$

Par le même raisonnement que précédemment, nous voyons que la période est $(1, 2, 3, 1, 0, 1)$, donc $\pi_U(4) = 6$ et $N_U(4) = 4$.

2.3 Borne supérieure sur la prépériode

Nous voulons maintenant obtenir une borne supérieure sur la prépériode de tout ensemble ultimement périodique U-reconnaissable reconnu par un AFD donné.

Proposition 2.3.1. *Soient $U = (U_i)_{i\geq 0}$ un système linéaire et $X \subseteq N$ un ensemble ultimement périodique U-reconnaissable de période p_X et de prépériode a_X. Alors tout AFD acceptant $\mathrm{rep}_U(X)$ a au moins $|\mathrm{rep}_U(a_X - 1)| - \iota_U(p_X)$ états.*

Démonstration. Sans perte de généralité, nous pouvons supposer que

$$|\mathrm{rep}_U(a_X - 1)| - \iota(p_X) > 0.$$

Puisque U satisfait une relation de récurrence linéaire d'ordre k, la suite $(U_i \bmod p_X)_{i\geq 0}$ est ultimement périodique de période $\pi_U(p_X)$ et de prépériode $\iota_U(p_X)$. Procédons par contradiction et supposons que $\mathcal{A}$ est un AFD avec moins de $|\mathrm{rep}_U(a_X - 1)| - \iota(p_X)$ états acceptant $\mathrm{rep}_U(X)$. Il existe des mots w et w_4 sur A_U tels que la U-représentation gloutonne de $a_X - 1$ peut être factorisée comme

$$\mathrm{rep}_U(a_X - 1) = ww_4 \quad \text{avec } |w| = |\mathrm{rep}_U(a_X - 1)| - \iota_U(p_X).$$

Puisque $\mathrm{rep}_U(X)$ est régulier, nous pouvons appliquer le lemme de la pompe. Le mot w peut s'écrire $w_1w_2w_3$ avec $w_2 \neq \varepsilon$ et pour tout $i \geq 0$ et tout $v \in A_U^*$, nous avons

$$wv \in \mathrm{rep}_U(X) \text{ si et seulement si } w_1w_2^iw_3v \in \mathrm{rep}_U(X).$$

En particulier, pour tout $i \geq 0$,

$$ww_4 \in \mathrm{rep}_U(X) \text{ si et seulement si } w_1w_2^iw_3w_4 \in \mathrm{rep}_U(X).$$

Par minimalité de a_X et p_X, soit $a_X - 1 \in X$ et donc, pour tout $n \geq 1$, $a_X + np_X - 1 \notin X$, soit $a_X - 1 \notin X$ et donc, pour tout $n \geq 1$, $a_X + np_X - 1 \in X$. En effet, supposons dans le premier cas qu'il existe un entier $n \geq 1$ tel que $a_X + np_X - 1 \in X$. Alors, puisque p_X est la période, nous avons pour tout $n \geq 1$, $a_X + np_X - 1 \in X$ et $a_X - 1 \in X$. Par conséquent a_X n'est pas la prépériode minimale de X, ce qui est une contradiction. Nous raisonnons de la même manière pour le deuxième cas.

En utilisant la période de $(U_i \bmod p_X)_{i\geq 0}$, nous voyons que, puisque $|w_4| = \iota_U(p_X)$, nous avons pour tout $i \geq 0$,

$$\begin{aligned}
&\mathrm{val}_U(w_1w_2^{i\pi_U(p_X)}w_2w_3w_4) \\
&\quad = \mathrm{val}_U(w_10^{|w_2^{i\pi_U(p_X)}|}w_2w_3w_4) + \mathrm{val}_U(w_2^{i\pi_U(p_X)}0^{|w_2w_3w_4|}) \\
&\quad \equiv \mathrm{val}_U(w_1w_2w_3w_4) + i\mathrm{val}_U(w_2^{\pi_U(p_X)}0^{|w_2w_3w_4|}) \pmod{p_X}.
\end{aligned}$$

Par conséquent, lorsque nous répétons un facteur dont la longueur est un multiple de $\pi_U(p_X)$ exactement p_X fois, la valeur $\mathrm{mod} p_X$ du mot ne change pas. Nous obtenons

$$\mathrm{val}_U(w_1w_2^{p_X\pi_U(p_X)}w_2w_3w_4) \equiv \mathrm{val}_U(w_1w_2w_3w_4) \pmod{p_X}.$$

Il en découle une contradiction par rapport à la minimalité de a_X et de p_X. □

Lemme 2.3.2. *Soient des entiers a, $b \geq 0$ et $U = (U_i)_{i\geq 0}$ un système de numération linéaire. Le langage*

$$\mathrm{val}_U^{-1}(a\,\mathbb{N} + b) = \{w \in A_U^* | \mathrm{val}_U(w) \in a\,\mathbb{N} + b\}$$

est régulier. En particulier, si $\mathbb{N}$ est U-reconnaissable, alors un AFD acceptant $\mathrm{rep}_U(a\,\mathbb{N} + b)$ peut être obtenu de manière effective et tout ensemble ultimement périodique est U-reconnaissable.

Remarque 2.3.3. Notons que pour tout entier $n \geq 0$, $\mathrm{val}_U^{-1}(n) \setminus 0^+A_U^*$ est un ensemble fini de mots $\{x_1, \ldots, x_{t_n}\}$ sur A_U tels que $\mathrm{val}_U(x_i) = n$ pour tout $i = 1, \ldots, t_n$. En particulier, nous avons $\mathrm{rep}_U(n) \in \{x_1, \ldots, x_{t_n}\}$.

Démonstration. Puisque les langages réguliers sont stables pour les modifications finies (i.e., ajouter ou enlever un nombre fini de mots dans le langage), nous pouvons supposer que $0 \leq b < a$. La suite $(U_i \bmod a)_{i\geq 0}$ est ultimement périodique (car U est un système de numération linéaire) de prépériode $\ell = \iota_U(a)$ et de période $p = \pi_U(a)$. Construisons un AFD $\mathcal{A}$ acceptant le miroir des mots de $\{w \in A_U^* \mid \mathrm{val}_U(w) \in a\,\mathbb{N} + b\}$.

- L'alphabet de $\mathcal{A}$ est A_U.
- Les états sont les paires (r, s) où $0 \leq r < a$ et $0 \leq s < \ell + p$.
- L'état initial est la paire $(0, 0)$.
- Les états finals sont les paires (b, s) où $0 \leq s < \ell + p$.
- Les transitions sont définies comme suit : pour tout $j \in A_U$,

$$\forall s < \ell + p : (r, s) \xrightarrow{j} (jU_s + r \bmod a, s + 1)$$

$$(r, \ell + p - 1) \xrightarrow{j} (jU_{\ell+p-1} + r \bmod a, \ell).$$

Nous remarquons que $\mathcal{A}$ ne tient pas compte du caractère glouton ou non des mots acceptés. La construction de $\mathcal{A}$ repose uniquement sur la valeur U-numérique des mots modulo a.

Supposons $\mathbb{N}$ U-reconnaissable. Le langage

$$\mathrm{rep}_U(a\,\mathbb{N} + b) = \mathrm{rep}_U(\mathbb{N}) \cap \mathrm{val}_U(a\,\mathbb{N} + b)$$

est régulier car les langages réguliers sont stables pour l'intersection. Donc, nous pouvons construire effectivement un automate qui imite simultanément les automates acceptant $\mathrm{val}_U(a\,\mathbb{N} + b)$ et $\mathrm{rep}_U(\mathbb{N})$. Ce type d'automate est appelé automate produit. Il sera décrit plus en détails dans le chapitre 4 en 4.3.4. □

Dans le résultat précédent, l'hypothèse sur le caractère U-reconnaissable de $\mathbb{N}$ a un intérêt particulier. En effet, pour un système U de numération linéaire arbitraire, en général, $\mathbb{N}$ n'est pas U-reconnaissable. De plus, nous avons la propriété suivante provenant de [18].

Proposition 2.3.4. *Soit un système de numération positionnel $U = (U_i)_{i\geq 0}$. Si $\mathbb{N}$ est U-reconnaissable, alors la suite $(U_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire à coefficients entiers, i.e., U est un système linéaire.*

Démonstration. Nous avons $\mathrm{rep}_U(U_\ell) = 10^\ell$ pour tout $\ell \geq 0$. Parmi les mots de longueur $\ell + 1$, le plus petit mot par rapport à l'ordre généalogique est 10^ℓ. Donc, $U_{\ell+1} - U_\ell$ est exactement le nombre de mots de longueur $\ell + 1$ dans $\mathrm{rep}_U(\mathbb{N})$. D'un autre côté, puisque le langage $\mathrm{rep}_U(\mathbb{N})$ est régulier, il est accepté par un automate fini déterministe. Le nombre de mots de longueur n dans $\mathrm{rep}_U(\mathbb{N})$ est égal au nombre de chemins de longueur n de l'état initial à un des états finals. Nous savons, par [24], que la suite $(\#(\mathrm{rep}_U(\mathbb{N})) \cap A_U^n)_{n\geq 0}$ satisfait une relation de récurrence linéaire avec des coefficients entiers. La conclusion en découle. □

Remarque 2.3.5. La réciproque du résultat précédent est fausse. Exhibons un contre-exemple provenant de [29]. Considérons le système de numération positionnel $U = (U_i)_{i\geq 0}$ défini par $U_i = (i + 1)^2$ pour tout $i \geq 0$. La suite U satisfait la relation de récurrence linéaire

$$U_{i+3} = 3U_{i+2} - 3U_{i+1} + U_i \text{ pour tout } i \geq 0.$$

L'ensemble L des U-représentations des nombres qui sont la somme de deux carrés est

$$\mathrm{rep}_U(\mathbb{N}) \cap 10^*10^* = \{10^a10^b \mid b^2 < 2a + 4\}.$$

Montrons que l'ensemble L n'est pas régulier.

Supposons que l'ensemble L est régulier. Alors, il est accepté par un automate fini. Soit k le nombre de ses états. Prenons $a \geq k$ tel que $2a + 3$ est un carré et $b = \sqrt{2a + 3}$. Le mot 10^a10^b est un mot de L. Puisque L est régulier, nous pouvons appliquer le lemme de la pompe à ce mot 10^a10^b. Il existe des mots x, y, z tels que $10^a10^b = xyz$ avec $y \neq \varepsilon$ et $|xy| < k \leq a$. Supposons qu'il existe des entiers i et j tels que $x = 10^i$, $y = 0^j$, avec $j \neq 0$, et $z = 0^{a-i-j}10^b$, le cas où $x = \varepsilon$ et $y = 10^j$ étant similaire. Par le lemme de la pompe, nous avons $xy^nz \in L$ pour tout entier $n \geq 0$. En particulier, pour $n = 0$, nous obtenons $xz = 10^{a-j}10^b \in L$. Or, nous avons $b^2 > 2(a - j) + 4$ car $j \geq 1$. D'où, nous avons $10^{a-j}10^b \notin L$, ce qui est une contradiction. Par conséquent, l'ensemble L n'est pas régulier.

Puisque la classe des langages réguliers est fermée pour l'intersection, $\mathrm{rep}_U(\mathbb{N})$ n'est pas régulier. Donc, $\mathbb{N}$ n'est pas U-reconnaissable.

2.4 Procédure de décision

Notre procédure de décision utilisera le résultat qui suit.

Lemme 2.4.1. *Soit $U = (U_i)_{i\geq 0}$ une suite croissante satisfaisant une relation de récurrence linéaire de la forme* (1.1)*. Les assertions suivantes sont équivalentes :*

i. $\lim_{m\to+\infty} N_U(m) = +\infty$,

ii. pour tout diviseur premier p *de* a_k, $\lim_{m\to+\infty} N_U(p^v) = +\infty$.

En particulier, si $a_k = \pm 1$, *alors* $\lim_{m\to+\infty} N_U(m) = +\infty$.

Démonstration. Il est évident que (i) implique (ii). Montrons que (ii) implique (i). Soit $|a_k| = p_1^{u_1}\cdots p_r^{u_r}$, avec $u_1,\ldots,u_r > 0$, la décomposition en facteurs premiers de $|a_k|$. Il est clair que si $m = p_1^{v_1}\cdots p_r^{v_r}c$ avec $v_1,\ldots,v_r \geq 0$ et $\operatorname{pgcd}(a_k, c) = 1$, alors

$$\pi_U(m) = \operatorname{ppcm}\{\pi_U(p_1^{v_1}),\ldots,\pi_U(p_r^{v_r}),\pi_U(c)\}.$$

Observons que si m tend vers l'infini, alors au moins un des v_j ou c tend vers l'infini.

Supposons que pour un $j \in \{1,\ldots,r\}$, nous avons $v_j \to +\infty$. Par hypothèse, nous avons $\lim_{m\to+\infty} N_U(p_j^{v_j}) = +\infty$. Vu la remarque 2.2.7, $N_U(p_j^{v_j}) \geq \pi_U(p_j^{v_j}) \geq (N_U(p_j^{v_j}))^k$. Nous obtenons alors $\lim_{v_j\to+\infty} \pi_U(p_j^{v_j}) = +\infty$. Donc $\pi_U(m)$ prend des valeurs plus grandes que n'importe quelle constante lorsque m est un entier divisible par une puissance suffisamment grande de p_j. Vu la remarque 2.2.7, la même conclusion vaut pour $N_U(m)$.

Posons $C = \{c_0 < c_1 < c_2 < \cdots\}$ l'ensemble des naturels premiers avec a_k. Pour tout $c \in C$, la suite $(U_i \bmod c)_{i\geq 0}$ est ultimement périodique (car U satisfait une relation de récurrence linéaire d'ordre k) et elle est même purement périodique. En effet, pour tout $i \geq 0$, U_{i+k} est déterminé par les k termes précédents $U_{i+k-1},\ldots,U_i$. Toutefois, puisque $\operatorname{pgcd}(a_k, c) = 1$, a_k est inversible modulo c et pour tout $i \geq 0$, $U_i \bmod c$ est aussi déterminé par les k termes suivants $U_{i+1},\ldots,U_{i+k}$.

Par définition de $N_U(c)$, la suite $(U_i \bmod c)_{i\geq 0}$ prend exactement $N_U(c)$ valeurs distinctes, car chaque terme apparaît infiniment souvent. Soit α la fonction qui envoie $m \in \mathbb{N}$ sur le plus petit indice $\alpha(m)$ tel que $U_{\alpha(m)} \geq m$. Puisque la suite U est croissante, la fonction α est croissante et

$$\lim_{m\to+\infty} \alpha(m) = +\infty.$$

De cette observation et de la périodicité de $(U_i \bmod c)_{i\geq 0}$, il découle que pour tout $c \in C$, $N_U(c) \geq \alpha(c)$ car $U_0 < \cdots < U_{\alpha(c)-1} < c$. Par conséquent, nous obtenons $\lim_{n\to+\infty} N_U(c_n) = +\infty$. Donc tout entier m contenant un nombre c suffisamment grand et premier avec a_k vérifie $\lim_{m\to+\infty} N_U(m) = +\infty$.

Puisque tout entier m suffisamment grand contient soit une grande puissance d'un p_j, soit un grand entier c premier avec a_k, l'assertion (i) est vérifiée. □

Théorème 2.4.2. *Soit $U = (U_i)_{i\geq 0}$ un système de numération linéaire tel que $\mathbb{N}$ est U-reconnaissable satisfaisant la condition* (2.1)*. Supposons que $\lim_{m\to+\infty} N_U(m) = +\infty$. Alors, nous pouvons décider si un ensemble U-reconnaissable est ou non ultimement périodique.*

Démonstration. La suite U satisfait une relation de récurrence linéaire d'ordre k de la forme (1.1). Soit $|a_k| = p_1^{u_1} \cdots p_r^{u_r}$, avec $u_1, \ldots, u_r > 0$, la décomposition en facteurs premiers de $|a_k|$. Considérons un AFD $\mathcal{A}$ avec d états acceptant un ensemble U-reconnaissable $X \subseteq \mathbb{N}$.

Supposons que X est ultimement périodique de période $p_X = p_1^{v_1} \cdots p_r^{v_r} c$ avec $v_1, \ldots, v_r \geq 0$ et $\mathrm{pgcd}(a_k, c) = 1$. Comme dans la preuve précédente puisque a_k et c sont premiers entre eux, la suite $(U_i \bmod c)_{i\geq 0}$ est purement périodique. Donc le terme $U_0 = 1$ apparaît un nombre infini de fois dans $(U_i \bmod c)_{i\geq 0}$. Puisque c est un diviseur de p_X, nous pouvons utiliser la proposition 2.2.5 et nous obtenons $c \leq d$.

Par la proposition 2.2.3, nous obtenons $N_U(p_X) \leq d$. Soit $j \in \{1, \ldots, r\}$. Vu la remarque 2.2.7, nous avons

$$N_U(p_j^{v_j}) \leq \pi_U(p_j^{v_j}) \leq \pi_U(p_X) \leq (N_U(p_X))^k \leq d^k.$$

L'hypothèse $\lim_{m\to+\infty} N_U(m) = +\infty$ implique $\lim_{v\to+\infty} N_U(p_j^v) = +\infty$. Remarquons que $N_U(p_j^v) \leq N_U(p_j^w)$ lorsque $v \leq w$. Par conséquent, l'exposant v_j apparaissant dans la décomposition de p_X est borné par s_j où s_j est le plus petit entier tel que pour tout $v \geq s_j$, $N_U(p_j^v) > d^k$. Cette borne s_j peut être effectivement calculée de la manière suivante. Pour tout $v \in \mathbb{N}$, $N_U(p_j^v)$ peut être calculé en un nombre fini d'opérations, en inspectant les premières valeurs de $(U_i \bmod p_j^v)_{i\geq 0}$ et en cherchant deux k-uples identiques composés de k éléments consécutifs. Une fois la période déterminée, nous pouvons immédiatement obtenir les valeurs qui sont répétées infiniment souvent. Puisque l'application $v \mapsto N_U(p_j^v)$ est croissante, nous pouvons calculer $N_U(p_j) \leq N_U(p_j^2) \leq \cdots$ jusqu'à trouver la première valeur s_j telle que $N_U(p_j^{s_j}) > d^k$.

Si X est ultimement périodique, alors les périodes admissibles sont bornées par la constante

$$P = p_1^{s_1} \cdots p_r^{s_r} d$$

qui est calculable effectivement. Par la proposition 2.3.1, les prépériodes admissibles a_X doivent satisfaire

$$|\mathrm{rep}_U(a_x - 1)| \leq d + \max_{p\leq P}(\iota_U(p))$$

où $|\mathrm{rep}_U(a)| \leq |\mathrm{rep}_U(b)|$ lorsque $a \leq b$. Cette dernière observation montre qu'une borne supérieure sur les périodes admissibles de X peut être effectivement donnée.

Par conséquent, l'ensemble des prépériodes et périodes admissibles que nous avons à vérifier est fini. Pour chaque paire (a, p) de prépériode et période admissibles, il y a au plus $2^a 2^b$ ensembles ultimement périodiques distincts. Par le lemme 2.3.2, puisque $\mathbb{N}$ est U-reconnaissable par hypothèse, tout ensemble ultimement périodique est U-reconnaissable. Donc, nous pouvons construire un automate pour chacun d'eux et ensuite comparer le langage L accepté par cet automate avec $\mathrm{rep}_U(X)$. Puisque tester si $L \setminus \mathrm{rep}_U(X) = \emptyset$ et $\mathrm{rep}_U(X) \setminus L = \emptyset$ est décidable par algorithme, nous pouvons décider si un ensemble U-reconnaissable est ultimement périodique ou non. □

Vu le résultat précédent, il est naturel de vouloir caractériser les suites U linéaires récurrentes telles que $\lim_{m\to+\infty} N_U(m) = +\infty$. Ce sera fait dans le chapitre 3. Toutefois, nous pouvons déjà nous intéresser à un cas particulier immédiat.

Corollaire 2.4.3. *Soit $U = (U_i)_{i\geq 0}$ un système de numération linéaire tel que $\mathbb{N}$ est U-reconnaissable et satisfaisant la condition* (2.1) *et une relation de récurrence linéaire d'ordre k de la forme* (1.1) *avec $a_k = \pm 1$. Nous pouvons décider si un ensemble U-reconnaissable est ultimement périodique ou non.*

Démonstration. Vu le cas particulier du lemme 2.4.1, puisque $a_k = \pm 1$, nous avons $\lim_{m\to+\infty} N_U(m) = +\infty$. La conclusion découle alors du théorème 2.4.2. □

Remarque 2.4.4. Nous avons donc obtenu une procédure de décision pour le problème 1 lorsque $\lim_{m\to+\infty} N_U(m) = +\infty$ et en particulier quand le coefficient a_k apparaissant dans (1.1) est égal à ± 1.

D'un autre côté, quand $\mathrm{pgcd}(a_1, \ldots, a_k) = g \geq 2$, pour tout $n \geq 1$ et tout i suffisamment grand, nous avons $U_i \equiv 0 \pmod{g^n}$ et l'hypothèse concernant $N_U(m)$ dans le théorème 2.4.2 n'est pas vérifiée. En effet, la seule valeur prise infiniment souvent par la suite $(U_i \bmod g^n)_{i\geq 0}$ est 0. Donc $N_U(m) = 1$ pour un nombre infini de m. Remarquons que la même observation peut être faite pour le système habituel en base entière $b \geq 2$. Dans ce cas, la seule valeur prise infiniment souvent par la suite $(b^i \bmod b^n)_{i\geq 0}$ est 0 pour tout $n \geq 1$.

Pour conclure ce chapitre, nous faisons une petite digression. Nous montrons comment utiliser le résultat d'Engstrom [10] sur les prépériodes pour obtenir des systèmes de numérations linéaires particuliers où $\lim_{m\to+\infty} N_U(m) = +\infty$.

Théorème 2.4.5. *[10] Soient $U = (U_i)_{i\geq 0}$ une suite linéaire récurrente d'ordre k de la forme* (1.1) *et p un diviseur premier de a_k. Si il existe $s(p) < k$ tel que $a_k, \ldots, a_{k-s(p)+1} \equiv 0 \pmod{p}$ et $a_{k-s(p)} \not\equiv 0 \pmod{p}$, alors $\iota_U(p^v) \leq vs(p)$.*

Remarque 2.4.6. Soit un système de numération linéaire $U = (U_i)_{i\geq 0}$ satisfaisant (1.1). Supposons que les hypothèses du théorème précédent sont vérifiées pour tous les diviseurs premiers p de a_k. C'est équivalent au fait que $\mathrm{pgcd}(a_1, \ldots, a_k) = 1$. Soit χ_U le polynôme caractéristique de U. Supposons que $\beta > 1$ est une racine de multiplicité $\ell \geq 1$ de $\chi_U(x)$ satisfaisant :

- pour tout autre racine $\gamma \in \mathbb{C}$ de $\chi_U(x)$, $|\gamma| < \beta$,
- $\beta < p^{1/s(p)}$ pour tout diviseur premier p de a_k

Soient $\gamma_1, \cdots, \gamma_r$ les autres racines de multiplicité $m_1, \ldots, m_r$ de $\chi_U(x)$. Vu la proposition 3.3.1, il existe des polynômes $P - \ell$ et Q_i respectivement de degré strictement inférieur à ℓ et m_i, pour $i \in \{1, \ldots, r\}$, tels que

$$U_i = Q_1(i)\gamma_1^i + \cdots + Q_r(i)\gamma_r^i + Q_\ell(i)\beta^i.$$

Vu les hypothèses sur β, il existe une constante c telle que $U_i \sim ci^{\ell-1}\beta^i$ $(i \to +\infty)$. Soit p un diviseur premier de a_k. Notons $j_p(v)$ le plus grand indice j tel que $U_j < p^v$. Soit $t > s(p)$ un réel tel que $\beta < p^{1/t} < p^{1/s(p)}$. Pour tout entier v suffisamment grand, nous avons $U_{\lfloor vt \rfloor} < p^v$. Donc, pour tout entier v suffisamment grand, il vient $j_p(v) \geq \lfloor vt \rfloor$. Vu le théorème 2.4.5, nous avons $\iota_U(p^v) \leq vs(p)$ pour tout $v \in \mathbb{N}$. D'où, pour tout entier v suffisamment grand,

$$U_{\iota_U(p^v)} < \cdots < U_{j_p(v)}$$

sont les premières valeurs de la partie périodique de la suite $(U_i \bmod p^v)_{i\geq 0}$ et nous obtenons $N_U(p^v) \geq \lfloor vt \rfloor - vs(p) + 1 \geq v(t - s(p))$. Par conséquent, pour tout diviseur premier p de a_k, nous avons $N_u(p^v) \to +\infty$ lorsque $v \to +\infty$. Par le lemme 2.4.1, il vient $\lim_{m\to+\infty} N_U(m) = +\infty$. Donc, nous pouvons appliquer notre procédure de décision donnée par le théorème 2.4.2, lorsque $\mathbb{N}$ est U-reconnaissable.

Exemple 2.4.7. Considérons la relation de récurrence linéaire donnée par

$$U_{i+3} = U_{i+1} + 3U_i \text{ pour tout } i \geq 0$$

et $U_i = i + 1$ pour $i \in \{0, 1, 2\}$. Les premiers termes de la suite sont

$$(1, 2, 3, 5, 9, 14, 24, 41, 66, 113, 189, 311, 528, 878, 1461, 2462, 4095, 6845, \cdots).$$

Le polynôme caractéristique est $\chi_U(x) = x^3 - x - 3$.

Avec les mêmes notations que précédemment, le seul diviseur premier de $a_k = 3$ est 3 et $\beta \simeq 1,6717 < 3$ est une racine de multiplicité 1. Les deux autres racines ont un module proche de $1,34 < \beta$. Nous avons $s(3) = 1$. Par le théorème 2.4.5, la prépériode $\iota_U(3^v)$ est bornée par v. D'un autre côté, nous avons $U_i \sim c\beta_i$ pour une constante $c > 0$. Remarquons aussi que

$\beta < 3^{1/2} < 3$. Alors, pour tout v suffisamment grand, $U_{2v} \sim c\beta^{2v} < 3^v$. D'où, les termes $U_v < \cdots < U_{2v}$ apparaissent dans la partie périodique de $(U_i \bmod 3^v)_{i\geq 0}$.

Par exemple, écrivons les premiers termes de la suite $(U_i \bmod 3^v)_{i\geq 0}$ lorsque $v = 3$. Nous avons

$$(1, 2, 3, 5, 9, 14, 24, 14, 12, 5, 0, 14, 15, 14, 3, 5, 18, 14, 6, 14, 21, 5, 9, 14, 24, \ldots).$$

Nous voyons que le triplet $(5, 9, 14)$ apparaît deux fois. Nous obtenons donc que la partie périodique est

$$(5, 9, 14, 24, 14, 12, 5, 0, 14, 15, 14, 3, 5, 18, 14, 6, 14, 21).$$

La prépériode est $(1, 2, 3)$ et $\iota_U(3^3) =\leq vs(p)$. Les termes $U_3 = 5$, $U_4 = 9$, $U_5 = 14$ et $U_{2\cdot 3} = 24$ sont bien les premiers termes de la partie périodique.

Nous pouvons procéder de la même façon pour $v = 4$ et $v = 5$. Les résultats sont repris dans la table suivante où les éléments $U_v < \ldots < U_{2v}$ ont été soulignés. Remarquons que les parties périodiques pour $v = 4$ et $v = 5$ ne sont pas écrites en entier car les périodes $\pi_U(3^4)$ et $\pi_U(3^5)$ sont grandes.

v	prépériode	période
3	$1, 2, 3$	$(\underline{5, 9, 14, 24}, 14, 12, 5, 0, 14, 15, 14, 3, 5, 18, 14, 6, 14, 21)$
4	$1, 2, 3, 5$	$(\underline{9, 14, 24, 41, 66}, 32, 27, 68, 42, 68, 3, 32, 45, 41, 60, 14, \cdots)$
5	$1, 2, 3, 5, 9$	$(\underline{14, 24, 41, 66, 113, 189}, 68, 42, 149, 3, 32, 207, 41, 60, \cdots)$

Chapitre 3

Suites linéaires récurrentes et classes de résidus

Comme cela a été dit dans la remarque 2.4.4, puisque notre approche pour résoudre le problème de décision 2 est d'exiger $\lim_{m\to+\infty} N_U(m) = +\infty$, nous pouvons seulement appliquer notre procédure à des suites linéaires de la forme (1.1) avec $\text{pgcd}(a_1, \ldots, a_k) = 1$. Dans ce chapitre, nous voulons déterminer quelles suites linéaires récurrentes U sont telles que $\lim_{m\to+\infty} N_U(m) = +\infty$. Dans ce but, il est nécessaire de présenter quelques notions d'analyse p-adique. Pour plus d'informations à ce sujet, nous pouvons consulter [16, 12, 27] ou encore les notes de cours [2]. Tout au long de la section suivant, p est un nombre premier fixé.

3.1 Nombres p-adiques

Rappelons qu'une *norme* sur un champ F est une application $\|\cdot\| : F \to \{r \in \mathbb{R} \mid r \geq 0\}$ telle que, pour $x, y \in F$,

i. $\|x\| = 0$ si et seulement si $x = 0$,

ii. $\|x \cdot y\| = \|x\| \cdot \|y\|$,

iii. $\|x + y\| \leq \|x\| + \|y\|$.

Une norme $\|\cdot\|$ est *non-archimédienne* si la troisième inégalité peut être remplacée par l'*inégalité ultramétrique* qui est une inégalité plus forte :

iv. $\|x + y\| \leq \max\{\|x\|, \|y\|\}$.

Définition 3.1.1. Pour rappel, p est un nombre premier fixé. Nous appelons *valuation p-adique* d'un nombre entier $a \neq 0$, notée $\text{ord}_p(a)$, l'exposant de p

dans la décomposition de a en produit de facteurs premiers. Nous pouvons alors l'étendre à tout nombre rationnel non nul $x = \frac{a}{b}$ en posant

$$\mathrm{ord}_p(x) = \mathrm{ord}_p(a) - \mathrm{ord}_p(b).$$

Cette définition dépend uniquement du nombre x et non de a et b. En effet, nous pouvons écrire $x = \frac{ac}{bc}$ et nous obtenons

$$\begin{aligned}\mathrm{ord}_p(x) &= \mathrm{ord}_p(ac) - \mathrm{ord}_p(bc)\\ &= \mathrm{ord}_p(a) + \mathrm{ord}_p(c) - \mathrm{ord}_p(b) - \mathrm{ord}_p(c)\\ &= \mathrm{ord}_p(a) - \mathrm{ord}_p(b).\end{aligned}$$

Nous pouvons maintenant définir la *valeur absolue p-adique* d'un rationnel x, notée $|x|_p$ comme suit :

$$|x|_p = \begin{cases} p^{-\mathrm{ord}_p(v)} & \text{si } x > 0\\ 0 & \text{si } x = 0.\end{cases}$$

Exemple 3.1.2. Nous avons

$$|12|_2 = \frac{1}{4}, |12|_3 = \frac{1}{3}, |12|_5 = |12|_7 = |12|_{11} = 1,$$

$$|2|_2 = |6|_2 = |14|_2 = \frac{1}{2}, |14|_7 = \frac{1}{7}, |25|_5 = |75|_5 = \frac{1}{25},$$

$$\left|\frac{3}{6}\right|_3 = \left|\frac{3}{6}\right|_7 = \left|\frac{10}{6}\right|_2 = 1, \left|\frac{14}{3}\right|_3 = \left|\frac{25}{12}\right|_3 = 3,$$

$$\left|\frac{9}{16}\right|_3 = \left|\frac{18}{25}\right|_3 = \frac{1}{9}, \left|\frac{14}{6}\right|_7 = \left|\frac{7}{15}\right|_7 = \frac{1}{7}, \left|\frac{20}{8}\right|_5 = \frac{1}{5},$$

$$\left|\frac{5}{36} + \frac{14}{3}\right|_3 = \left|\frac{5\cdot 3 + 2^3\cdot 3^2\cdot 7}{2^2\cdot 3^3}\right|_3 = 9 = \max\left\{\left|\frac{5}{36}\right|_3, \left|\frac{14}{3}\right|_3\right\}.$$

Proposition 3.1.3. *La valeur absolue p-adique est une norme non-archimédienne sur* $\mathbb{Q}$.

Démonstration. Soient x et y deux rationnels. Il est clair que $|x|_p = 0$ si et seulement si $x = 0$. Montrons que $|x\cdot y|_p = |x|_p|y|_p$. Cela découle du fait que $\mathrm{ord}_p(xy) = \mathrm{ord}(x) + \mathrm{ord}(y)$. En effet, $\mathrm{ord}_p(xy)$ est le plus grand $n \geq 0$ tel que p^n divise xy. Donc, nous avons $n = k + \ell$ où k (resp. ℓ) est le plus grand naturel tel que p^k divise x (resp. p^ℓ divise y). D'où la conclusion.

Il reste à montrer que $|x+y|_p \leq |x|_p + |y|_p$. Si $x = 0$ ou $y = 0$, alors l'inégalité est directe. De même, si $x + y = 0$, nous avons toujours $|x+y|_p = 0 \leq |x|_p + |y|_p$. Supposons donc $x, y \neq 0$ et $x + y \neq 0$. Nous pouvons écrire $x = \frac{a}{b}$ et $y = \frac{c}{d}$ avec $a, b, c, d \in \mathbb{Z}$. Nous avons

$$\mathrm{ord}_p(x+y) = \mathrm{ord}_p(\frac{ad+cb}{bd}) = \mathrm{ord}_p(ad+cb) - \mathrm{ord}_p b - \mathrm{ord}_p d.$$

Or la plus grande puissance divisant la somme de deux nombres est au moins le minimum de la plus grande puissance divisant le premier nombre et de la plus grande puissance divisant le deuxième nombre. D'où, il vient

$$\begin{aligned}
\mathrm{ord}_p(x+y) &\geq \min\{\mathrm{ord}_p(ad), \mathrm{ord}_p(cb)\} - \mathrm{ord}_p b - \mathrm{ord}_p d \\
&= \min\{\mathrm{ord}_p(a) + \mathrm{ord}_p(d), \mathrm{ord}_p(c) + \mathrm{ord}_p(b)\} - \mathrm{ord}_p b - \mathrm{ord}_p d \\
&= \min\{\mathrm{ord}_p(a) - \mathrm{ord}_p(b), \mathrm{ord}_p(c) - \mathrm{ord}_p(d)\} \\
&= \min\{\mathrm{ord}_p(x), \mathrm{ord}_p(y)\}.
\end{aligned}$$

Par conséquent, nous avons

$$\begin{aligned}
|x+y|_p &= p^{-\mathrm{ord}_p(x+y)} \\
&\leq \max\{p^{-\mathrm{ord}_p(x)}, p^{-\mathrm{ord}_p(y)}\} \\
&= \max\{|x|_p, |y|_p\} \\
&\leq |x|_p + |y|_p.
\end{aligned}$$

En particulier, puisque $|x+y|_p \leq \max\{|x|_p, |y|_p\}$, la valeur absolue p-adique est non-archimédienne. □

Rappelons, en toute généralité, quelques notions de convergence. Soit une suite $(a_n)_{n\geq 0}$ d'éléments d'un anneau R avec une norme N.

Définition 3.1.4. La suite $(a_n)_{n\geq 0}$ *tend vers la limite* $a \in R$ si

$$\forall \delta > 0, \exists M \in \mathbb{N} \text{ tel que } \forall n > M, N(a - a_n) < \delta.$$

Nous notons alors $\lim^{(N)}_{n\to+\infty} a_n = a$. La suite $(a_n)_{n\geq 0}$ est *de Cauchy* si

$$\forall \delta > 0, \exists M \in \mathbb{N} \text{ tel que } \forall m, n > M, N(a_m - a_n) < \delta.$$

La suite $(a_n)_{n\geq 0}$ est une *suite nulle* si $\lim^{(N)}_{n\to+\infty} a_n = 0$, en supposant que la limite existe.

Proposition 3.1.5. *Si* $\lim^{(N)}_{n\to+\infty} a_n$ *existe, alors la suite* $(a_n)_{n\geq 0}$ *est de Cauchy.*

Démonstration. Soient $\lim_{n\to+\infty}^{(N)} a_n = a$ et $\delta > 0$. Il existe un $M \in \mathbb{N}$ tel que pour tout $n > M$, nous avons $N(a - a_n) < \frac{\delta}{2}$. D'où, pour tous $m, n > M$, nous obtenons

$$\begin{aligned} N(a_m - a_n) &= N(a_m - a + a - a_n) \\ &\leq N(a_m - a) + N(a - a_n) \\ &< \frac{\delta}{2} + \frac{\delta}{2} = \delta \end{aligned}$$

Donc, $(a_n)_{n\geq 0}$ est de Cauchy. □

Rappelons quelques définitions de la théorie des champs.

Définition 3.1.6. Un champ K est *complet par rapport à une norme* N si toute suite de Cauchy composée d'éléments de K a une limite dans K par rapport à N. L'ensemble $X \subseteq K$ est *dense* si tout élément de K est une limite (par rapport à la norme N) d'une suite d'éléments de X. Dans notre cas, la norme considérée est la valeur absolue p-adique.

Un champ L est une *extension du champ* K si K est isomorphe à un sous-champ de L. Soit L une extension du champ K. L'élément $x \in L$ est *algébrique sur* K s'il existe un polynôme $f \in K[x] \setminus \{0\}$ tel que $f(x) = 0$. Un champ K est *algébriquement clos* si tout polynôme non constant $f \in K[x]$ a une racine dans K. Une extension L d'un champ K est appelé la *clôture algébrique de* K si L est la plus petite extension de K qui est algébriquement close.

Exemple 3.1.7. Plaçons nous dans le champ $\mathbb{Q}$ muni de la norme euclidienne $|\cdot|$. Considérons le développement décimal de $\sqrt{2}$:

$$\sqrt{2} = 1.414213562\cdots$$

Prenons la suite $a = (a_n)_{n\geq 0}$ des approximations

$$a = (1, \frac{14}{10}, \frac{141}{100}, \frac{1412}{1000}, \ldots)$$

de $\sqrt{2}$. Les éléments de cette suite sont rationnels mais nous avons

$$\lim_{n\to+\infty}^{(|\cdot|)} a_n = \sqrt{2} \notin \mathbb{Q}\,.$$

L'exemple précédent montre que $\mathbb{Q}$ n'est pas complet pour la norme euclidienne. Par rapport à la norme $|\cdot|_p$, en exhibant une suite de rationnels de Cauchy qui n'a pas de limite dans $\mathbb{Q}$, F. Q. Gouvêa a montré dans [12] que $\mathbb{Q}$ n'est pas complet pour la valeur absolue p-adique.

Il est naturel de vouloir "compléter" $\mathbb{Q}$. Tout d'abord, plaçons nous dans un cadre plus général avec un anneau R muni d'une norme N. Notons une suite $(a_n)_{n\geq 0} = (a_n)$ si le contexte est clair. Les résultats présentés ci-après peuvent être trouvés dans [2].

Définition 3.1.8. Nous définissons $\mathrm{CS}(R,N)$ l'ensemble des suites de Cauchy de R (par rapport à la norme N) et $\mathrm{Null}(R,N)$ l'ensemble des suites nulles de R (par rapport à la norme N). En particulier, nous avons $\mathrm{CS}(R,N) \subseteq \mathrm{CS}(R,N)$. Nous définissons l'addition et la multiplication d'éléments de $\mathrm{CS}(R,N)$ par

$$(a_n) + (b_n) = (a_n + b_n) \text{ et } (a_n) \times (b_n) = (a_n)(b_n)$$

avec $(a_n), (b_n) \in \mathrm{CS}(R,N)$. Ce sont bien des opérations de $\mathrm{CS}(R,N) \times \mathrm{CS}(R,N) \to \mathrm{CS}(R,N)$. Avec les éléments $0_{CS} = (0_R)$ et $1_{CS} = (1_R)$ où 0_R et 1_R sont les neutres de $+$ et $\cdot$ dans R, $(\mathrm{CS}(R,N), +, \times)$ a une structure d'anneau avec 0_{CS} (resp. 1_{CS}) neutre pour $+$ (resp. $\times$). Remarquons que si $(a_n) \in \mathrm{CS}(R,N)$ et $(b_n) \in \mathrm{Null}(R,N)$, alors $(a_n b_n)$ et $(b_n a_n)$ sont dans $\mathrm{Null}(R,N)$. L'anneau quotient $\mathrm{CS}(R,N)/\mathrm{Null}(R,N)$ est appelé le *complété de R par rapport à la norme N* et il est noté $\hat{R}_N$ ou $\hat{R}$. Pour $(a_n) \in \mathrm{CS}(R,N)$, nous notons sa classe d'équivalence $\{a_n\}$ dans $\hat{R}$.

Vu le résultat suivant, nous pouvons étendre la norme N au complété de R.

Théorème 3.1.9. *L'anneau $\hat{R}_N$ possède une somme et un produit définis par*

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \text{ et } \{a_n\} \times \{b_n\} = \{a_n b_n\},$$

avec $(a_n), (b_n) \in CS(R,N)$. L'anneau $\hat{R}$ est commutatif si R l'est. De plus, il y a une unique norme $\hat{N}$ sur $\hat{R}_N$ qui satisfait $\hat{N}(\{a\}) = N(a)$ où $(a) = (a)_{n\geq 0}$ est une suite de Cauchy constante avec $a \in R$. Cette norme est définie par

$$\hat{N}(\{c_n\}) = \lim_{n\to+\infty} N(c_n) \text{ avec } (c_n) \in CS(R,N)$$

comme une limite dans les nombres réels. Enfin, la norme $\hat{N}$ est non-archimédienne si et seulement si N est non-archimédienne.

Les deux résultats suivants nous seront utiles par la suite.

Proposition 3.1.10. *Le complété $\hat{R}$ de R est complet par rapport à la norme $\hat{N}$. De plus, R est dense dans $\hat{R}$.*

Théorème 3.1.11. *Si R est un champ, alors $\hat{R}$ est aussi un champ.*

Considérons à nouveau le champ $\mathbb{Q}$ avec la valeur absolue p-adique. Nous complétons $\mathbb{Q}$ par rapport cette valeur absolue et nous obtenons le *champ des rationnels p-adiques*, noté $\mathbb{Q}_p$. La norme sur $\mathbb{Q}_p$ est notée $|\cdot|_p$.

Définition 3.1.12. Le disque unité fermé

$$\{x \in \mathbb{Q}_p : |x|_p \leq 1\}$$

est l'ensemble des *entiers p-adiques*, noté $\mathbb{Z}_p$.

Proposition 3.1.13. *L'ensemble des entiers ordinaires $\mathbb{Z}$ est dense dans $\mathbb{Z}_p$.*

Démonstration. Soient un entier p-adique $x \in \mathbb{Z}_p$ et un entier $n \geq 1$. Puisque $\mathbb{Q}$ est dense dans $\mathbb{Q}_p$ par la proposition 3.1.10, nous pouvons trouver un rationnel $\frac{a}{b} \in \mathbb{Q}$, avec $a, b \in \mathbb{Z}$, assez proche de x tel que $|x - \frac{a}{b}|_p \geq p^{-n} < 1$. Montrons que nous pouvons même choisir un entier assez proche de x. Pour a/b défini ci-dessus, puisque $x \in \mathbb{Z}_p$, nous avons

$$\left|\frac{a}{b}\right|_p = \left|x + (\frac{a}{b} - a)\right|_p \leq \max\left\{|x|_p, \left|x - \frac{a}{b}\right|_p\right\} < 1.$$

Cela signifie que p ne divise pas b. Donc, il existe un entier c tel que $bc \equiv 1 \pmod{p^n}$ et $ac \in \mathbb{Z}$. D'où, il existe un naturel $\lambda < p^n$ tel que $bc = \lambda p^n + 1$ et nous obtenons

$$\left|\frac{a}{b} - ac\right|_p = \left|\frac{a}{b}\right|_p |1 - bc|_p \leq \left|\frac{a}{b}\right|_p |\lambda p^n|_p \leq p^{-n}.$$

Pour finir, nous choisissons un entier α tel que $0 \leq \alpha \leq p^n - 1$ et $\alpha \equiv bc \pmod{p^n}$, ce qui est toujours possible. Donc, il existe un naturel $\mu < p^n$ tel que $\alpha = bc + \mu p^n$. Il vient

$$\begin{aligned} |x - \alpha|_p &\leq \max\left\{\left|x - \frac{a}{b}\right|_p, \left|\frac{a}{b} - \alpha\right|_p\right\} \\ &\leq \max\left\{\left|x - \frac{a}{b}\right|_p, \left|\frac{a}{b} - ac\right|_p, |\mu|_p\right\} \leq p^{-n}. \end{aligned}$$

□

De façon équivalente, les rationnels p-adiques peuvent être vu comme les expressions formelles de la forme

$$c_{-N}p^{-N} + \cdots + c_{-1}p^{-1} + c_0 + c_1 p + c_2 p^2 + \cdots$$

avec $c_j \in \{0, \ldots, p-1\}$, $N \in \mathbb{Z}$. Les entiers p-adiques sont identifiés avec les expressions formelles qui ne font intervenir que des puissances positives

de p. Ces expressions sont appelées *développements p-adiques* de nombres p-adiques. En particulier, une propriété intéressante de cette représentation des nombres p-adiques est que le développement p-adique d'un entier positif est toujours fini et correspond à sa décomposition habituelle en base entière p.

Dans [27], nous voyons que le champ $\mathbb{Q}_p$ n'est pas algébriquement clos. Nous pouvons prendre la clôture algébrique de $\mathbb{Q}_p$, notée $\bar{\mathbb{Q}}_p$, et étendre la valeur absolue $|\cdot|_p$ à cette clôture algébrique. Cependant, cette clôture algébrique $\bar{\mathbb{Q}}_p$ n'est pas complète. En la complétant et en étendant à nouveau $|\cdot|_p$, nous obtenons un champ complet algébriquement clos $\mathbb{C}_p$ avec valeur absolue $|\cdot|_p$, qui, restreinte sur $\mathbb{Q}$, est la valeur absolue p-adique. De plus, $\mathbb{C}_p$ est algébriquement clos, cf. [12].

Présentons quelques concepts d'analyse p-adique tels que la convergence de suites et de séries dans $\mathbb{Q}_p$ par rapport à la valeur absolue p-adique $|\cdot|_p$. Les limites par rapport à $|\cdot|_p$ sont notées $\lim^{(p)}$. Supposons que $(\alpha_n)_{n\geq 0}$ est une suite dans $\mathbb{Q}_p$.

Proposition 3.1.14. *La suite $(\alpha_n)_{n\geq 0}$ est de Cauchy dans $\mathbb{Q}_p$ si et seulement si $(\alpha_{n+1}-\alpha_n)_{n\geq 0}$ est une suite nulle.*

Démonstration. Soit $\delta > 0$. Supposons que $(\alpha_n)_{n\geq 0}$ est de Cauchy. Il existe donc $M \in \mathbb{N}$ tel que pour tous $m, n > M$, nous avons $|\alpha_m - \alpha_n|_p < \delta$. En particulier, pour tout $n > M$, $|\alpha_{n+1} - \alpha_n|_p < \delta$. Donc, nous avons

$$\lim_{n\to+\infty}{}^{(p)}(\alpha_{n+1}-\alpha_n) = 0$$

et $(\alpha_{n+1}-\alpha_n)_{n\geq 0}$ est une suite nulle.

Montrons l'autre direction et supposons que $(\alpha_{n+1}-\alpha_n)_{n\geq 0}$ est une suite nulle. Il existe alors $M \in \mathbb{N}$ tel que pour tout $n > M$, nous avons $|\alpha_{n+1} - \alpha_n|_p < \delta$. Pour tous $m \geq n > M$, nous obtenons

$$\begin{aligned} |\alpha_m - \alpha_n|_p &= |\alpha_m - \alpha_{m-1} + \alpha_{m-1} - \cdots - \alpha_{n+1} + \alpha_{n+1} - \alpha_n|_p \\ &\leq \max_{j\in\{n+1,\ldots,m\}}\{|\alpha_j - \alpha_{j-1}|_p\} \\ &< \delta \end{aligned}$$

puisque la valeur absolue p-adique est non-archimédienne. Donc, la suite $(\alpha_n)_{n\geq 0}$ est de Cauchy. □

Définition 3.1.15. Considérons la *n-ième somme partielle* de la série $\sum_{n\geq 0}\alpha_n$, $s_n = \alpha_0 + \cdots + \alpha_n$. Si la suite $(s_n)_{n\geq 0}$ dans $\mathbb{Q}_p$ a pour limite

$$S = \lim_{n\to+\infty}{}^{(p)} s_n,$$

alors la série $\sum_{n\geq 0}\alpha_n$ *converge vers la limite* S et nous écrivons

$$\sum_{n\geq 0}\alpha_n = S.$$

Si la série n'a pas de limite, alors elle *diverge*.

Remarque 3.1.16. Le résultat suivant diffère de ce qui se passe dans l'analyse réelle. Dans le cadre de l'analyse réelle, le fait que la suite $(\alpha_n)_{n\geq 0}$ est une suite nulle n'implique pas la convergence de la série $\sum_{n\geq 0}\alpha_n$. Prenons par exemple la série harmonique

$$\sum_{n\geq 1}\frac{1}{n}.$$

Son terme général $\alpha_n = \frac{1}{n}$ tend vers 0 lorsque $n \to +\infty$. Mais la suite des sommes partielles $(s_n)_{n\geq 0}$ n'est pas de Cauchy. En effet, nous avons pour $n \in \mathbb{N}$,

$$|s_{2n} - s_n| = \left|\sum_{j=n+1}^{2n}\frac{1}{j}\right| \geq \left|n\frac{1}{2n}\right| = \frac{1}{2}.$$

Comme $\mathbb{R}$ est complet, la suite des sommes partielles est de Cauchy si et seulement si la série converge. Donc, la série harmonique diverge.

Par contre, lorsqu'on se place dans $\mathbb{Q}_p$ (et même dans $\mathbb{C}_p$) en considérant la norme $|\cdot|_p$, alors la série harmonique est convergente vu la proposition suivante.

Proposition 3.1.17. *La série* $\sum_{n\geq 0}\alpha_n$ *dans* $\mathbb{Q}_p$ *converge si et seulement si* $(\alpha_n)_{n\geq 0}$ *est une suite nulle.*

Démonstration. Supposons que la série $\sum_{n\geq 0}\alpha_n$ converge. Donc la limite $\lim^{(p)}_{n\to+\infty} s_n$ existe. Par la proposition 3.1.5, la suite $(s_n)_{n\geq 0}$ est de Cauchy. Cela signifie que la suite $(s_{n+1} - s_n)_{n\geq 0}$ est nulle vu la proposition 3.1.14. Or nous avons $\alpha_n = s_n - s_{n-1}$. Donc la suite $(\alpha_n)_{n\geq 0}$ est une suite nulle.

L'autre direction utilise les mêmes arguments. Si la suite $(\alpha_n)_{n\geq 0}$ est une suite nulle, alors la suite $(s_{n+1} - s_n)_{n\geq 0}$ l'est aussi. Donc, $(s_n)_{n\geq 0}$ est de Cauchy vu la proposition 3.1.14. D'où la limite $\lim^{(p)}_{n\to+\infty} s_n$ existe et la série $\sum_{n\geq 0}\alpha_n$ converge. □

Donc, pour vérifier la convergence d'une série $\sum_{n\geq 0}\alpha_n$ dans $\mathbb{Q}_p$, il suffit de vérifier si $\lim^{(p)}_{n\to+\infty}\alpha_n = 0$, i.e., si $|\alpha_n|_p \to 0$ lorsque $n \to +\infty$.

3.2 A propos des groupes abéliens finiment engendrés

Comme dans la section précédente, nous rappelons des résultats utiles pour la démonstration du théorème 3.4.3 ci-dessous. Puisque l'objet de ce mémoire n'est pas de faire de l'algèbre, nous nous contenterons d'énoncer les résultats sans les démontrer. Pour de plus amples informations, le lecteur peut consulter [5].

Définition 3.2.1. Un *groupe abélien de torsion* est un groupe abélien G tel que chaque élément x de G a un ordre fini, i.e., pour chaque $x \in G$, il existe $n \in \mathbb{Z}$ tel que $nx = 0$ et $n \neq 0$, où 0 est le neutre de G. Un *groupe abélien sans torsion* est un groupe abélien G tel que le seul élément ayant un ordre fini est l'élément neutre.

Lemme 3.2.2. *Tout groupe abélien finiment généré de torsion est fini.*

Définition 3.2.3. Soient G un groupe abélien et $(g_i)_{i\in I}$ une famille d'éléments de G. Si chaque élément x de G peut être décomposé de façon unique comme $x = \sum_{i\in I} x_i g_i$, avec $x_i \in \mathbb{Z}$ pour tout i et avec $x_i = 0$ pour tout, sauf un nombre fini d'indices, i, alors $(g_i)_{i\geq 0}$ est une *base* de G. Un *groupe abélien libre* est un groupe abélien qui possède une base.

Lemme 3.2.4. *Tout groupe abélien libre est sans torsion.*

Lemme 3.2.5. *Tout groupe abélien libre finiment généré est isomorphe à $\mathbb{Z}^e$ pour un $e \in \mathbb{N}$. Dans ce cas, e est appelé le rang du groupe abélien libre.*

Théorème 3.2.6. *Tout groupe abélien finiment généré sans torsion est un groupe abélien libre.*

Théorème 3.2.7 (Théorème fondamental des groupes commutatifs finiment engendrés)**.** *Tout groupe abélien finiment engendré est isomorphe au produit direct $\mathbb{Z}^e \times T$ où T est un groupe abélien de torsion, pour un $e \in \mathbb{N}$.*

3.3 Suites linéaires récurrentes et déterminants de Hankel

Dans cette section, nous considérons des suites définies sur un champ arbitraire $\mathbb{K}$. Nous nous intéressons, plus particulièrement à une caractérisation des suites linéaires récurrentes et quelques propriétés de stabilité les concernant. Ce sujet est développé dans [23, 4].

Proposition 3.3.1. *Soit un entier $k > 1$. Si une suite $(U_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire à coefficients constants de la forme : pour tout $i \in \mathbb{N}$,*

$$U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \text{ avec } a_1, \ldots, a_k \in \mathbb{K} \text{ et } a_k \neq 0$$

et si $\alpha_1, \ldots, \alpha_r$ sont les racines de multiplicité $m_1, \ldots, m_r$ du polynôme caractéristique de la récurrence $\chi_U(x)$, alors il existe des polynômes Q_i de degré strictement inférieur à m_i, pour $i \in \{1, \ldots, r\}$, tels que

$$U_i = Q_1(i)\alpha_1^i + \cdots + Q_r(i)\alpha_r^i.$$

En particulier les polynômes $Q_1, \ldots, Q_r$ sont entièrement déterminés par les conditions initiales $U_0, \ldots, U_{k-1}$.

Théorème 3.3.2. *La suite $U = (U_i)_{i\geq 0} \in \mathbb{K}^{\mathbb{N}}$ satisfait la relation de récurrence linéaire d'ordre k :*

$$\forall n \in \mathbb{N}, U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \text{ où } a_1, \ldots, a_k \in \mathbb{K}$$

avec les conditions initiales $U_0, \ldots, U_{k-1}$ si et seulement si la série génératrice est la fraction rationnelle :

$$\mathbf{U}(x) := \sum_{i\geq 0} U_i x^i = \frac{\sum_{i=0}^{k-1} U_i x^i - \sum_{0<i+j<k} a_i U_j x^{i+j}}{1 - a_1 x - \cdots - a_k x^k}.$$

Démonstration. Nous avons

$$\begin{aligned}
\mathbf{U}(x) &= \sum_{i\geq 0} U_i x^i \\
&= \sum_{i=0}^{k-1} U_i x^i + \sum_{n\geq 0} U_{n+k} x^{n+k} \\
&= \sum_{i=0}^{k-1} U_i x^i + \sum_{n\geq 0} \left(\sum_{i=1}^{k} a_i U_{n+k-i} \right) x^{n+k} \\
&= \sum_{i=0}^{k-1} U_i x^i + \sum_{i=1}^{k} a_i x^i \sum_{n\geq 0} U_{n+k-i}\ x^{n+k-i} \\
&= \sum_{i=0}^{k-1} U_i x^i + \sum_{i=1}^{k} a_i x^i \left(\mathbf{U}(x) - \sum_{j=0}^{k-i-1} U_j x^j \right)
\end{aligned}$$

où nous avons utilisé le fait qu'il s'agit de sommations formelles et que nous pouvons sans problème permuter les différents symboles sommatoires. Nous obtenons donc

$$\left(1 - \sum_{i=1}^{k} a_i x^i\right) \mathbf{U}(x) = \sum_{i=0}^{k-1} U_i x^i - \sum_{i=1}^{k} \sum_{j=0}^{k-i-1} a_i U_j x^j$$
$$= \sum_{i=0}^{k-1} U_i x^i - \sum_{0<i+j<k} a_i U_j x^{i+j}.$$

□

Exemple 3.3.3. Considérons la suite $(U_i)_{i\geq 0}$ satisfaisant $U_0 = 1$, $U_1 = 2$ et la relation de récurrence linéaire $U_{i+2} = 3U_{i+1} - U_{i+1}$ pour tout $i \geq 0$. La série génératrice de $(U_i)_{i\geq 0}$ est

$$\begin{aligned}\mathbf{U}(x) &= \sum_{i\geq 0} U_i x^i \\ &= 1 + 2x + \sum_{i\geq 0} \underbrace{U_{i+2}}_{=3U_{i+1}+U_i} x^{i+2} \\ &= 1 + 2x + 3x \sum_{i\geq 0} U_{i+1} x^{i+1} - x^2 \sum_{i\geq 0} U_i x^i \\ &= 1 + 2x + 3x(\mathbf{U}(x) - 1) - x^2\, \mathbf{U}(x)\end{aligned}$$

Donc, nous avons $(1 - 3x + x^2)\, \mathbf{U}(x) = 1 + 2x - 3x$. Par conséquent, nous obtenons

$$\mathbf{U}(x) = \frac{1 - x}{x^2 - 3x + 1}.$$

Proposition 3.3.4. *Soient des entiers $p \geq 1$ et $\ell \geq 0$. Si la suite $U = (U_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire, alors il en est de même pour la sous-suite $U^{(\ell)} = (U_{pi+\ell})_{i\geq 0}$.*

Démonstration. Puisque la suite $U = (U_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire, sa série génératrice $\mathbf{U}(x)$ est une fraction rationnelle : $\mathbf{U}(x) = P(x)/Q(x)$ où $P(x)$ et $Q(x)$ sont des polynômes. Si w est une racine p-ième de l'unité, nous avons

$$\frac{1}{p} \sum_{t=0}^{p-1} \mathbf{U}(w^t x) = \frac{1}{p} \sum_{t=0}^{p-1} \sum_{i=0}^{\infty} U_i\, x^i$$
$$= \sum_{i=0}^{\infty} U_i (1 + w^i + w^{2i} + \cdots + w^{(p-1)i}) x^i.$$

Or, si i est un multiple de p, nous avons

$$1 + w^i + w^{2i} + \cdots + w^{(p-1)i} = p,$$

sinon

$$1 + w^i + w^{2i} + \cdots + w^{(p-1)i} = \frac{(w^i)^p - 1}{w^i - 1} = 0.$$

Donc, nous obtenons

$$\frac{1}{p} \sum_{t=0}^{p-1} \mathbf{U}(w^t x) = \sum_{i=0}^{\infty} U_{ip}\, z^{ip}.$$

D'un autre côté, nous avons la fraction rationnelle suivante

$$\frac{1}{p} \sum_{t=0}^{p-1} \mathbf{U}(w^t x) = \frac{1}{p} \sum_{t=0}^{p-1} \frac{P(w^t x)}{Q(w^t x)} =: \frac{A(x)}{B(x)}.$$

En réduisant au même dénominateur, nous obtenons $B(x) = \prod_{t=0}^{p-1} Q(w^t x)$. Notons $Q(x) = \alpha_0(z - \alpha_1) \cdots (x - \alpha_k)$ où les zéros de Q sont répétés selon leur multiplicité. Dans ce cas,

$$B(x) = \prod_{t=0}^{p-1} Q(w^t x) = \alpha_0 \prod_{t=0}^{p-1} \prod_{i=1}^{k} (w^t x - \alpha_i) = \alpha_0 \prod_{i=1}^{k} w^{p(p-1)/2} \prod_{t=0}^{p-1} (x - \frac{\alpha_i}{w^t}).$$

Nous observons que $w^{p(p-1)/2} = \pm 1$ selon la parité de p et que

$$\{1, 1/w, \ldots, 1/w^{p-1}\} = \{1, w, \ldots, w^{p-1}\}.$$

Nous pouvons alors écrire

$$B(x) = \alpha_0 \prod_{i=1}^{k} w^{p(p-1)/2} \prod_{t=0}^{p-1} (x - w^t \alpha_i) = \alpha_0 \prod_{i=1}^{k} w^{p(p-1)/2} (z^p - \alpha_i^p)$$

car les p racines p-ième de α_i^p sont exactement $\alpha_i, w\alpha_i, \ldots, w^{p-1}\alpha_i$. Par conséquent, le polynôme $B(x)$ a tous ses termes de degré multiple de p. Comme nous avons

$$\frac{A(x)}{B(x)} = \frac{1}{p} \sum_{t=0}^{p-1} \mathbf{U}(w^t x) = \sum_{i=0}^{\infty} U_{ip} z^{ip},$$

le polynôme $A(x) = B(x) \sum_{i=0}^{\infty} U_{ip} z^{ip}$ a lui aussi tous ses termes de degré multiple de p. Nous pouvons donc considérer sans problème la fraction rationnelle $A(x^{1/p})/B(x^{1/p})$ où $A(x^{1/p})$ et $B(x^{1/p})$ sont des polynômes en x.

De plus, cette fraction rationnelle est égale à la série génératrice de la suite $(U_{pi})_{i\geq 0}$,

$$\frac{A(x^{1/p})}{B(x^{1/p})} = \sum_{i\geq 0} U_{ip} x^i.$$

Donc la suite $(U_{pi})_{i\geq 0}$ satisfait une relation de récurrence linéaire. Pour conclure, il suffit de remarquer que si une suite $(V_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire, alors la suite $(V_{i+\ell})_{i\geq 0}$ satisfait exactement la même relation, seules les conditions initiales changent. □

Définition 3.3.5. Soient $U = (U_i)_{i\geq 0}$ et $V = (V_i)_{i\geq 0}$ deux suites sur $\mathbb{K}$. La *somme* de U et V est la suite $(U_i + V_i)_{i\geq 0}$. Pour tout $c \in \mathbb{K}$, la *multiplication par c* de U est la suite $(cU_i)_{i\geq 0}$. Le *produit de Hadamard* de U et V est la suite $(U_i V_i)_{i\geq 0}$. Le *produit de Cauchy* de U et V est la suite $(\sum_{j=0}^{i} U_j V_{i-j})_{i\geq 0}$.

Proposition 3.3.6. *L'ensemble des suites linéaires récurrentes est fermé pour la somme, la multiplication par un scalaire et le produit de Cauchy.*

Démonstration. Soient deux suites $U = (U_i)_{i\geq 0}$ et $V = (V_i)_{i\geq 0}$ sur $\mathbb{K}$ qui satisfont des relations de récurrence linéaire. Notons S_1 (resp S_2) la série formelle de puissance de U (resp. V). Il existe des polynômes P_1, P_2, Q_1, Q_2 tels que $S_1 = P_1/Q_1$ et $S_2 + P_2/Q_2$.

Vu le théorème 3.3.2, il suffit de montrer que les séries correspondant à la somme de U et V, la multiplication par un scalaire $c \in \mathbb{K}$ de U, le produit de Hadamard et le produit de Cauchy de U et V, sont encore des fonctions rationnelles.

Nous avons directement

$$S_1 + S_2 = \frac{(P_1Q_2 + P_2Q_1)}{Q_1Q_2}, cS_1 = \frac{cP_1}{Q_1} \text{ et } S_1S_2 = \frac{P_1P_2}{Q_1Q_2}.$$

La conclusion en découle. □

Proposition 3.3.7. *L'ensemble des suites linéaires récurrentes est fermé pour le produit de Hadamard.*

Nous pouvons trouver une démonstration de la proposition précédente par M.-P. Schützenberger dans [28].

Définition 3.3.8. Soit $U = (U_i)_{i\geq 0}$ une suite. Nous définissons les déterminants de Hankel, pour $n \in \mathbb{N}$,

$$D_n^{(k+1)} := \det \begin{pmatrix} U_n & \cdots & U_{n+k} \\ \vdots & & \vdots \\ U_{n+k} & \cdots & U_{n+2k} \end{pmatrix}.$$

Lemme 3.3.9. *Soient $(U_i)_{i\geq 0}$ une suite et un entier $n \in \mathbb{N}$. Si $D_n^{(k)} = 0$ et $D_n^{(k+1)} = 0$, alors nous avons $D_{n+1}^{(k)} = 0$.*

Le résultat suivi est parfois connu comme le *théorème de Kronecker*.

Théorème 3.3.10. *Une suite non nulle $(U_i)_{i\geq 0} \in \mathbb{K}^{\mathbb{N}}$ satisfait une relation de récurrence linéaire à coefficients dans $\mathbb{K}$ d'ordre minimal k si et seulement si $D_0^{(k)} \neq 0$ et pour tout $m \geq k+1$, $D_0^{(m)} = 0$.*

Exemple 3.3.11. Considérons la suite

$$(U_i)_{i\geq 0} = (100, 200, 1, 1, 2, 3, 5, 8, 13, \cdots)$$

satisfaisant la relation de récurrence linéaire d'ordre 4

$$U_{i+4} = U_{i+3} + U_{i+2} + 0U_{i+1} + 0U_i \qquad \forall i \geq 0.$$

Remarquons que, pour tout $i > 1$, la suite satisfait une relation de récurrence d'ordre 2 : $U_{i+2} = U_{i+1} + U_i$. Donc, la suite translatée $(U_{i+2})_{i\geq 0}$ satisfait une relation de récurrence d'ordre 2. Vérifions que l'ordre minimal de la relation de récurrence satisfaite par $(U_i)_{i\geq 0}$ est 4. Procédons par l'absurde et supposons que il existe $a, b, c \in \mathbb{Z}$ tels que, pour tout $i \geq 0$,

$$U_{i+3} = aU_{i+2} + bU_{i+1} + CU_i.$$

Dans ce cas, nous avons

$$\begin{cases} 1 = a + 200b + 100c \\ 2 = a + b + 200c \\ 3 = 2a + b + c \\ 5 = 3a + 3b + c \end{cases}.$$

Puisque le système n'a pas de solution, $(U_i)_{i\geq 0}$ ne satisfait pas de relation de récurrence d'ordre 3 (ni d'ordre 2).

Calculons les déterminants de Hankel de cette suite. Nous avons

$$D_0^{(2)} = \det \begin{pmatrix} 100 & 200 \\ 200 & 1 \end{pmatrix} = -39900$$

$$D_0^{(3)} = \det \begin{pmatrix} 100 & 200 & 1 \\ 200 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} = -79501$$

$$D_0^{(4)} = \det \begin{pmatrix} 100 & 200 & 1 & 1 \\ 200 & 1 & 1 & 2 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 5 \end{pmatrix} = -40000$$

$$D_0^{(5)} = \det \begin{pmatrix} 100 & 200 & 1 & 1 & 2 \\ 200 & 1 & 1 & 2 & 3 \\ 1 & 1 & 2 & 3 & 5 \\ 1 & 2 & 3 & 5 & 8 \\ 2 & 3 & 5 & 8 & 13 \end{pmatrix} = 0$$

Vu le théorème 3.3.10, cela signifie que la suite $(U_i)_{i\geq 0}$ ne peut satisfaire aucune relation d'ordre inférieur à 4 et, a priori, la suite peut satisfaire une relation d'ordre 4.

3.4 Théorème de caractérisation

Dans cette section, nous caractérisons les suites linéaires récurrentes U telles que $\lim_{m\to+\infty} N_U(m) = +\infty$ en utilisant les résultats vus dans les sections précédentes. Pour ce faire, vu le lemme 2.4.1, il suffit de se concentrer sur le comportement de $N_U(p^v)$ pour tout nombre premier p qui divise a_k. Tout au long de cette section, nous supposons que $(U_i)_{i\geq 0}$ est une suite linéaire récurrente satisfaisant une relation de récurrence linéaire de la forme (1.1)

$$\forall n \in \mathbb{N}, U_{i+k} = a_1 U_{i+k-1} + \cdots + a_k U_i, \text{ où } a_1, \ldots, a_k \in \mathbb{Z}\,.$$

De plus, nous supposons que $U = (U_i)_{i\geq 0}$ ne satisfait pas de récurrence d'ordre inférieur à k. Vu le théorème 3.3.10, sous l'hypothèse $a_k \neq 0$, supposer que $(U_i)_{i\geq 0}$ ne satisfait pas de récurrence d'ordre inférieur à k, est équivalent à supposer que k est le plus grand entier satisfaisant

$$\det \begin{pmatrix} U_0 & \cdots & U_{k-1} \\ \vdots & & \vdots \\ U_{k-1} & \cdots & U_{2k-2} \end{pmatrix} \neq 0.$$

Notons $P_U(x)$ le polynôme défini comme

$$P_U(x) := x^k \chi_U(1/x) = 1 - a_1 x - \cdots - a_k x^k,$$

où $\chi_U(x)$ est polynôme caractéristique de $U = (U_i)_{i\geq 0}$. Puisque nous supposons $a_k \neq 0$, nous remarquons que, si $\alpha_1, \ldots, \alpha_s$ sont les racines de χ_U, alors les inverses $1/\alpha_1, \ldots, 1/\alpha_s$ sont exactement les racines de P_U.

Remarque 3.4.1. Vu le théorème 3.3.2, il existe un polynôme $Q(x)$ telle que la fonction génératrice de $(U_i)_{i\geq 0}$ est

$$\sum_{i\geq 0} U_i x^i = \frac{Q(x)}{P_U(x)}.$$

Puisque nous supposons que la relation de récurrence linéaire est d'ordre minimal k, les polynômes $Q(x)$ et $P_U(x)$ sont premiers entre eux. Donc, les pôles de $\sum_{i\geq 0} U_i x^i$ sont exactement les racines de $P_U(x)$, i.e., les inverses des racines de $\chi_U(x)$.

Remarque 3.4.2. Puisque la suite $(U_i)_{i\geq 0}$ satisfait une relation de récurrence linéaire, par la proposition 3.3.1, nous avons $U_i = \sum_{j=1}^{s} q_j(i)\alpha_j^i$ où les q_j sont des polynômes de $\mathbb{C}_p[x]$ et les α_j sont les racines du polynôme caractéristique de la suite $(U_i)_{i\geq 0}$, i.e., les inverses des pôles de $\mathbf{U}(x)$. Dans ce cas, nous obtenons

$$U_i^{(b)} = U_{ia+b} = \sum_{j=1}^{s} q_j(ai+b)\alpha_j^{ai+b} = \sum_{j=1}^{s} \alpha_j^b q_j(ai+b)(\alpha_j^a)^i.$$

Donc, les pôles de la série de puissances $\mathbf{U}^{(b)}(x)$ de $(U_i^{(b)})_{i\geq 0}$ sont les a-ièmes puissances des pôles de $\mathbf{U}(x)$.

Théorème 3.4.3. *Nous avons $N_U(p^v) \not\to +\infty$ lorsque $v \to +\infty$ si et seulement si $P_U(x) = A(x)B(x)$ avec $A(x)$, $B(x) \in \mathbb{Z}[x]$ tels que :*

i. $B(x) \equiv 1 \pmod{p\,\mathbb{Z}[x]}$,

ii. *$A(x)$ n'a pas de racine multiple et toutes ses racines sont racines de l'unité.*

Dans ce cas, nous avons en plus $A(0) = B(0) = 1$.

Avant d'en faire la démonstration, nous rappelons la notion de corps de décomposition.

Définition 3.4.4. Soient un champ L et un polynôme monique $f \in L[x]$. Une extension K du champ L est appelée un *corps de décomposition de* f si K est la plus petite extension de champ telle que f se scinde en facteurs linéaires dans $K[x]$.

Démonstration du théorème 3.4.3. Montrons que la condition est suffisante. Supposons qu'il existe une telle factorisation de $P_U(x)$. Vu (ii), il existe un naturel d tel que $A(x)$ divise $(x^d - 1)$. Vu le théorème 3.3.2, il existe un polynôme $Q(x)$ tel que $\sum_{i\geq 0} U_i x^i = \frac{Q(x)}{P_U(x)}$. De ces deux observations, nous déduisons qu'il existe un polynôme $R(x)$ tel que

$$(x^d - 1)\sum_{i\geq 0} U_i x^i = \frac{(x^d-1)Q(x)}{P_U(x)} = \frac{(x^d-i)Q(x)}{A(x)B(x)} = \frac{R(x)}{B(x)}.$$

Par l'hypothèse (i), il existe un polynôme à coefficients entiers $B_1(x)$ tel que $B(x) = 1 - pB_1(x)$. D'où, nous avons

$$(x^d - 1)\sum_{i\geq 0} U_i x^i = \frac{R(x)}{1 - pB_1(x)} = \sum_{i\geq 0} p^i R(x) B_1(x)^i.$$

En particulier, pour tout entier fixé v, $(x^d - 1)\sum_{i\geq 0} U_i x^i$ est congruent à un polynôme (mod p^v). Ceci signifie que $U_{i+d} \equiv U_i \pmod{p^v}$ pour tout i suffisamment grand car

$$(x^d - 1)\sum_{i\geq 0} U_i x^i = -\sum_{i=0}^{d-1} U_i x^i + \sum_{i\geq 0}(U_i - U_{i+d})x^{i+d}.$$

Par conséquent, il y a au plus d valeurs qui peuvent être prises infiniment souvent dans la suite $(U_i \bmod p^v)_{i\geq 0}$, i.e., $N_U(p^v) \leq d$ pour tout v. Donc, nous avons $N_U(p^v) \not\to +\infty$ lorsque $v \to +\infty$.

Montrons maintenant que la condition est nécessaire. Cette direction nécessite plus de travail et emploie des méthodes p-adiques.

La fonction $v \mapsto N_U(p^v)$ est croissante, i.e.,

$$N_U(p^w) \geq N_U(p^v) \text{ lorsque } w \geq v.$$

En particulier, si $N_U(p^v) \not\to +\infty$, alors il existe un naturel d tel que $N_U(p^v) = d$ pour tout naturel v suffisamment grand. Donc, nous pouvons choisir des entiers $a_{1,v}, \ldots, a_{d,v}$ tels que, si $U_i \equiv a \pmod{p^v}$ pour un nombre infini de i, alors $a \equiv a_{j,v} \pmod{p^v}$ pour un $j \in \{1, \ldots, d\}$. Puisque nous avons

$$\{(a_{1,w} \bmod p^v), \ldots, (a_{d,w} \bmod p^v)\} = \{(a_{1,v} \bmod p^v), \ldots, (a_{d,v} \bmod p^v)\}$$

pour tout $w \geq v$, nous pouvons supposer, sans perte de généralité, que

$$a_{j,w} \equiv a_{j,v} \pmod{p^v} \text{ pour } w \geq v, 1 \leq j \leq d.$$

Par conséquent, pour $1 \leq j \leq d$, la suite $(a_{j,v})_{v \geq 1}$ est de Cauchy dans $\mathbb{Z}_p$. En effet, puisque $a_{j,w} \equiv a_{j,v} \pmod{p^v}$ pour $w \geq v$, il existe des entiers k, l et $t \in \{0, \ldots, p^v - 1\}$ tels que $a_{j,v} = kp^v + t$ et $a_{j,w} = \ell p^v + t$. Nous obtenons alors

$$|a_{j,v} - a_{j,w}|_p = |(k - \ell)p^v|_p = p^{-v}.$$

Donc il existe $b_1, \ldots, b_d \in \mathbb{Z}_p$ tels que pour $1 \leq j \leq d$,

$$a_{j,v} \to b_j \text{ lorsque } v \to +\infty.$$

Posons

$$V_i = \prod_{j=1}^{d} (U_i - b_j) \in \mathbb{Z}_p\,.$$

Vu la proposition 3.3.6, la suite $(V_i)_{i \geq 0}$ satisfait une récurrence linéaire sur $\mathbb{Z}_p$. De plus, par construction, V_i est finalement dans $p\,\mathbb{Z}_p$ pour tout v fixé. En effet, chaque valeur de U_i qui n'est pas congrue à un des $b_1, \ldots, b_d \bmod p^v$ n'apparaît qu'un nombre fini de fois, vu les définitions des b_j, $1 \leq j \leq d$. Ceci signifie qu'il existe un $k \in \mathbb{Z}_p$ qui peut s'écrire $k = p^\ell k'$ tel que $V_i = p^v k = p^{k+\ell} k'$. D'où nous avons $|V_i|_p = p^{-(v+\ell)} \leq p^{-v}$. Donc, pour tout v fixé, $|V_i|_p \leq p^v$. Ceci implique $|V_i|_p \to 0$ lorsque $i \to +\infty$. Puisque $(V_i)_{i \geq 0}$ satisfait une relation de récurrence linéaire, la série

$$\mathbf{V}(x) := \sum_{i \geq 0} V_i x^i$$

est une série de puissances rationnelle dans $\mathbb{Q}_p(x)$. De plus, puisque $|V_i|_p \to 0$ lorsque $i \to +\infty$, $|V_i x^i|_p = |V_i|_p |x|_p^i \to 0$ lorsque $i \to +\infty$ et $|x|_p \leq 1$. Donc, vu la proposition 3.1.17, $\mathbf{V}(x)$ converge sur le disque unité fermé

$$\{x \in \mathbb{C}_p : |x|_p \leq 1\}.$$

Comme $\mathbf{V}(x)$ est une série rationnelle et converge sur $\mathbb{Z}_p$, ses pôles $\beta_1, \ldots,$ $\beta_r \in \mathbb{C}_p$ doivent satisfaire

$$|\beta_j|_p > 1 \text{ pour } 1 \leq j \leq r.$$

Pour finir la preuve, nous utilisons le lemme 3.4.5. Il reste encore à vérifier que les pôles de $\mathbf{U}(x) = \sum_{i \geq 0} U_i x^i$ génèrent un sous-groupe libre abélien. En général, les pôles de $\mathbf{U}(x)$ génèrent un sous-groupe abélien finiment engendré

de $\mathbb{C}_p^\times$. Par le théorème fondamental des groupes abéliens finiment engendrés 3.2.7, ce groupe est isomorphe à $\mathbb{Z}^e \times T$, où T est un groupe abélien fini et où $e \geq 0$ est un entier.

Montrons comment se débarrasser du groupe de torsion T pour pouvoir appliquer le lemme 3.4.5. Soit $a = \#T$. Pour $0 \leq b < a$, au lieu de prendre la suite $(U_i)_{i\geq 0}$, considérons la suite $(U_i^{(b)})_{i\geq 0} := (U_{ai+b})_{i\geq 0}$. Par la proposition 3.3.4, cette suite satisfait une récurrence linéaire. Vu la remarque 3.4.2, les pôles de la fonction génératrice $\mathbf{U}^{(b)}(x)$ de $(U_i^{(b)})_{i\geq 0}$ sont les a-ièmes puissances des pôles de $\mathbf{U}(x)$. Par conséquent, les pôles de $\mathbf{U}^{(b)}(x)$ génèrent un groupe abélien finiment engendré sans torsion vu le choix de a. De plus, ce groupe est nécessairement un groupe abélien libre vu le théorème 3.2.6.

Puisque les pôles $\beta_1, \ldots, \beta_r \in \mathbb{C}_p$ de $\mathbf{V}(x)$ satisfont $|\beta_j|_p > 1$ pour $1 \leq j \leq r$, avec le même raisonnement, nous obtenons que les pôles de la fonction rationnelle $\mathbf{V}^{(b)}(x) = \sum_{i\geq 0} V_{ai+b}x^i$ sont les a-ièmes puissances de $\beta_1, \ldots, \beta_r$ et pour $1 \leq j \leq r$, $|\beta_j^a|_p = |\beta_j|_p^a > 1$.

Nous pouvons maintenant appliquer le lemme 3.4.5 à la suite $(U_i^{(b)})_{i\geq 0}$ et déduire que tout pôle de $\gamma \in \mathbb{C}_p$ de $\mathbf{U}^{(b)}(x)$ satisfait soit $|\gamma|_p > 1$, soit $\gamma = 1$. Les pôles distincts de $\mathbf{U}(x)$ sont les $1/\alpha_1, \ldots, 1/\alpha_s$, avec $\alpha_1, \ldots, \alpha_s \in \mathbb{C}_p$. Comme ce sont les a-ièmes racines des pôles de $\mathbf{U}^{(b)}(x)$, tout pôle $1/\alpha_j$ est tel que $|1/\alpha_j|_p > 1$ ou tel qu'il est racine de l'unité. Vu la minimalité de l'ordre k de la récurrence satisfaite par $(U_i)_{i\geq 0}$ et la remarque 3.4.1, les pôles de $\mathbf{U}(x)$ sont précisément les racines de $P_U(x)$.

Donc, nous pouvons factoriser

$$P_U(x) = (1 - \delta_1 x) \cdots (1 - \delta_k x)$$

avec tout δ_j égal à un des $\alpha_1, \ldots, \alpha_s$ éventuellement répétés. Factorisons $P_U(x)$ comme $A(x)B(x)$ en posant

$$A(x) = \prod_{j:|\delta_j|_p=1} (1 - \delta_j x) \text{ et } B(x) = \prod_{j:|\delta_j|<1} (1 - \delta_j x).$$

Alors, $P_U(x) \in \mathbb{Z}[x]$. De plus, si K est un corps de décomposition de $P_U(x)$ sur $\mathbb{Q}$, alors tout automorphisme de K doit permuter l'ensemble des δ_j tels que $|\delta_j|_p < 1$, puisque l'automorphisme doit permuter l'ensemble des δ_j et envoyer les racines de l'unité sur elles-mêmes. Donc, $B(x)$ est un polynôme rationnel car il est fixé par tout automorphisme de K. Remarquons que pour $n > 0$, le coefficient de x^n dans $B(x)$ est la somme des produits de n éléments de $\{\delta_j : |\delta_j|_p < 1\}$. L'ensemble des entiers algébriques est un sous-anneau de $\mathbb{C}_p$ et les seuls rationnels qui sont des entiers algébriques sont en fait entiers. Puisque les δ_j sont des entiers algébriques, $B(x)$ est un polynôme à coefficients entiers. De plus, puisque la valeur absolue p-adique

est non-archimédienne, le coefficient de x^n dans $B(x)$, $n > 0$, a une valeur absolue p-adique strictement inférieure à 1. Observons qu'un entier m tel que $|m|_p < 1$ est nécessairement un multiple de p. Donc, $B(x) \equiv 1 \pmod{p\,\mathbb{Z}[x]}$.

Intéressons-nous maintenant au polynôme $A(x)$. Ses racines sont racines de l'unité. De plus, par le même raisonnement que ci-dessus, $A(x) \in \mathbb{Z}[x]$.

Il reste à montrer que $A(x)$ n'a pas de racine multiple. Pour ce faire, nous montrons que les pôles de $\mathbf{U}(x)$, qui sont racines de l'unité, sont simples. Rappelons que $1/\alpha_1, \ldots, 1/\alpha_s$ sont les pôles distincts de $\mathbf{U}(x)$. Supposons qu'il existe un entier $0 \leq t \leq s$ tel que $\alpha_1, \ldots, \alpha_t$ sont racines de l'unité et, pour $j = t+1, \ldots, s$, $|\alpha_j|_p < 1$. Alors, vu la proposition 3.3.1, il existe des polynômes $q_j \in \mathbb{C}_p[x]$, avec $j = 1, \cdots, s$ tels que, pour tout i,

$$U_i = \sum_{j=1}^{s} q_j(i)\alpha_j^i = \underbrace{\sum_{j=1}^{t} q_j(i)\alpha_j^i}_{:=T_i} + \underbrace{\sum_{j=t+1}^{s} q_j(i)\alpha_j^i}_{:=W_i}.$$

Puisque pour $j = T+1, \ldots, s$, $|\alpha_j|_p < 1$, nous avons

$$|U_i - T_i|_p = |W_i|_p \to 0 \text{ lorsque } i \to +\infty.$$

En effet, pour tout j, l'ensemble $\{|q_j(i)|_p : i \in \mathbb{N}\}$ est borné par une constante. Comme pour $1 \leq j \leq t$, α_j est racine de l'unité, il existe un naturel a tel que $\alpha_j^a = 1$ pour tout $j \in \{1, \ldots, t\}$. Comme précédemment, posons $T_i^{(b)} = T_{ai+b}$ pour $0 \leq b < a$. Alors,

$$T_i^{(b)} = \sum_{j=1}^{t} q_j(ai+b)\alpha_j^{ai+b} = \sum_{j=1}^{t} \alpha_j^b q_j(ai+b)$$

est un polynôme avec coefficient dans $\mathbb{C}_p$. Nous le notons $g_b(i)$.

Soit $\varepsilon > 0$. Pour tout i suffisamment grand, il existe $\ell(i) \in \{1, \ldots, d\}$ tel que $|U_i - b_{\ell(i)}|_p < \varepsilon$. En effet, par définition des b_j, nous avons pour tout v suffisamment grand et tout entier i suffisamment grand,

$$|U_i - b_{\ell(i)}|_p \leq |U_i - a_{\ell(i)}, v|_p + |a_{\ell(i)}, v - b_{\ell(i)}|_p \leq p^{-v} + \varepsilon_v$$

où $\varepsilon_v \to 0$ lorsque $v \to +\infty$.

Puisque $|U_i - T_i|p \to 0$ quand $i \to +\infty$, nous avons

$$\begin{aligned} 0 \leq |(T_i - b_1)\cdots(T_i - b_d)|_p &= \left|\prod_{j=1}^{d}(T_i - U_i + U_i - b_{\ell(i)} + b_{\ell(i)} - b_j)\right|_p \\ &= \prod_{j=1}^{d}|T_i - U_i + U_i - b_{\ell(i)} + b_{\ell(i)} - b_j|_p \\ &\leq \prod_{j=1}^{d}|T_i - U_i|_p + |U_i - b_{\ell(i)}|_p + |b_{\ell(i)} - b_j|_p \to 0 \end{aligned}$$

lorsque $i \to +\infty$, car chaque facteur est borné par une constante et un facteur tend vers zéro. Donc, pour $0 \leq b < a$, comme $T_i^{(b)} = g_b(i)$, on a

$$|(g_b(i) - b_1)\cdots(g_b(i) - b_d)|_p \to 0 \text{ lorsque } i \to +\infty.$$

Considérons le polynôme défini par

$$h_b(i) := (g_b(i) - b_1)\cdots(g_b(i) - b_d).$$

Soit n_0 un naturel. Nous avons

$$|h_b(n_o + p^v) - h_b(n_0)|_p \leq Cp^v \text{ lorsque } v \to +\infty$$

où C est le maximum des valeurs absolues p-adiques des coefficients de h_b. En effet, si le polynôme $h_b(i)$ s'écrit $h_b(i) = c_d i^d + \cdots + c_1 i + c_0$, alors

$$|h_b(n_o + p^v) - h_b(n_0)|_p = c_d((n_o + p^v)^d - n_0^d) + \cdots + c_1((n_o + p^v) + n_0)$$

et chacun des termes entre parenthèses contient au moins un facteur p^v par le théorème binomial, ce qui suffit pour déduire l'inégalité précédente. Puisque $|h_b(i)|_p \to 0$ lorsque $i \to +\infty$, nous avons $h_b(n_0) = 0$ car

$$0 \leq |h_b(n_0)|_p \leq |h_b(n_0) - h_b(n_0 + p^v)|_p + |h_b(n_0 + p^v)|_p \to 0$$

lorsque $v \to +\infty$. Donc, pour tout entier i, chaque $h_b(i) = 0$, avec $0 \leq b < a$. En conséquence du théorème fondamental d'algèbre, chaque g_b, avec $0 \leq b < a$, prend infiniment souvent la même valeur parmi $b_1, \ldots, b_d$. Donc, chaque g_b, avec $0 \leq b < a$, est un polynôme constant. Puisque $T_i^{(b)} = g_b(i)$ est constant, la suite des T_i est périodique :

$$\forall i \geq 0, T_{i+a} = T_i.$$

D'où,

$$|U_{i+a} - U_i|_p \leq |T_{i+a} - T_i|_p + |W_{i+a} - W_i|_p \to 0$$

lorsque $i \to +\infty$. Vu la remarque 3.1.17, $(x^a - 1)\,\mathbf{U}(x) = \sum_{i\geq 0}(U_{i+a} - U_i)x^i$ converge sur le disque unité fermé $\{x \in \mathbb{C}_p : |x|_p \leq 1\}$. Par conséquent, $(x^a - 1)\,\mathbf{U}(x)$ n'a pas de pôles sur le disque unité fermé et en particulier, sur le cercle unité. Or $\mathbf{U}(x) = Q(x)/P_U(x)$, avec Q et P_U premiers entre eux, par minimalité de l'ordre de la récurrence satisfaite par $(U_i)_{i\geq 0}$. Donc $A(x)$ divise $(x^a - 1)$ puisque $(x^a - 1)Q(x)/P_U(x)$ n'a pas de racine sur le cercle unité. Ceci implique que $A(x)$ n'a pas de racine multiple. □

Lemme 3.4.5. *Soit $(U_i)_{i\geq 0}$ une suite d'entiers linéaire récurrente satisfaisant (1.1). Soit $\mathbf{U}(x) = \sum_{i\geq 0} U_i x^i \in \mathbb{Z}_p[[x]]$ la série rationnelle de puissances correspondante. Supposons que le sous-groupe multiplicatif de $\mathbb{C}_p^\times$ généré par les pôles (en nombre fini) de $\mathbf{U}(x)$ est un groupe abélien libre. Soit $V_i = \prod_{j=1}^d (U_i - b_j)$ pour $i \geq 0$ avec $b_1, \ldots, b_d \in \mathbb{Z}_p$. Si la série rationnelle de puissances $\mathbf{V}(x) = \sum_{i\geq 0} U_i x^i \in \mathbb{Z}_p[[x]]$ a pour racines $\beta_1, \ldots, \beta_r \in \mathbb{C}_p$ satisfaisant $|\beta_j|_p > 1$ pour $1 \leq j \leq r$, alors chaque pôle $\gamma \in \mathbb{C}_p$ de $\mathbf{U}(x)$ est tel que soit $|\gamma|_p > 1$, soit $\gamma = 1$.*

Démonstration. Soient $1/\alpha_1, \ldots, 1/\alpha_s$ les pôles disctincts de $\mathbf{U}(x)$, avec $\alpha_1, \ldots, \alpha_s \in \mathbb{C}_p$. Remarquons que 0 ne peut pas être un pôle de $\mathbf{U}(x)$. Nous voulons montrer tout d'abord que $|\alpha_j| \leq 1$ pour $1 \leq j \leq s$. Pour ce faire, observons que pour être un pôle, chaque $1/\alpha_j$ doit satisfaire $P_U(1/\alpha_j) = 0$, car $\mathbf{U}(x) = Q(x)/P_U(x)$, où $Q(x)$ est un polynôme. Ce qui signifie que

$$P_U(\frac{1}{\alpha_j}) = 1 - \frac{a_1}{\alpha_j} - \cdots - \frac{a_k}{\alpha_j^k} = 0.$$

Par conséquent, pour tout $j \in \{1, \ldots, s\}$, nous avons

$$\left|\frac{a_1}{\alpha_j} + \cdots + \frac{a_k}{\alpha_j^k}\right|_p = |1|_p = 1.$$

Puisque $|\cdot|_p$ est non-archimédienne, nous avons $|a_\ell/\alpha_j^\ell|_p \geq |\frac{a_1}{\alpha_j} + \cdots + \frac{a_k}{\alpha_j^k}|_p = 1$ pour un $\ell \in \{1, \ldots, k\}$. Donc, $|\alpha_j|_p^\ell \leq |a_\ell|_p$. Puisque $a_\ell \in \mathbb{Z}$, $|a_\ell|_p \leq 1$. Ce qui implique $|\alpha_j|_p \leq 1$ pour $1 \leq j \leq s$. Nous savons que les pôles de $\mathbf{U}(x)$ ont une valeur absolue p-adique plus grande ou égale à 1. Nous pouvons supposer qu'il existe $0 \leq t \leq s$ tel que $|\alpha_1|_p = \cdots = |\alpha_t|_p = 1$ et $|\alpha_j|_p < 1$ pour $t < j \leq s$.

Il existe des polynômes $q_1(x), \cdots, q_s(x) \in \mathbb{C}_p[x]$ tels que

$$U_i = \sum_{j=1}^{s} q_j(i)\alpha_j^i.$$

De plus, nous définissons, pour $i \geq 0$,

$$V_i := \prod_{j=1}^{d}(U_i - b_j) = c_d U_i^d + c_{d-1}U_i^{d-1} + \cdots + c_0$$

pour des $c_0, \cdots, c_{d-1} \in \mathbb{Z}_p$ et $c_d = 1$. Par le théorème multinomial, nous obtenons

$$\begin{aligned} V_i &= \sum_{j=0}^{d} c_j \sum_{\substack{j_1+\cdots+j_s=j \\ j_1,\ldots,j_s\geq 0}} \binom{j}{j_1,\ldots,j_s} \prod_{l=1}^{s}(q_\ell(i)\alpha_\ell^i)^{j_\ell} \\ &= \sum_{j=0}^{d} c_j \sum_{\substack{j_1+\cdots+j_s=j \\ j_1,\ldots,j_s\geq 0}} \binom{j}{j_1,\ldots,j_s} \prod_{l=1}^{s} q_\ell(i)^{j_\ell} \left(\prod_{\ell=1}^{s}\alpha_\ell^{j_\ell}\right)^i \end{aligned}$$

Puisque les racines du polynôme caractéristique sont les inverses des pôles de la série de puissance rationnelle correspondante, il suit que l'ensemble des pôles de $\mathbf{V}(x) = \sum_{i\geq 0} V_i x^i$ est inclus dans l'ensemble

$$\left\{\prod_{\ell=1}^{s}\alpha_\ell^{j_\ell} : j_1,\ldots,j_s \geq 0,\, j_1+\cdots+j_s \leq d\right\}.$$

Par hypothèse, les pôles de $\mathbf{V}(x)$ ont tous des valeurs absolues strictement plus grande que 1. Remarquons que

$$\left|\prod_{\ell=1}^{r}\alpha_\ell^{j_\ell}\right|_p = \left|\prod_{\ell=1}^{t}\alpha_\ell^{j_\ell}\right|_p \left|\prod_{\ell=t+1}^{r}\alpha_\ell^{j_\ell}\right|_p > 1$$

si et seulement si $j_\ell > 0$ pour un $\ell > t$. Par conséquent, nous pouvons conclure que les pôles possibles de $\mathbf{V}(x)$ sont ceux qui ne contiennent pas uniquement les $\alpha_1, \ldots, \alpha_t$. Par exemple, α_1^d ne peut être un pôle de $\mathbf{V}(x)$.

Notons G le sous-groupe multiplicatif de $\mathbb{C}_p^\times$ généré par $\alpha_1, \ldots, \alpha_t$. Par hypothèse, G est un sous-groupe d'un groupe abélien finiment engendré. Donc, $G \cong \mathbb{Z}^e$ pour un naturel $e \geq 0$. Pour conclure la preuve, il suffit de montrer que $e = 0$ car, dans ce cas, nous déduisons que le seul pôle possible γ de $\mathbf{U}(x)$ tel que $|\gamma|_p = 1$ est 1.

Supposons $e > 0$. Soient $\gamma_1, \ldots, \gamma_e$ les générateurs d'un groupe abélien libre de rang e. Alors, pour $1 \leq i \leq t$, nous pouvons écrire

$$\alpha_i = \prod_{j=1}^{e}\gamma_j^{b_{i,j}},$$

où les $b_{i,j}$ sont des entiers. Nous renommons les $b_{i,j}$ si nécessaire de sorte que

– $|b_{1,1}| = \max\{|b_{i,1}|\} > 0$;
– $|b_{1,j}| = \max\{|b_{i,j} : b_{i,\ell} = b_{1,\ell} \text{ pour } \ell < j\}$ pour $j \in \{2, \ldots, e\}$.

Par construction, α_1^d ne peut pas être écrit comme un mot différent dans $\alpha_1, \ldots, \alpha_t$ de longueur au plus d. Alors, dans l'expression de $\mathbf{V}_i$ ci-dessus, il y a une occurrence de

$$c_d(q_1(i))^d(\alpha_1^d)^i$$

qui ne peut être annulée par aucun autre pôle vu notre choix de α_1. Il en découle que α_1^d devrait être un pôle de $\mathbf{V}$, ce qui est une contradiction. Donc, $e = 0$ et le résultat suit. □

Exemple 3.4.6. Considérons la suite linéaire récurrente $U = (U_i)_{i\geq 0}$ donnée par

$$U_{i+4} = 3U_{i+3} + 2U_{i+2} + 3U_i \text{ pour } i \in \mathbb{N} \text{ et } U_i = i + 1 \text{ pour } i \in \{0, \ldots, 3\}.$$

Comme montré dans [11], l'addition dans ce système de numération linéaire n'est pas calculable par un automate fini. Cependant, nous pouvons montrer que $N_U(3^v) \to +\infty$ si $v \to +\infty$, en appliquant le théorème 3.4.3. Nous avons $\chi_U(x) = x^4 - 3x^3 - 2x^2 - 3$ et donc $P_U(x) = 1 - 3x - 2x^2 - 3x^4$. Nous pouvons facilement vérifier que $P_U(x)$ ne se factorise pas comme $A(x)B(x)$ avec $A(x)$ et $B(x)$ satisfaisant les hypothèse du théorème 3.4.3. Donc, nous pouvons appliquer que notre procédure de décision peut s'appliquer à des systèmes de numération qui ne peuvent pas être traités par [11].

Exemple 3.4.7. Considérons la suite linéaire récurrente $U = (U_i)_{i\geq 0}$ satisfaisant

$$U_{i+5} = 6U_{i+4} + 3U_{i+3} - U_{i+2} + 6U_{i+1} + 3U_i, \forall i \geq 0.$$

Nous avons

$$\chi_U(x) = x^5 - 6x^4 - 3x^3 + x^2 - 6x - 3$$

et

$$P_U(x) = 1 - 6x - 3x^2 + x^3 - 6x^4 - 3x^5 = \underbrace{(x^3 + 1)}_{A(x)}\underbrace{(-3x^2 - 6x + 1)}_{B(x)}.$$

Avec les conditions initiales $U_i = i + 1$ pour $i \in \{0, \ldots, 4\}$, vu le théorème 3.3.10, la suite U ne satisfait pas de relation de récurrence linéaire d'ordre inférieur à 5 puisque

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 54 \\ 3 & 4 & 5 & 54 & 359 \\ 4 & 5 & 54 & 359 & 2344 \\ 5 & 54 & 359 & 2344 & 15129 \end{pmatrix} = 8458240 \neq 0.$$

Même si le pgcd des coefficients de la relation de récurrence vaut 1, nous avons $N_U(3^v) \not\to +\infty$ lorsque $v \to +\infty$ car $P_U(x)$ ne satisfait pas les hypothèses du théorème 3.4.3. Le tableau suivant reprend les premières valeurs de $N_U(3^v)$. Pour rappel, il suffit d'inspecter les premières valeurs de la suite $(U_i \bmod 3^v)_{i\geq 0}$ à la recherche de deux k-uples d'éléments consécutifs qui sont identiques pour trouver la période.

v	période	$N_U(3^v)$
1	(1,0,1,2,0,2)	3
2	(4,0,1,5,0,8)	5
3	(22,9,19,5,18,8)	6
4	(49,63,19,32,18,62)	6
5	(211,225,19,32,18,224)	6
6	(224,697,468,505,32,261)	6
⋮	⋮	⋮

Chapitre 4

Une procédure de décision pour une classe de systèmes de numération abstraits

4.1 Introduction

Définition 4.1.1. Un *système de numération abstrait* est la donnée d'un triplet

$$S = (L, \Sigma, <)$$

où L est un langage régulier infini sur un alphabet totalement ordonné $(\Sigma, <)$. Pour tout naturel n, $\mathrm{rep}_S(n)$ dénote le $(n+1)$-ième mot de L selon l'ordre généalogique induit par l'ordre $<$ sur Σ. Nous appelons $\mathrm{rep}_S(n)$ la *S-représentation* de n. Nous obtenons ainsi une bijection $\mathrm{rep}_S : \mathbb{N} \to L$ de $\mathbb{N}$ sur L. L'application réciproque, qui à un mot $w \in L$ associe son indice dans le langage L ordonné selon l'ordre généalogique, est notée $\mathrm{val}_S(w)$. Nous appelons val_S la *valeur S-numérique* de w.

Définition 4.1.2. Un ensemble $X \subseteq \mathbb{N}$ d'entiers est *S-reconnaissable* si le langage $\mathrm{rep}_S(X)$ sur Σ est régulier.

Remarquons que par définition, pour tout système de numération abstrait, $\mathbb{N}$ est S-reconnaissable.

Exemple 4.1.3. Considérons le système de numération abstrait

$$S = (a^*b^*, \{a, b\}, a < b).$$

Si nous énumérons les premiers mots du langage a^*b^*, nous avons

ε	0	bb	5	$aaaa$	10
a	1	aaa	6	$aaab$	11
b	2	aab	7	$aabb$	12
aa	3	abb	8	$abbb$	13
ab	4	bbb	9	$bbbb$	14

Nous obtenons, par exemple, $\mathrm{rep}_S(6) = aaa$ et $val_S(ab) = 4$.

Définition 4.1.4. Soit $S = (L, \Sigma, <)$ un système de numération abstrait construit sur un langage régulier infini L dont l'automate minimal est

$$\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L).$$

Nous notons $\mathbf{u}_i(q)$ (resp. $\mathbf{v}_i(q)$) le nombre de mots de longueur i (resp. ≤ 1) accepté à partir de $q \in Q_L$ dans $\mathcal{M}_L$.

Définition 4.1.5. Soit $\mathcal{A} = (Q, q_0, \Sigma, \delta, F)$ un automate fini déterministe. La *matrice d'adjacence* de $\mathcal{A}$ est la matrice donnée par

$$M_{q,r} = \#\{\sigma \in \Sigma \mid \delta(q, \sigma) = r\}, \qquad q, r \in Q.$$

Remarque 4.1.6. Observons que les suites $(\mathbf{u}_i(q))_{i\geq 0}$ satisfont la même relation de récurrence linéaire homogène pour tout $q \in Q_L$. En effet, pour $q \in Q_L$, la suite $(\mathbf{u}_i(q))_{i\geq 0}$ satisfait une relation de récurrence linéaire homogène d'ordre k dont le polynôme caractéristique est le polynôme caractéristique de la matrice d'adjacence de $\mathcal{M}_L$. Une preuve de ce résultat se trouve dans [24].

Exemple 4.1.7. Considérons l'automate représenté à la figure 4.1. Sa matrice d'adjacence est

$$M = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Nous savons que le puits de l'automate n'influence pas la suite $(\mathbf{u}_i(q_0))_{i\geq 0}$. Donc, nous pouvons regarder la matrice correspondante

$$M' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Calculons son polynôme caractéristique

$$\det(M' - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix}.$$

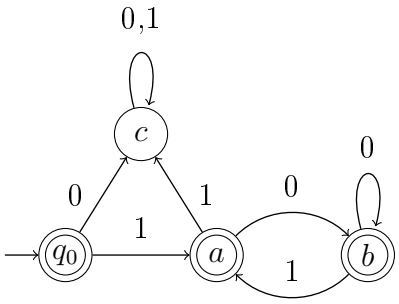

FIGURE 4.1 – Automate acceptant le langage $\mathrm{rep}_F(\mathbb{N})$.

D'où, nous avons

$$\det(M' - \lambda I) = -\lambda(-\lambda(1-\lambda) - 1) = -\lambda^3 + \lambda^2 + \lambda.$$

Par le théorème de Cayley-Hamilton, cf. [24], nous obtenons $0 = -M'^3 + M'^2 + M'$. D'où, il vient $M'^{i+3} = M'^{i+2} + M'^{i+1}$. De là, nous déduisons que la suite $(\mathbf{u}_i(q_0))_{i\geq 0}$ satisfait la relation de récurrence linéaire

$$\mathbf{u}_{i+3}(q_0) = \mathbf{u}_{i+2} + \mathbf{u}_{i+1} \, \forall i \geq 0.$$

De plus, nous remarquons que $\mathbf{u}_0(q_0) = 1$, $\mathbf{u}_1(q_0) = 1$, $\mathbf{u}_2(q_0) = 1$ et $\mathbf{u}_3(q_0) = 2$. Par conséquent, $(\mathbf{u}_{i+2}(q_0))_{i\geq 0} = (F_i)_{i\geq 0}$ où les F_i sont les nombres de Fibonacci.

Remarque 4.1.8. Nous remarquons aussi que les suites $(\mathbf{v}_i(q))_{i\geq 0})$ satisfont la même relation de récurrence linéaire homogène pour tout $q \in Q_L$. Puisque $(\mathbf{u}_i(q))_{i\geq 0}$ satisfait une relation de récurrence linéaire homogène d'ordre k, il existe $a_1, \ldots, a_k \in \mathbb{Z}$ tels que pour tout $i \geq 0$,

$$\mathbf{u}_{i+k}(q) = a_1 \, \mathbf{u}_{i+k-1}(q) + \cdots + a_k \, \mathbf{u}_i(q).$$

Donc, pour tout $i \geq 0$, nous avons

$$\begin{aligned} \mathbf{v}_{i+k-1}(q) - \mathbf{v}_{i+k}(q) &= \mathbf{u}_{i+k+1}(q) \\ &= a_1(\mathbf{v}_{i+k}(q) - \mathbf{v}_{i+k-1}(q)) + \cdots + a_k(\mathbf{v}_{i+1}(q) - \mathbf{v}_i(q)). \end{aligned}$$

Par conséquent, la suite $(\mathbf{v}_i(q))_{i\geq 0}$ satisfait une relation de récurrence linéaire homogène d'ordre $k+1$.

Les systèmes de numération abstraits sont une généralisation des systèmes de numération positionnels $U = (U_i)_{i\geq 0}$ pour lesquels $\mathbb{N}$ est U-reconnaissable.

Exemple 4.1.9. Considérons le langage $L = \{\varepsilon\} \cup 1\{0, 01\}^*$ et supposons $0 < 1$. Les premiers mots ordonnés généalogiquement sont

$$\varepsilon < 1 < 10 < 100 < 101 < 1000 < 1001 < \cdots$$

Le système de numération obtenu $S = (L, \{0, 1\}, 0 < 1)$ est le système de Fibonacci défini par la suite $F = (F_i)_{i \geq 0}$ où les F_i sont les nombres de Fibonacci. En effet,

$$\mathrm{rep}_S(1) = 1 = \mathrm{rep}_F(1), \mathrm{rep}_S(2) = 10 = \mathrm{rep}_F(2), \mathrm{rep}_S(3) = 100 = \mathrm{rep}_F(3), \ldots$$

Exemple 4.1.10. Prenons le système de numération positionnel en base 2. Lorsque les zéros de tête sont interdits, l'ensemble de toutes les représentations est $\mathrm{rep}_U(\mathbb{N}) = 1\{1, 0\}^* \cup \{\varepsilon\}$. C'est donc un langage régulier. Considérons le système de numération abstrait composé de $\Sigma = \{a, b\}$, $a < b$ et $L = b\{a, b\}^* \cup \{\varepsilon\}$. Pour obtenir les représentations habituelles, définissons un morphisme h par $h(a) = 0$ et $h(b) = 1$. Alors $h(L) = \mathrm{rep}_U(\mathbb{N})$. Remarquons que, si nous autorisons les zéros de tête dans les représentations, alors les mots 0, 00, et 000 ont des valeurs numériques différentes.

$h(b\{a, b\}^* \cup \{\varepsilon\})$	$\mathbb{N}$	$h(\{a, b\}^*)$
ε	0	ε
1	1	0
10	2	1
11	3	00
100	4	01
101	5	10
110	6	11
111	7	000

L'exemple suivant montre que la classe des systèmes de numération abstraits est strictement plus grande que la classe des systèmes de numération linéaires pour lesquels $\mathbb{N}$ est reconnaissable

Exemple 4.1.11. Considérons l'alphabet $\Sigma = \{a, b, c, d\}$, avec $a < b < c < d$, et le langage $L = \{\varepsilon\} \cup \{a, ab\}^* \cup \{c, cd\}^*$. Si nous ordonnons les premiers mots de L, nous obtenons

ε	0	cc	5	ccc	10	$aaba$	15
a	1	cd	6	ccd	11	$abaa$	16
c	2	aaa	7	cdc	12	$abab$	17
aa	3	aab	8	$aaaa$	13	$cccc$	18
ab	4	aba	9	$aaab$	14	$cccd$	19

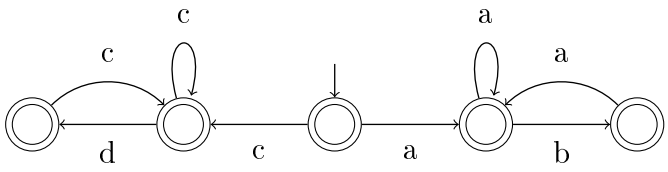

FIGURE 4.2 – Automate admettant L.

Nous remarquons qu'il n'y a pas de bijection ν entre $\{a, b, c, d\}$ et un ensemble d'entiers conduisant à un système de numération positionnel linéaire. Autrement dit, a, b, c, d ne peuvent pas être identifiés aux "chiffres" habituels. En effet, supposons qu'il existe une suite $U = (U_i)_{i\geq 0}$ d'entiers telle que, pour tout mot $x_1 \cdots x_n \in L$, avec $x_i \in \{a, b, c, d\}$ pour tout i,

$$\mathrm{val}_U(\nu(x_1) \cdots \nu(x_n)) = \mathrm{val}_S(x_1 \cdots x_n).$$

Puisque $1 = \mathrm{val}_S(a) = val_U(\nu(a)) = \nu(a)U_0$ et $2 = \mathrm{val}_S(c) = \nu(c)U_0$, nous obtenons $U_0 = 1$, $\nu(a) = 1$ et $\nu(c) = 2$. De plus, de $3 = \mathrm{val}_S(aa) = \nu(a)U_1 + \nu(a)U_0$, nous déduisons $U_1 = 2$. Alors, $\mathrm{val}_U(\nu(c)\nu(c)) = 2U_1 + 2U_0 = 6$ mais $\mathrm{val}_S(cc) = 5$, ce qui est une contradiction.

Calculons les valeurs de $\mathbf{u}_i(q_{0,L})$ et $\mathbf{v}_i(q_{0,L})$. Pour tout $i \geq 1$, nous avons

$$\mathbf{u}_0(q_{0,L}) = 1 \text{ et } \mathbf{u}_i(q_{0,L}) = 2F_i \text{ pour tout } i \geq 1$$

où F_i est le i-ième nombre de Fibonacci. Donc, pour $i \geq 1$, nous avons

$$\mathbf{v}_i(q_{0,L}) = \sum_{n=0}^{i} \mathbf{u}_n(q_{0,L}) = 1 + 2\sum_{n=1}^{i} F_n.$$

De plus, nous observons pour $i \geq 1$

$$\mathbf{v}_i(q_{0,L}) - \mathbf{v}_{i-1}(q_{0,L}) = 2F_i = \mathbf{u}_i(q_{0,L}).$$

Par définition de la suite de Fibonacci, nous obtenons pour $i \geq 3$

$$\mathbf{v}_i(q_{0,L}) - \mathbf{v}_{i-1}(q_{0,L}) = (\mathbf{v}_{i-1}(q_{0,L}) - \mathbf{v}_{i-2}(q_{0,L})) + (\mathbf{v}_{i-2}(q_{0,L}) - \mathbf{v}_{i-3}(q_{0,L})).$$

D'où, il vient pour $i \geq 3$

$$\mathbf{v}_i(q_{0,L}) = 2\,\mathbf{v}_{i-1}(q_{0,L}) - \mathbf{v}_{i-3}(q_{0,L}),$$

$$\text{avec } \mathbf{v}_0(q_{0,L}) = 1, \mathbf{v}_1(q_{0,L}) = 3, \mathbf{v}_2(q_{0,L}) = 7.$$

Puisque les systèmes de numération abstraits sont une généralisation des systèmes de numération positionnels $U = (U_i)_{i\geq 0}$ pour lesquels $\mathbb{N}$ est U-reconnaissable, il est naturel de s'intéresser au problème de décision suivant qui est analogue au problème 1, avec quelques hypothèses supplémentaires sur les systèmes de numération abstraits.

Problème 2. Étant donné un système de numération abstrait S et un ensemble $X \subseteq \mathbb{N}$ tel que $\mathrm{rep}_S(X)$ est accepté par un automate fini (déterministe), pouvons-nous décider si X est ou non un ensemble ultimement périodique, i.e., si X est ou non une union finie de progressions arithmétiques ?

4.2 Calcul de val_S et S-reconnaissabilité d'ensembles ultimement périodiques

Dans cette section, étant donné un système de numération abstrait $S = (L, \Sigma, <)$, nous montrons comment calculer la fonction val_S. De plus, nous prouvons que toute progression arithmétique $a + \mathbb{N}\, b$ est S-reconnaissable, cf [19, 26].

Définition 4.2.1. Soient $S = (L, \Sigma, <)$ un système de numération abstrait et $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_l, F_L)$ l'automate minimal de L. Un état $q \in Q_L$ est de la forme $w^{-1}.L$ si et seulement si $q = \delta_L(q_{0,L}, w)$. Nous écrivons souvent $q'.w$ au lieu de $\delta_L(q', w)$. Dans ce cas, $w^{-1}.L$ est l'ensemble L_q des mots acceptés par $\mathcal{M}_L$ à partir de q, i.e.,

$$w^{-1}.L = L_q = \{x \in \Sigma^* : \delta_L(q, x) \in F_L\}.$$

En particulier, nous avons $L_{q_{0,L}} = L$.

Chaque $q \in Q_L$ pour lequel L_q est infini donne lieu à un système de numération $S_q = (L_q, \Sigma, <)$. Les fonctions rep_{S_q} et val_{S_q} sont simplement notées rep_q et val_q si le contexte est clair. Dans le cas où L_q est fini, les fonctions rep_q et val_q sont définies comme dans le cas infini mais leur domaine est restreint à $\{0, \ldots, \#L_q - 1\}$.

Lemme 4.2.2. *Soit $s = (L, \Sigma, <)$ un système de numération abstrait. Si $xy \in L_q$, avec $x, y \in \Sigma \setminus \{\varepsilon\}$, alors*

$$\mathrm{val}_q(xy) = \mathrm{val}_{q.x}(y) + \mathbf{v}_{|xy|-1}(q) - \mathbf{v}_{|y|-1}(q.x) + \sum_{\substack{x'<x \\ |x'|=|x|}} \mathbf{u}_{|y|}(q.x').$$

Démonstration. Nous devons calculer le nombre de mots de L_q qui sont généalogiquement strictement inférieur à xy. Il y en a de trois sortes. Premièrement, les mots de longueurs strictement inférieures à $|xy|$. Il y en a $\mathbf{v}_{|xy|}(q)$ par définition. La deuxième sorte contient les mots de longueurs exactement $|xy|$ admettant comme préfixe x. Puisque nous avons $x'y' \in L_q$ si et seulement si $y' \in L_{q.y'}$, nous voyons qu'il y en a $\mathrm{val}_{q.x}(y) - \mathbf{v}_{|y|-1}(q.x)$.

Enfin, les mots de la dernière sorte sont ceux qui ont la même longueur que xy et qui ne commence pas par x. Leur nombre est

$$\#\{w \in L_q : w = x'y', |x'| = |x|, |y'| = |y|, x' < x\} = \sum_{\substack{x'<x \\ |x'|=|x|}} \mathbf{u}_{|y|}(q.x').$$

La conclusion suit directement. □

Comme nous le montrons dans le lemme suivant, dans un système de numération abstrait, les différentes suites $(\mathbf{u}_i(q))_{i\geq 0}$, pour $q \in Q_L$, jouent le rôle de l'unique suite $(U_i)_{i\geq 0}$ définissant un système de numération positionnel comme dans la définition 1.3.1.

Lemme 4.2.3. *Soit $w = \sigma_1 \cdots \sigma_n \in L$. Nous avons*

$$\mathrm{val}_S(w) = \sum_{q\in Q_L} \sum_{i=1}^{|w|} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \tag{4.1}$$

où

$$\beta_{q,i}(w) := \#\{\sigma < \sigma_i \mid \delta_L(q_{0,L}, \sigma_1 \cdots \sigma_{i-1}\sigma) = q\} + \mathbf{1}_{q,q_{0,L}} \tag{4.2}$$

pour $i = 1, \ldots, |w|$.

Rappelons que

$$\mathbf{1}_{q,q'} = \begin{cases} 1 & \text{si } q = q' \\ 0 & \text{si } q \neq q'. \end{cases}$$

Démonstration. Écrivons simplement $q_{0,L} = q_0$ pour ne pas alourdir les notations. Appliquons le lemme 4.2.2 au mot $w = \sigma_1 \cdots \sigma_n$ en prenant $x = \sigma_1$ et $y = \sigma_2 \cdots \sigma_n$. Nous obtenons

$$\mathrm{val}_{q_0}(w) = \mathrm{val}_{q_0.\sigma_1}(\sigma_2 \cdots \sigma_n) + \mathbf{v}_{n-1}(q_0) - \mathbf{v}_{n-2}(q_0.\sigma_1) + \sum_{\substack{x'<\sigma_1 \\ |x'|=1}} \mathbf{u}_{n-1}(q_0.x').$$

En l'appliquant à nouveau pour calculer $\mathrm{val}_{q_0.\sigma_1}(\sigma_2\cdots\sigma_n)$ avec $x=\sigma_2$ et $y=\sigma_3\cdots\sigma_n$, nous obtenons

$$\begin{aligned}\mathrm{val}_{q_0}(w) &= \mathrm{val}_{q_0.\sigma_1\sigma_2}(\sigma_3\cdots\sigma_n) + \mathbf{v}_{n-2}(q_0.\sigma_1) - \mathbf{v}_{n-3}(q_0.\sigma_1\sigma_2)\\ &\quad + \sum_{\substack{x'<\sigma_2\\ |x'|=1}} \mathbf{u}_{n-2}(q_0.\sigma_1 x') + \mathbf{v}_{n-1}(q_0) - \mathbf{v}_{n-2}(q_0.\sigma_1) + \sum_{\substack{x'<\sigma_1\\ |x'|=1}} \mathbf{u}_{n-1}(q_0.x').\end{aligned}$$

En appliquant le lemme 4.2.2 plusieurs fois et en simplifiant le résultat obtenu, nous avons

$$\begin{aligned}\mathrm{val}_{q_0}(w) &= \mathbf{v}_{n-1}(q_0) + \sum_{\substack{x'<\sigma_1\\ |x'|=1}} \mathbf{u}_{n-1}(q_0.x') + \cdots + \sum_{\substack{x'<\sigma_{n-1}\\ |x'|=1}} \mathbf{u}_1(q_0.\sigma_1\cdots\sigma_{n-2}x')\\ &\quad + \mathbf{v}_0(q_0.\sigma_1\cdots\sigma_{n-1}) + \mathrm{val}_{q_0.\sigma_1\cdots\sigma_{n-1}}(\sigma_n)\end{aligned}$$

Remarquons que si σ est une lettre de L_q, alors nous avons

$$val_q(\sigma) = \mathbf{u}_0(q) + \sum_{\substack{x'<\sigma\\ |x'|=1}} \mathbf{u}_0(q.x').$$

Nous obtenons donc

$$\mathrm{val}_{q_0}(w) = \mathbf{v}_{n-1}(q_0) + \sum_{\substack{x'<\sigma_1\\ |x'|=1}} \mathbf{u}_{n-1}(q_0.x') + \cdots + \sum_{\substack{x'<\sigma_n\\ |x'|=1}} \mathbf{u}_1(q_0.\sigma_1\cdots\sigma_{n-1}x').$$

Puisque $\mathbf{v}_{n-1}(q_0) = \mathbf{u}_0(q_0) + \cdots + \mathbf{u}_{n-1}(q_0)$, en réorganisant les sommes en fonction des états, il vient

$$\mathrm{val}_S(w) = \mathrm{val}_{q_0}(w) = \sum_{q\in Q_L}\sum_{i=1}^{|w|} \beta_{q,i}(w)\,\mathbf{u}_{|w|-i}(q).$$

□

Proposition 4.2.4. *Soit $S=(L,\Sigma,<)$ un système de numération abstrait. Soient des entiers $a,b\geq 0$. La progression arithmétique $a+\mathbb{N}\,b$ est S-reconnaissable.*

Démonstration. Nous pouvons supposer $a<b$. Notons, comme d'habitude, $\mathcal{M}_L=(Q_L,q_{0,L},\Sigma,\delta_L,F_L)$ l'automate minimal de L. Montrons que l'automate minimal de $rep_S(a+\mathbb{N}\,b)$ a un nombre fini d'états. Ses états sont les ensembles

$$w^{-1}.\mathrm{rep}_S(a+\mathbb{N}\,b) = \{x\in\Sigma^* : \mathrm{val}_S(wx)\equiv a \pmod b\} \text{ pour } w\in\Sigma^*.$$

Or, nous avons

$$\mathrm{val}_S(wx) = \mathrm{val}_{q_{0,L}.w}(x) + \mathbf{v}_{|wx|-1}(q_{0,L}) - v_{|x|-1}(q_{0,L}.w) + \sum_{\substack{w'<w \\ |w'|=|w|}} \mathbf{u}_{|x|}(q_{0,L}.w').$$

Nous savons que les suites $(\mathbf{u}_n(q))_{n\geq 0}$ (resp. $(\mathbf{v}_n(q))_{n\geq 0}$), pour $q \in Q_L$, satisfont une relation de récurrence linéaire à coefficients dans $\mathbb{Z}$. Donc, $(\mathbf{v}_n(q_{0,L}) \bmod b)_{n\geq 0}$ est ultimement périodique de période $\pi_{\mathbf{v}}(b)$. Pour $|w|$ suffisament grand, nous avons

$$\mathbf{v}_{|wv|-1}(q_{0,L}) \equiv \mathbf{v}_{|x|+i}(q_{0,L}) \pmod b \text{ pour un } j \in \{0, \ldots, \pi_{\mathbf{v}}(b)\}.$$

Puisque le langage L est régulier, $q_{0,L}.w$ prend un nombre fini de valeurs dans Q_L. De plus, nous avons

$$\sum_{\substack{w'<w \\ |w'|=|w|}} \mathbf{u}_{|x|}(q_{0,L}.w') \equiv \sum_{q'\in Q_L} j_{q'}\, \mathbf{u}_{|x|}(q') \text{ pour des } j_{q'} \in \{0, \ldots, b-1\}.$$

Donc, pour $|w|$ suffisamment grand, les états $w^{-1}.\mathrm{rep}_S(a + \mathbb{N}\, b$ sont de la forme

$$\{x \in \Sigma^* : \mathrm{val}_q(x) + \mathbf{v}_{|x|+i}(q_{0,L}) - v_{|x|-1}(q) + \sum_{q'\in Q_L} j_{q'}\, \mathbf{u}_{|x|}(q') \equiv a \pmod b\}$$

pour un $q \in Q_L$, des $j_{q'} \in \{0, \ldots, b-1\}$ et un $i \in \{0, \ldots, \pi_{\mathbf{v}}(b)\}$. Ces ensembles sont en nombre fini. Donc, le langage $\mathrm{rep}_S(a + \mathbb{N}\, b)$ est régulier. □

Le résultat suivant est une conséquence directe de la proposition 4.2.4.

Proposition 4.2.5. *Soit $S = (L, \Sigma, <)$ un système de numération abstrait construit sur un langage régulier infini L sur Σ. Tout ensemble X ultimement périodique est S-reconnaissable et un AFD acceptant $\mathrm{rep}_S(X)$ peut être effectivement obtenu.*

4.3 Bornes sur la période et prépériode d'un ensemble ultimement périodique

Définition 4.3.1. Soit un automate fini déterministe $\mathcal{A} = (Q, q_0, F, \Sigma, \delta)$. L'automate $\mathcal{A}$ est *accessible* si pour tout état $q \in Q$, il existe un mot $w \in \Sigma^*$ tel que $\delta(q_0, w) = q$. L'automate $\mathcal{A}$ est *coaccessible* si pour tout état $q \in Q$, il existe un mot $u \in \Sigma^*$ tel que $\delta(q, u) \in F$. Enfin, l'automate $\mathcal{A}$ est dit *emondé* s'il est accessible et coaccessible.

Proposition 4.3.2. *Soit $S = (L, \Sigma, <)$ un système de numération abstrait tel que, pour tous les états q de l'automate minimal émondé $\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ de L, nous avons*

$$\lim_{i \to +\infty} \mathbf{u}_i(q) = +\infty$$

et $\mathbf{u}_i(q_{0,L}) > 0$ pour tout $i \geq 0$. Si $X \subseteq \mathbb{N}$ est un ensemble ultimement périodique de période p_X, alors tout AFD acceptant $\mathrm{rep}_S(X)$ a au moins $\lceil N_{\mathbf{v}}(p_X)/\#Q_L \rceil$ états, où $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$.

Démonstration. Soit a_X la prépériode de X. Puisque, pour tout état q de $\mathcal{M}_L$, nous avons $\lim_{i \to +\infty} \mathbf{u}_i(q) = +\infty$, il existe une constante minimale $J > 0$ telle que, pour tout $i \geq J$ et tout état $q \in Q_L$, nous avons $\mathbf{u}_i(q) \geq p_X$. Considérons pour tout $i \in \mathbb{N}$, le mot

$$w_i = \mathrm{rep}_S(\mathbf{v}_i(q_{0,L}))$$

correspondant au premier mot de longueur $i+1$ dans le langage L ordonné généalogiquement. Par conséquent, pour tout entier $i \geq J-1$, le mot w_i peut être factorisé comme $w_i = x_i y_i$, avec $|y_i| = J$ et nous définissons $q_i := \delta_L(q_{0,L}, x_i)$. Remarquons que y_i est le plus petit mot de longueur J accepté à partir de q_i. Par définition de J, à partir de chaque q_i avec $i \geq J-1$, il y a au moins p_X mots de longueur J menant à un état final. Ordonnons les généalogiquement et notons les p_X premiers mots par

$$y_i = y_{i,0} < y_{i,1} < \cdots < y_{i,p_X - 1}.$$

Remarquons que, pour tout $t \in \{0, \ldots, p_X - 1\}$, nous avons

$$\mathrm{val}_S(x_i y_{i,t}) = \mathrm{val}_S(x_i y_i) + t = \mathbf{v}_i(q_{0,L}) + t.$$

Puisque la suite $\mathbf{v} := (\mathbf{v}_i(q_{0,L}))_{i \geq 0}$ satisfait une relation de récurrence linéaire, la suite $(\mathbf{v}_i(q_{0,L}) \bmod p_X)_{i \geq 0}$ est ultimement périodique. Elle prend infiniment souvent $N_{\mathbf{v}}(p_X) =: N$ valeurs distinctes. Soient $h_1, \ldots, h_N \geq J-1$ des entiers tels que, pour tous $i, j \in \{1, \ldots, N\}$, nous avons

$$\mathbf{v}_{h_i}(q_{0,L}) \geq a_X \text{ et } (i \neq j \Rightarrow \mathbf{v}_{h_i}(q_{0,L}) \not\equiv \mathbf{v}_{h_j}(q_{0,L}) \pmod{p_X}).$$

En particulier, nous avons $\mathrm{rep}_S(\mathbf{v}_{h_i}(q_{0,L}) = w_{h_i} = x_{h_i} y_{h_i}$ et $q_{h_i} = \delta_L(q_{0,L}, x_{h_i})$. Les éléments de l'ensemble $\{q_{h_1}, \ldots, q_{h_N}\}$ peuvent prendre au maximum $\#Q_L$ valeurs distinctes. Donc, au moins $\sigma := \lceil N/\#Q_L \rceil$ d'entre eux sont égaux. Sans perte de généralité, nous pouvons supposer que les éléments égaux sont les $q_{h_1}, \ldots, q_{h_\sigma}$. Pour tous $i, j \in \{1, \ldots, \sigma\}$ et tout $t \in \{0, \ldots, p_X - 1\}$, nous avons donc $b_{h_i,t} = b_{h_j,t}$. Pour tous $i, j \in \{1, \ldots, \sigma\}$ tels que $i \neq j$, par le lemme 2.1.17, il existe $t_{i,j} < p_X$ tel que

– soit $\mathbf{v}_{h_i}(q_{0,L}) + t_{i,j} \in X$ et $\mathbf{v}_{h_j}(q_{0,L}) + t_{i,j} \notin X$,
– soit $\mathbf{v}_{h_i}(q_{0,L}) + t_{i,j} \notin X$ et $\mathbf{v}_{h_j}(q_{0,L}) + t_{i,j} \in X$.

Par conséquent, les mots x_{h_i} et x_{h_j} n'appartiennent pas à la même classe d'équivalence pour $\sim_{\mathrm{rep}_S(X)}$. En effet, nous le montrons en concaténant le mot $y_{h_i,t_{i,j}} = y_{h_j,t_{i,j}}$ comme dans les preuves des propositions 2.2.3 et 2.2.5. Donc, vu la définition 1.1.9, l'automate minimal de $\mathrm{derep}_S(X)$ a au moins $\sigma = \lceil N/\#Q_L \rceil$ états. □

Corollaire 4.3.3. *Soit $S = (L, \Sigma, <)$ un système de numération abstrait ayant les mêmes propriétés que dans la proposition 4.3.2. Supposons que la suite $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i\geq 0}$ est telle que*

$$\lim_{m\to+\infty} N_{\mathbf{v}}(m) = +\infty.$$

Alors la période d'un ensemble ultimement périodique $X \subseteq \mathbb{N}$ tel que $\mathrm{rep}_S(X)$ est accepté par un AFD à d états est bornée par le plus petit entier s_0 tel que pour tout $m \geq s_0$, $N_{\mathbf{v}}(m) > d\#Q_L$, où Q_L est l'ensemble des états de l'automate minimal (émondé) de L.

Démonstration. Vu la proposition 4.3.2, nous savons que $d \geq N_U(p_X)/\#Q_L$ car d est plus grand ou égal au nombre d'états de l'automate minimal acceptant $\mathrm{rep}_S(X)$. La conclusion en découle. □

Définition 4.3.4. Soient deux automates finis déterministes

$$\mathcal{A} = (Q^{(a)}, q_0^{(a)}, \Sigma, \delta^{(a)}, F^{(a)}) \text{ et } \mathcal{B} = (Q^{(b)}, q_0^{(b)}, \Sigma, \delta^{(b)}, F^{(b)}).$$

Nous définissons l'automate produit $\mathcal{P} = \mathcal{A} \times \mathcal{B}$ comme suit
– l'ensemble fini d'états est $Q^{(a)} \times Q^{(b)}$,
– l'état initial est $(q_0^{(a)}, q_0^{(b)})$,
– l'ensemble des états finals est $F^{(a)} \times F^{(b)}$

et la fonction de transition Δ est définie par

$$\Delta : (Q^{(a)} \times Q^{(b)}) \times \Sigma \to (Q^{(a)} \times Q^{(b)}) : ((q, q'), \sigma) \mapsto (\delta^{(a)}(q, \sigma), \delta^{(b)}(q', \sigma)).$$

Les mots acceptés par $\mathcal{P}$ sont exactement les mots $w \in \Sigma^*$ tels que

$$\Delta((q_0^{(a)}, q_0^{(b)}), w) \in F^{(a)} \times F^{(b)},$$

$$\text{i.e., } \delta^{(a)}(q_0^{(a)}, w) \in F^{(a)} \text{ et } \delta^{(b)}(q_0^{(b)}, w) \in F^{(b)}.$$

Cela signifie que le langage accepté par $\mathcal{P}$ est l'intersection des langages acceptés par $\mathcal{A}$ et $\mathcal{B}$.

Proposition 4.3.5. *Soit $S = (L, \Sigma, <)$ un système de numération abstrait. Si $X \subseteq \mathbb{N}$ est un ensemble ultimement périodique de période p_X tel que $\mathrm{rep}_S(X)$ est accepté par un AFD à d états, alors la prépériode a_X de X est bornée par une constante C qui peut être effectivement calculée et qui dépend seulement de d et de p_X.*

Démonstration. Soient $\mathcal{A} = (Q, q_0, \Sigma, \delta, F)$ un AFD à d états acceptant $\mathrm{rep}_S(X)$ et $\mathcal{M} = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ l'automate minimal de L. Pour tout $q \in Q_L$, $\mathbf{u}_i(q)$ est le nombre de mots de longueur i acceptés à partir de $q \in Q_L$. Pour tout $q \in Q_L$, la suite $\mathbf{u}(q) = (\mathbf{u}_i(q))_{i\geq 0}$ satisfait une relation de récurrence linéaire. Donc, les suites $(\mathbf{u}_i(q) \bmod p_X)_{i\geq 0}$, avec $q \in Q_L$, sont ultimement périodiques. Notons, comme dans le chapitre 2, $\iota_{\mathbf{u}(q)}(p_X)$ (resp. $\pi_{\mathbf{u}(q)}(p_X)$) la prépériode (resp. la période) de $(\mathbf{u}_i(q) \bmod p_X)_{i\geq 0}$. Posons

$$I(p_X) := \max_{q\in Q_L} \iota_{\mathbf{u}(q)}(p_X) \text{ et } P(p_X) := \mathrm{ppcm}_{q\in Q_L} \pi_{\mathbf{u}(q)}(p_X).$$

Nous voulons montrer que

$$|\mathrm{rep}_S(a_X - 1)| \leq d\#Q_L + I(p_X).$$

Si nous avons $|\mathrm{rep}_S(a_X - 1)| \leq d\#Q_L$, alors l'inégalité précédente est forcément vérifier. Donc, à partir de maintenant, nous supposons $|\mathrm{rep}_S(a_X - 1)| > d\#Q_L$. Appliquons le lemme de la pompe à l'automate produit $\mathcal{A} \times \mathcal{M}_L$. Il existe des mots $x, y, z \in \Sigma^*$ avec $y \neq \varepsilon$ tels que nous avons

$$\begin{gathered}
\mathrm{rep}_S(a_X - 1) = xyz \\
|xy| \leq d\#Q_L; \\
\delta(q_0, x) = \delta(q_0, xy); \\
\delta_L(q_{0,L}, x) = \delta_L(q_{0,L}, xy); \\
\forall n \in \mathbb{N}, xy^n z \in \mathrm{rep}_S(X) \Leftrightarrow xyz \in \mathrm{rep}_S(X). \qquad (4.3)
\end{gathered}$$

Par conséquent, il suffit de montrer que $|z| \leq I(p_X)$. Procédons par contradiction et supposons que $|z| > I(p_X)$.

Prouvons

$$\mathrm{val}_S(xy^{p_X P(p_X)} yz) \equiv \mathrm{val}_S(xyz) \pmod{p_X}. \qquad (4.4)$$

Supposons $x = x_1 \cdots x_r$, $y = y_1 \cdots y_s$ et $z = z_1 \cdots z_t$. Pour tous les entiers $n \geq 1$, nous obtenons, en utilisant le lemme 4.2.3 pour $w = xy^n z$,

$|w| = r + ns + t$ et

$$\begin{aligned}\mathrm{val}_S(xy^n z) = \sum_{q\in Q_L} \Bigg(& \sum_{i=1}^{r} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \\ & + \sum_{i=r+1}^{r+s} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) + \cdots + \sum_{i=r+(n-1)s+1}^{r+ns} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \\ & \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \Bigg)\end{aligned}$$

où la première (resp. deuxième, troisième) ligne correspond à la contribution de x (resp. y, z) comme expliqué ci-dessous. Par définition (4.2) des coefficients $\beta_{q,i}(w)$, nous savons que

$$\beta_{q,1}(w) = \#\{\sigma < x_1 \mid \delta_L(q_{0,L}, \sigma) = q\} + \mathbf{1}_{q,q_{0,L}}$$

ne dépend que de x_1, mais

$$\beta_{q,2}(w) = \#\{\sigma < x_2 \mid \delta_L(q_{0,L}, x_1\sigma) = q\} + \mathbf{1}_{q,q_{0,L}}$$

dépend seulement de x_2 et $\delta_L(q_{0,L}, x_1)$. En continuant de cette façon, nous voyons que $\beta_{q,r}(w)$ dépend seulement de x_r et $\delta_L(q_{0,L}, x_1 \cdots x_{r-1})$ et que pour $1 \leq j \leq s$, $\beta_{q,r+j}(w)$ dépend de y_j et $\delta_L(q_{0,L}, xy_1 \cdots x_{j-1})$. Considérons maintenant $\beta_{q,r+s+1}(w)$. Il dépend de y_1 et de

$$\delta_L(q_{0,L}, xy_1 \cdots y_s) = \delta_L(q_{0,L}, xy) = \delta_L(q_{0,L}, x).$$

Donc, $\beta_{q,r+s+1}(w) = \beta_{q,r+1}(w)$. De la même manière, nous avons pour tout $j \in \{1, \ldots, s\}$, $\beta_{q,r+s+j}(w) = \beta_{q,r+j}(w)$. Cet argument est valable pour chaque copie de y dans w. D'où nous avons

$$\beta_{q,r+\ell s+j}(w) = \beta_{q,r+j}(w) \qquad \forall q \in Q_L, \ell \in \{0, \ldots, n-1\}, j \in \{1, \ldots, s\}.$$

Par conséquent, le développement précédent devient

$$\begin{aligned}\mathrm{val}_S(xy^n z) = \sum_{q\in Q_L} \Bigg(& \sum_{i=1}^{r} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \\ & + \sum_{i=r+1}^{r+s} \beta_{q,i}(w) \sum_{\ell=0}^{n-1} \mathbf{u}_{|w|-i-\ell s}(q) \\ & \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \Bigg).\end{aligned}$$

Choisissons maintenant $n = p_X P(p_X) + 1$. D'où, nous avons $|w| = r + p_X P(p_x)s + s + t$. Pour $q \in Q_L$ et $i \in \{r+1, \ldots, r+s\}$, nous avons

$$\sum_{\ell=0}^{n-1} \mathbf{u}_{|w|-i-\ell s}(q) = \mathbf{u}_{|w|-i} + \sum_{\ell=1}^{p_X P(p_X)} \mathbf{u}_{|w|-i-\ell s}(q)$$

et le second terme est congruent à 0 modulo p_X par les définitions de $P(p_X)$ et de $I(p_X)$ puisque nous considérons $|z| = t > I(p_X)$. Il en découle

$$\begin{aligned} \mathrm{val}_S(xy^n z) \equiv \sum_{q \in Q_L} \Bigg(& \sum_{i=1}^{r} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \\ & + \sum_{i=r+1}^{r+s} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \\ & \sum_{i=r+ns+1}^{r+ns+t} \beta_{q,i}(w)\, \mathbf{u}_{|w|-i}(q) \Bigg) \pmod{p_X}. \end{aligned}$$

Par le même raisonnement, nous remarquons que pour tout $j \in \{1, \ldots, t\}$, nous avons $\beta_{q,r+ns+j}(w) = \beta_{q,r+s+j}(xyz)$. Donc, nous pouvons facilement déduire (4.4).

Utilisons la minimalité de a_X pour obtenir une contradiction. Supposons que $a_X - 1 \in X$, l'autre cas étant similaire. Alors, pour tout entier $n \geq 1$, $a_X + np_X - 1 \notin X$, sinon a_X n'est pas la prépériode minimale. Vu (4.3), $xy^{p_X P(p_X)} yz \in \mathrm{rep}_S(X)$. Mais vu (4.4), ce mot représente un nombre de la forme $a_X + np_X - 1$ avec $n \in \mathbb{N} \setminus 0$, qui ne peut appartenir à X. Il en découle une contradiction. Par conséquent, nous avons $|z| \leq I(p_X)$ et $|\mathrm{rep}_S(a_X - 1)| \leq d\#Q_L + I(p_X)$. La prépériode de X est donc bornée par une constante C telle que si $a_X > C$, alors

$$|\mathrm{rep}_S(a_X - 1)| > d\#Q_L + I(p_X).$$

De plus, cette borne peut être effectivement calculée. En effet, puisque le système S de numération abstrait, la période p_X de X et le nombre d d'états de $\mathcal{A}$ sont donnés, $I(p_X)$ et $rep_S(n)$, pour tout $n \in \mathbb{N}$ sont effectivement calculables. □

4.4 Procédure de décision

Théorème 4.4.1. *Soit $S = (L, \Sigma, <)$ un système de numération abstrait tel que pour tous les états q de l'automate minimal émondé*

$\mathcal{M}_L = (Q_L, q_{0,L}, \Sigma, \delta_L, F_L)$ *de* L*, nous avons*

$$\lim_{i\to+\infty} \mathbf{u}_i(q) = +\infty$$

et $\mathbf{u}_i(q_{0,L}) > 0$ *pour tout* $i \geq 0$*. Supposons aussi que la suite* $\mathbf{v} = (\mathbf{v}_i(q_{0,L}))_{i\geq 0}$ *est telle que* $\lim_{m\to+\infty} N_{\mathbf{v}}(m) = +\infty$*. Alors, nous pouvons décider si un ensemble* S*-reconnaissable est ou non ultimement périodique.*

Démonstration. Cette preuve est analogue à celle du théorème 2.4.2. La suite $\mathbf{v}$ satisfait une relation de récurrence linéaire de forme (1.1) avec $a_k \neq 0$ comme dernier coefficient. De plus, $\mathbf{v}$ est une suite strictement croissante car, par hypothèse, $\mathbf{u}_i(q_{0,L}) > 0$ pour tout $i \in \mathbb{N}$.

Soit $|a_k| = p_1^{u_1} \cdots p_r^{u_r}$, avec $u_1, \ldots, u_r > 0$, la décomposition en facteurs premiers de $|a_k|$. Considérons un AFD $\mathcal{A}$ avec d états acceptant un ensemble S-reconnaissable $X \subseteq \mathbb{N}$. Supposons que X est ultimement périodique de période $p_X = p_1^{v_1} \cdots p_r^{v_r} c$ avec $v_1, \ldots, v_r \geq 0$ et $\mathrm{pgcd}(a_k, c) = 1$.

Par la proposition 4.3.2, nous avons $N_{\mathbf{v}}(p_X) \leq d\#Q_L$. Vu la remarque 2.2.7, nous obtenons

$$N_{\mathbf{v}}(c) \leq \pi_{\mathbf{u}}(c) \leq \pi_{\mathbf{v}}(p_X) \leq (N_{\mathbf{v}}(p_X))^k \leq (d\#Q_L)^k.$$

Soit $\alpha(m)$ le plus grand indice i tel que $\mathbf{v}_i(q_{0,L}) < m$. Puisque la suite $\mathbf{v}$ est strictement croissante, l'application $m \mapsto \alpha(m)$ est croissante et

$$\lim_{m\to+\infty} \alpha(m) = +\infty.$$

Puisque $\mathrm{pgcd}(a_k, c) = 1$, nous voyons comme précédemment que la suite $\mathbf{v} = (\mathbf{v}_i(q_{0,L}) \bmod c)_{i\geq 0}$ est purement périodique. Alors, nous avons $N_{\mathbf{v}} \geq \alpha(c)$ puisque $\mathbf{v}_0(q_{0,L}) < \cdots < \mathbf{v}_{\alpha(c)-1}(q_{0,L}) < c$. Donc, nous obtenons $\alpha(c) \leq (d\#Q_L)^k$, ce qui nous permet de donner une borne supérieure sur c.

Donnons maintenant une borne supérieure sur les v_j, $j \in \{1, \ldots, r\}$. Vu la remarque 2.2.7, nous avons pour tout $j \in \{1, \ldots, r\}$, $N_{\mathbf{v}}(p_j^{s_j}) \leq (d\#Q_L)^k$. De plus, l'hypothèse $\lim_{m\to+\infty} N_{\mathbf{v}}(m) = +\infty$ implique $\lim_{v\to+\infty} N_{\mathbf{v}}(p_j^v) = +\infty$ par le lemme 2.4.1. Nous faisons alors exactement le même raisonnement que dans la preuve du théorème 2.4.2. Puisque $N_U(p_j^v) \leq N_U(p_j^w)$ lorsque $v \leq w$, l'exposant v_j apparaissant dans la décomposition de p_X est borné par s_j où s_j est le plus petit entier tel que pour tout $v \geq s_j$, $N_U(p_j^v) > (d\#Q_L)^k$. Cette borne s_j peut être effectivement calculée.

Donc, si X est ultimement périodique, alors sa période p_X est bornée par une constante qui peut être effectivement calculée. Vu la proposition

4.3.5, la prépériode de X est aussi bornée par une constante qui peut être effectivement calculée.

Par conséquent, les périodes et prépériodes admissibles que nous devons vérifier sont en nombre fini. Vu la proposition 4.2.5, nous pouvons construire un automate pour chaque ensemble ultimement périodique correspondant à une paire des prépériodes et périodes admissibles. Il reste à comparer les langages acceptés par ces automates avec $\mathrm{rep}_S(X)$. Or tester si $L \setminus \mathrm{rep}_U(X) = \emptyset$ et $\mathrm{rep}_U(X) \setminus L = \emptyset$ est décidable par algorithme. Il s'en suit que nous pouvons décider si un ensemble S-reconnaissable est ou non ultimement périodique. □

Exemple 4.4.2. Considérons le système de numération abstrait donné dans l'exemple 4.1.11, $S = (L, \{a, b, c, d\}, a < b < c < d)$ avec $L = \{\varepsilon\} \cup \{a, ab\}^* \cup \{c, cd\}^*$, satisfait les hypothèses du théorème 4.4.1. En effet, nous avons pour tout $i \geq 1$, $\mathbf{u}_i(q_{0,L}) = 2F_i > 0$, $\mathbf{u}_0(q_{0,L}) = 1 > 0$ et $\lim_{i\to+\infty} \mathbf{u}_i(q_{0,L}) = +\infty$. De plus, nous avons montré, dans l'exercice 4.1.11 que la suite $\mathbf{v}_i(q_{0,L})$ satisfait $\mathbf{v}_i(q_{0,L}) = 2\,\mathbf{v}_{i-1}(q_{0,L}) - \mathbf{v}_{i-3}(q_{0,L})$ pour tout $i \geq 3$. Vu le lemme 2.4.1, nous avons, $\lim_{m\to+\infty} N_{\mathbf{v}}(m) = +\infty$.

4.5 Lien avec les problèmes de périodicité HD0L

Dans cette dernière section, nous montrons comment le théorème 4.4.1 peut être utilisé pour certains problèmes de décisions concernant les systèmes HD0L. Tout d'abord, donnons quelques définitions.

Définition 4.5.1. Un *système D0L* est un triplet $G = (\Delta, f, w)$ où Δ est un alphabet fini, $f : \Delta \to \Delta^*$ est un morphisme et w est un mot sur Δ.

Définition 4.5.2. Un *système HD0L* est un quintuple $G = (\Delta, \Gamma, f, g, w)$ où (Δ, f, w) est un système D0L, Γ est un alphabet fini et $g : \Delta^* \to \Gamma^*$ est un morphisme.

Soit $G = (\Delta, \Gamma, f, g, w)$ un système HD0L. Si le mot w est préfixe de $f(w)$ et l'ensemble $\{f^n(w) \mid n \geq 0\}$ est infini, nous notons

$$f^\omega(w) = \lim_{n\to+\infty} f^n(w).$$

De même, si w est un préfixe de $f(w)$ et si $g(f^\omega(w))$ est un mot infini sur Γ, alors nous définissons le *mot infini généré par* G comme suit

$$\omega(G) := g(f^\omega(w)) = w_0 w_1 w_2 w_3 \cdots$$

Deux systèmes HD0L G_1 et G_2, tels que $\omega(G_1)$ et $\omega(G_2)$ existent, sont ω-équivalent si $\omega(G_1) = \omega(G_2)$.

Nous nous intéressons à la question suivante.

Problème 3 (HD0L periodicity problem)**.** Etant donné un système HD0L G tel que $\omega(G)$ existe, pouvons-nous décider si le mot infini $\omega(G)$ est ultimement périodique ou non ?

Vu le lemme suivant, nous pouvons supposer que w est une lettre a.

Lemme 4.5.3. *[15] Soit $G = (\Delta, \Gamma, f, g, w)$ un système HD0L tel que $w(G)$ existe et $|w| > 1$. Alors, il existe un système HD0L ω-équivalent $G' = (\Delta', \Gamma', f'g', a)$, avec $a \in \Delta'$ tel que a est un préfixe de $f(a)$, i.e., f est prolongeable sur a.*

De plus, il est bien connu que nous pouvons supposer que f est un morphisme non-effaçant et g est un codage, i.e., $f(\sigma) \neq \varepsilon$ pour tout $\sigma \in \Delta$ et $g(\Delta) \subseteq \Gamma$. Dans [15], J. Honkala et M. Rigo ont montré que ce problème 3 de périodicité est équivalent au problème 2. Grâce à [21], étant donné un système HD0L G, nous pouvons construire canoniquement un système de numération abstrait $S = (L, \Sigma, <)$ et un automate fini déterministe avec sortie $\mathcal{M} = (Q, q_0, \Sigma, \delta, \Gamma, \tau)$ où τ est la fonction de sortie telle que

$$\forall n \geq 0, w_n = \tau(\delta(q_0, \mathrm{rep}_S(n))).$$

Une telle suite $(w_n)_{n\geq 0}$ est appelée une suite *S-automatique*. De [25], nous savons que les ensembles

$$X_b = \{n \in \mathbb{N} \mid w_n = b\} \text{ pour } b \in \Gamma$$

sont S-reconnaissables. Par conséquent, si S satisfait les hypothèses du théorème 4.4.1, nous pouvons décider si les ensembles X_b sont ultimement périodiques ou non. Enfin, remarquons que le mot infini $\omega(G) = g(f^\omega(a))$ est ultimement périodique si et seulement si, pour chaque $b \in \Gamma$, les ensembles X_b sont ultimement périodiques.

Donc, si G est tel que le système S de numération abstrait associé satisfait les hypothèse du théorème 4.4.1, alors nous pouvons décider si $\omega(G)$ est ultimement périodique ou non.

Bibliographie

[1] J.-P. Allouche et J. Shallit, *Automatic sequences. theory, applications, generalizations*, Cambridge University Press, Cambridge, 2003.

[2] A. J. Baker, *An Introduction to p-adic Numbers and p-adic Analysis*, 2007, `http://www.maths.gla.uk/~ajb`.

[3] J. Bell, E. Charlier, A. S. Fraenkel, et M. Rigo, *A Decision Problem for Ultimately Periodic Sets in Non-standard Numeration Systems*, Internat. J. Algebra and Computation **19** (2009), 809–839.

[4] J. Berstel, *Noncommutative rational series with applications*, Springer-Verlag, 2010.

[5] P. B. Bhattacharya, S. K. Jain, et S. R. Nagpaul, *Basic Abstract Algebra*, 2e éd., Cambridge University Press, Cambridge, 1994.

[6] V. Bruyère, G. Hansel, C. Michaux, et R. Villemaire, *Logic and p-recongnizable sets of integers*, Bull. Belg. Math. Soc. **1** (1994), 191–238.

[7] E. Charlier, *Abstract numeration systems : Recognizability, decidability, multidimensional s-automatic words, and real numbers*, Thèse de Doctorat, Université de Liège, 2009.

[8] A. Cobham, *On the base-dependence of sets of numbers recognizable by finite automata*, Math. Systems Theory **3** (1969), 186–192.

[9] E. Duchêne et M. Rigo, *Cubic pisot unit combinatorial games*, Monat. fur Math. **155** (2008), 217–249.

[10] H. T. Engstrom, *On Sequences Defined by Linear Recurrence Relations*, Trans. of the AMS **33** (1931), 210–218.

[11] C. Frougny, *On the sequentiality of the successor function*, Inform. and Computation **139** (1997), 17–38.

[12] F. Q. Gouvêa, *p-adic Numbers : An Introduction*, 2e éd., Springer-Verlag, Berlin, 1997.

[13] M. Hollander, *Greedy numeration systems and regularity*, Theory Comput. Syst. **31** (1998), 111–133.

[14] J. Honkala, *A decision method for the recognizability of sets defined by number systems*, Theoret. Inform. Appl. **20** (1986), 395–403.

[15] J. Honkala et M. Rigo, *Decidability questions related to abstract numeration systems*, Discrete Mathematics **285** (2004), 329–333.

[16] N. Koblitz, *p-adic numbers, p-adic analysis, and zeta-functions*, Springer-Verlag, New York, 1977.

[17] P. Lecomte et M. Rigo, *Real numbers having ultimately periodic representations in abstract numeration*, Inform. and Comput. **192** (2004), 57–83.

[18] P. Lecomte et M Rigo, *Abstract numeration systems*, Combinatorics, Automata and Number Theory (V. Berthé et M. Rigo, éds), Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, 2010.

[19] P. B. A. Lecomte et M. Rigo, *Numeration systems on a regular language*, Theory Comput. Syst. **34** (2001), 27–44.

[20] M. Lothaire, *Algebraic combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002.

[21] A. Maes et M. Rigo, *More on generalized automatic sequences*, J. Autom. Lang. Comb. **7** (2002), 351–376.

[22] A. Muchnik, *The definable criterion for definability in presburger arithmetic and its applications*, Theoret. Comput. Sci. **290** (2003), 1433–1444.

[23] M. Rigo, *Mathématiques discrètes : Notes de cours 2009–2010*, Ulg, Faculté des sciences.

[24] M Rigo, *Théorie des automates et langages formels : Notes de cours 2007–2008*, Ulg, Faculté des sciences.

[25] M. Rigo, *Generalization of automatic sequences for numeration systems on a regular language*, Theoret. Comput. Sci. **244** (2000), 271–281.

[26] M Rigo, *Abstract numeration systems on a regular languages and recognizability*, Thèse de Doctorat, Université de Liège, 2001.

[27] A. M. Robert, *A Course in p-adic Analysis*, Graduate Texts in Mathematics, vol. 198, Springer-Verlag, New York, 2000.

[28] M.-P. Schutzenberger, *On a theorem of R. Jungen*, Proc. Amer. Math. Soc. **13** (1962), 885–890.

[29] J. Shallit, *Numeration systems, linear recurrences, and regular sets*, Inform. and Comput. **113** (1994), 331–347.