UNIVERSITE DE LIEGE
Faculté des Sciences
Institut de Mathématique

Les systèmes de numération en base rationnelle

Mémoire présenté par
Julien LEROY
dans le cadre du Master
à finalité approfondie en
sciences mathématiques

Année académique 2007–2008

Je souhaite tout d'abord remercier M. Rigo
pour sa disponibilité, son aide et ses conseils
tout au long de ce travail, ainsi que pour
la confiance qu'il me témoigne en appuyant
mes projets de doctorat.

Merci également à Marion et à Rémi pour leur aide.

Enfin, je tiens à remercier Anne
pour cette année passée ensemble.

Introduction

Les systèmes de numération sont un outil classique permettant de représenter les nombres entiers ou réels par des mots sur un alphabet fini de chiffres. Un tel sytème peut être vu comme un couple composé d'une suite strictement croissante de nombres (comme par exemple la suite $(k^i)_{i\in\mathbb{Z}}$ pour la numération usuelle en base entière $k \geq 2$) et d'un alphabet fini de chiffres. Ainsi, tout nombre positif x est décomposé (le plus souvent algorithmiquement) en une combinaison linéaire (voire une série) d'éléments de la suite dont les coefficients appartiennent à l'alphabet considéré. La concaténation des chiffres apparaissant dans cette décomposition est alors une représentation de x. L'étude des propriétés de ces représentations en tant que mots sur un alphabet conduit alors à des résultats arithmétiques sur les nombres. L'utilisation de certains systèmes permet également de réaliser algorithmiquement des opérations sur les nombres de manière efficace. Par exemple, dans le système binaire (il s'agit d'écrire les nombres en base 2), l'addition de deux entiers peut se faire en un temps proportionnel à la taille des données, mais si nous considérons l'alphabet plus large $\{-1, 0, 1\}$, alors l'addition se fait en un temps indépendant de cette taille.

Dans les systèmes de β-numération, où β est un réel strictement supérieur à 1 (c'est par exemple le cas du système binaire avec $\beta = 2$), la suite utilisée est la suite $(\beta^i)_{i\in\mathbb{Z}}$ des puissance de β et l'alphabet est l'ensemble des entiers positifs strictement inférieurs à β ; nous disons que β est la *base* du système. Tout nombre réel positif admet alors une représentation, appelée le *β-développement*, calculée par un algorithme glouton de gauche à droite, i.e. le chiffre le plus significatif en premier. Lorsque β est un entier, les représentations sont dites *standards* et le langage des β-développements des entiers jouit de belles propriétés : il est accepté par un automate fini, la conversion de mots écrits sur un alphabet plus large dans l'alphabet canonique est réalisée par un transducteur fini,... Dans le cas où β est un nombre de Pisot [1], le langage des β-développements des entiers positifs jouit également de la plupart de ces propriétés.

Ce travail est principalement basé sur l'article de S. Akiyama, C. Frougny et J. Sakarovitch, [9], intitulé *Powers of rationals modulo* 1 *and rational base number systems* et s'intéresse aux systèmes de numération en base $\beta = \frac{p}{q}$, où p et q sont des entiers permiers entre eux, $p > q \geq 1$. Etrangement le cas des bases rationnelles a

1. i.e. un entier algébrique strictement supérieur à 1 dont tous les conjugués sont en module strictement inférieur à 1.

été très peu étudié jusqu'ici et, comme nous allons le voir, il s'avère que certaines des propriétés des bases entières ou de Pisot ne sont plus valables.

Le premier chapitre rappelle les notions de théorie des automates et langages formels et de théorie des nombres utilisées tout au long de ce travail. En particulier, nous rappelons comment munir l'ensemble $A^{\mathbb{N}}$ des mots infinis (sur un alphabet fini A) d'une distance afin d'en faire un espace métrique complet compact (une annexe a été rajoutée au travail afin de développer certaines notions de topologie sur l'espace $A^{\mathbb{N}}$). Nous rappelons également les notions d'automate, de transducteur et de grammaire algébrique. Nous introduisons enfin les systèmes de numération et présentons l'exemple de la β-numération.

Le deuxième chapitre concerne plus spécifiquement les représentations des entiers. L'algorithme que nous utilisons pour calculer le $\frac{p}{q}$-développement d'un entier n'est pas celui utilisé classiquemement dans la β-numération et calcule ces développements de droite à gauche, i.e. le chiffre le moins significatif en premier ; nous l'appelons l'*algorithme MD* ou *algorithme de division modifié*. L'alphabet A du système est également différent de celui utilisé usuellement et consiste en l'ensemble des entiers positifs strictement inférieurs à p. Un premier résultat important est que pour tout entier positif N, le $\frac{p}{q}$-développement calculé par l'algorithme MD est l'unique représentation finie dans ce système de numération. Comme annoncé, certaines propriétés des bases entières ou de Pisot ne sont plus valables ici. Par exemple, le langage $L_{\frac{p}{q}}$ des $\frac{p}{q}$-développements des entiers positifs n'est pas régulier, ni même algébrique. Cependant, certaines propriétés restent vraies, notamment la conversion entre alphabets qui est réalisée par un transducteur à droite fini ou, la fonction successeur qui est également calculée par un transducteur à droite fini.

Le troisième chapitre fournit les prémisses de la représentation des réels que nous étudions au quatrième chapitre. Par le biais d'une famille de fonctions partielles τ_a, nous y construisons un arbre $T_{\frac{p}{q}}$, sous-arbre du graphe de Cayley du monoïde A^*, représentant le langage $L_{\frac{p}{q}}$. Nous considérons alors l'ensemble $W_{\frac{p}{q}}$ des mots infinis sur A qui étiquetent les chemins infinis dans $T_{\frac{p}{q}}$ et nous définissons des mots particuliers : les mots *ultimement minimaux* (resp. *ultimement maximaux*) qui sont des mots de $W_{\frac{p}{q}}$ qui, à partir d'un noeud de l'arbre, empruntent toujours la branche d'étiquette minimum (resp. maximum). Nous présentons alors quelques résultats sur l'évaluation de ces mots infinis. Nous terminons le chapitre par une application des systèmes de numération en base $\frac{p}{q}$: l'étude du *problème de Josephus*.

Le quatrième chapitre s'intéresse aux $\frac{p}{q}$-développements des réels non-entiers. Par opposition à ce qui se fait dans les systèmes de β-numération classiques, nous définissons *a priori*, et non algorithmiquement, l'ensemble des $\frac{p}{q}$-développements des réels et nous montrons ensuite que ce choix est cohérent, i.e. nous montrons que tout réel admet au moins un $\frac{p}{q}$-développement. Le $\frac{p}{q}$-développement d'un réel n'est pas nécessairement unique, mais nous montrons que l'ensemble des réels admettant plus d'un développement est infini dénombrable et que ces réels n'admettent qu'un nombre fini de développements. Enfin, nous définissons une autre représentation des réels, la *représentation compagnon*, qui est calculée algorithmiquement sur un alphabet C plus large que l'alphabet canonique et comportant des chiffres négatifs. Nous présentons alors une méthode permettant de calculer une approximation d'un $\frac{p}{q}$-développement à partir de la représentation compagnon.

Le cinquième et dernier chapitre est une application des systèmes de numération en base rationnelle. Il concerne un problème bien connu de K. Mahler : la démonstration de l'existence, ou de l'inexistence, de *Z-nombres*. Il s'agit de nombres réels z non nuls tels que la partie fractionnaire de $z\left(\frac{3}{2}\right)^n$ appartient au semi-intervalle $[0, \frac{1}{2}[$ pour tout entier positif n. Nous démontrons ici que l'étude de ces nombres peut en fait être transposée à l'étude des chemins ultimement minimaux dans l'arbre $T_{\frac{p}{q}}$.

Chapitre 1

Premières notions

Nous rappelons dans ce chapitre les définitions et résultats de base dont nous avons besoin dans ce travail. Nous présentons par exemple les premières définitions de la théorie des automates et langages formels comme les mots finis ou infinis, les transducteurs, les automates finis et les langages régulier ainsi que les grammaire algébrique et les langages algébriques. Nous rappelons également comment munir l'ensemble $A^{\mathbb{N}}$ des mots infinis d'une distance afin d'en faire un espace métrique complet compact. Enfin, nous introduisons en quelques mots les systèmes de numération, notament les systèmes de β-numération, β étant un réel strictement supérieur à 1.

1.1 Mots finis et infinis

Un *alphabet* est un ensemble fini A dont les éléments sont appelés *lettres* ou *symboles*. Nous ne considérons dans ce travail que des alphabets composés de *chiffres*, i.e. nous ne considérons que des alphabets qui sont des sous-ensembles de $\mathbb{Z}$. Une suite finie d'éléments de A est un *mot* et l'ensemble des mots, muni de la concaténation, est un monoïde libre noté A^*. Le *mot vide* est le neutre de A^* et est noté ε.

La *longueur* d'un mot w sur A, notée $|w|$ est le nombre de chiffres qui composent ce mot. L'ensemble des mots sur A de longueur n (resp. de longueur inférieure à n) est noté A^n (resp. $A^{\leq n}$). La concaténation de n copies d'un mot w est notée w^n. Un mot u est un *facteur* d'un mot w s'il existe des mots x et y sur A vérifiant $w = xuy$. Si x (resp. y) est le mot vide, alors le mot u est un *préfixe* (resp. un *suffixe*) du mot w.

Un *langage* sur A est une partie L de A^*. Un langage est *préfixiel* (resp. *suffixiel*) s'il contient tous les préfixes (resp. tous les suffixes) de chacun de ses éléments. Si L est un langage sur A, alors $\mathrm{Pref}(L)$ (resp. $\mathrm{Suff}(L)$) est un langage sur A composé de tous les préfixes (resp. suffixes) des mots de L. Le langage $\mathrm{Pref}(L)$ (resp. $\mathrm{Suff}(L)$) est donc un langage préfixiel (resp. suffixiel) et par conséquent, un langage L est préfixiel (resp. suffixiel) si et seulement s'il est égal au langage $\mathrm{Pref}(L)$ (resp. $\mathrm{Suff}(L)$).

Supposons l'alphabet A ordonné par l'ordre total[1] $\leq$. L'ensemble A^* est alors totalement ordonné par l'*ordre généalogique* $\preceq$ défini comme suit : $v \prec w$ si $|v| < |w|$ ou si $|v| = |w|$ et s'il existe des chiffres $a < b$ telles que $v = uav'$ et $w = ubw'$, où u, v' et w' sont des mots sur A. L'ensemble A^* est également totalement ordonné par l'*ordre lexicographique* $\sqsubseteq$ défini comme suit : $v \sqsubset w$ si v est un préfixe de w, ou s'il existe des chiffres $a < b$ tels que $v = uav'$ et $w = ubw'$, où u, v' et w' sont des mots sur A. L'ordre lexicographique est donc l'orde du dictionnaire[2], alors que l'ordre généalogique classe d'abord les mots par longueur et utilise l'ordre du dictionnaire à l'intérieur d'une classe de mots de même longueur. Remarquons que ces deux ordres coïncident pour les couples de mots de même longueur. Remarquons également que l'ordre généalogique est un bon ordre[3], ce qui se montre sans peine, tandis que si l'alphabet A contient plus d'un élément, l'ordre lexicographique n'en est pas un. En effet, considérons le langage $L = \{a^n b \mid n \geq 0\}$ sur l'alphabet $\{a, b\}$ totalement ordonné par $a < b$. Pour tout entier positif i, le mot $a^i b$ est strictement inférieur au mot $a^{i+1} b$ et le langage L n'admet donc pas de minimum.

Un *mot infini* sur A est une suite infinie d'éléments de A. Dans ce travail, nous ne travaillons qu'avec des mots infinis à droite dont les indices sont des entiers positifs ou négatifs suivant le contexte. Cependant, dans les deux cas, nous notons $A^{\mathbb{N}}$ l'ensemble des mots infinis sur A et pour autant que les notations le permettent, nous notons les mots infinis par des caractères gras. Nous notons également par $\boldsymbol{a}_{[n]}$ le préfixe de longueur n du mot infini $\boldsymbol{a}$ et le n-ième chiffre se note $\boldsymbol{a}_n$.

Un mot infini $\boldsymbol{a}$ est *ultimement périodique* s'il peut s'écrire sous la forme

$$\boldsymbol{a} = uw^{\omega} = uwww\cdots ,$$

où u et w sont des mots de A^*, $w \neq \varepsilon$.

Il est évident que l'ordre généalogique n'est pas défini sur $A^{\mathbb{N}}$. L'ordre lexicographique par contre se généralise de la manière suivante : $\boldsymbol{a} \sqsubset \boldsymbol{b}$ s'il existe un mot u dans A^*, deux mots infinis $\boldsymbol{x}$ et $\boldsymbol{y}$ dans $A^{\mathbb{N}}$ et deux chiffres a et b dans A, $a < b$, tels que $\boldsymbol{a} = ua\boldsymbol{x}$ et $\boldsymbol{b} = ub\boldsymbol{y}$.

L'ensemble $A^{\mathbb{N}}$ est muni d'une distance δ définie comme suit : si $\boldsymbol{a} = (a_n)_{n\in\mathbb{N}}$ et $\boldsymbol{b} = (b_n)_{n\in\mathbb{N}}$, alors

$$\delta(\boldsymbol{a}, \boldsymbol{b}) = \begin{cases} 0 & \text{si } \boldsymbol{a} = \boldsymbol{b} \\ 2^{-r} & \text{si } \boldsymbol{a} \neq \boldsymbol{b}, \end{cases}$$

où $r = \min\{n \mid a_n \neq b_n\}$. En d'autres termes, plus deux mots ont un long préfixe en commun, plus la distance entre ces deux mots est petite.

1. Cette supposition est naturelle puisque nous travaillons avec des alphabets qui sont des sous-ensembles de $\mathbb{Z}$; l'ordre $\leq$ est donc l'ordre habituel sur les entiers.

2. Il s'agirait exactement de l'ordre du dictionnaire si nous travaillions avec des lettres, ce qui n'est pas le cas ici.

3. Un *bon ordre* sur A^* est un ordre total tel que tout sous-ensemble non-vide de A^* admet un élément minimum.

Il est facile de voir que δ est bien une distance et la topologie induite par celle-ci est la topologie produit de la topologie discrète sur A, ce qui fait de $A^{\mathbb{N}}$ un espace métrique complet compact. Nous étudions plus en détails la topologie de $A^{\mathbb{N}}$ dans l'annexe ; nous y caractérisons entre autres les ensembles fermés de $A^{\mathbb{N}}$, ainsi que les ensembles relativement compacts.

1.2 Automates et transducteurs

Un automate sur un alphabet A est la donnée d'un quintuple $\mathcal{A} = (Q, A, E, I, T)$, où

i) Q est un ensemble fini dont les éléments sont les états de l'automate ;

ii) $I \subseteq Q$ est l'ensemble des états initiaux ;

iii) $F \subseteq Q$ désigne l'ensemble des états finals ;

iv) E est l'ensemble des transitions de l'automate.

Les transitions de l'automate sont étiquetées par des éléments de A et l'ensemble E peut donc être vu comme un sous-ensemble de $Q \times A \times Q$.

Un automate $\mathcal{A}$ est *déterministe* s'il ne possède qu'un état initial et si pour tout couple $(q, a) \in Q \times A$, il existe au plus un état p dans Q tel que le triplet (q, a, p) appartient à E et dans ce cas, l'ensemble E des transition définit une fonction partielle appelée *fonction de transition* :

$$\delta : Q \times A \mapsto Q$$

et nous notons $\mathcal{A} = (Q, A, \delta, \{q_0\}, T)$.

Soit $\mathcal{A}$ un automate sur A. L'*automate transposé* de $\mathcal{A}$, noté $\mathcal{A}^t$, est défini par $\mathcal{A}^t = (Q, A, E^t, T, I)$, où le triplet (p, a, q) est une transition dans l'automate $\mathcal{A}^t$ si et seulement si le triplet (q, a, p) en est une dans l'automate $\mathcal{A}$. Un automate est *co-déterministe* si son automate transposé est déterministe.

Un état q de $\mathcal{A}$ est *accessible* s'il existe un chemin dans $\mathcal{A}$ partant d'un état initial et aboutissant dans l'état q. La *partie accessible* de $\mathcal{A}$ est le sous-automate induit par l'ensemble des états acessibles. Un automate est alors *accessible* si tous ses états le sont, i.e. s'il est égal à sa partie accessible. L'étiquette d'un chemin est la concaténation (de gauche à droite) des étiquettes des différentes transitions qui composent le chemin. Un *chemin accepteur* dans $\mathcal{A}$ est un chemin partant d'un état initial et aboutissant dans un état final. Un mot est alors *accepté* par $\mathcal{A}$ s'il est l'étiquette d'un chemin accepteur dans $\mathcal{A}$. Enfin, le *langage accepté par l'automate* $\mathcal{A}$ est l'ensemble des mots acceptés par $\mathcal{A}$.

La figure suivante représente un automate sur l'alphabet $\{0, 1\}$ qui accepte la représentation binaire des multiples de 3. Les états initiaux sont indiqués par une flèche entrante et les états finals par une flèche sortante.

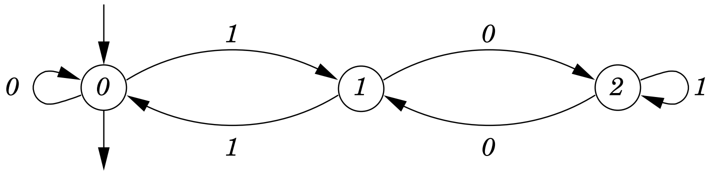

FIGURE 1.1 – Automate acceptant la représentation binaire des entiers multiples de 3.

Un automate $\mathcal{A}$ sur A est *fini* s'il n'a qu'un nombre fini d'états[4]. Un langage L sur A est *régulier* si et seulement s'il existe un automate fini sur A dont le langage accepté est exactement L. Cependant, il peut exister plussieurs automates finis qui acceptent un langage donné. Par exmple, le langage sur l'alphabet $\{a, b\}$ formé des mots ne contenant pas deux a consécutifs est accepté par les automates représentés aux figures 1.2 et 1.3. Nous définissons alors l'*automate minimal* $\mathcal{A}_L$ d'un langage L qui est un automate déterministe et accessible qui accepte le langage L. Cet automate est fini si et seulement si le langage est régulier et dans ce cas, tout automate fini qui accepte L possède au moins autant d'états que $\mathcal{A}_L$. Nous ne présentons ici que la construction de l'automate minimal d'un langage ; les démonstrations des propriétés enoncées ci-dessus sont disponibles dans [7].

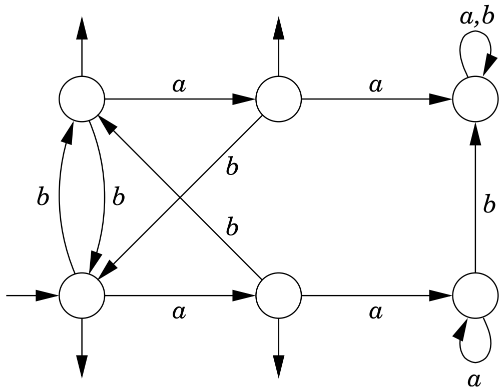

FIGURE 1.2 – Automate fini acceptant le langage $A^* \setminus (A^* aa A^*)$.

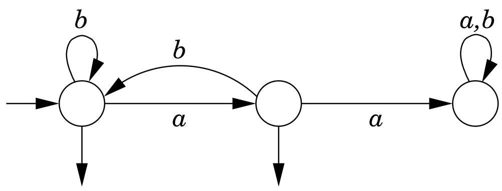

FIGURE 1.3 – Automate fini acceptant le langage $A^* \setminus (A^* aa A^*)$.

Soit L un langage sur l'alphabet A. Pour tout mot w de A^*, nous notons $w^{-1}.L$ l'ensemble des mots de A^* qui, concaténés avec w, appartiennent au langage L, i.e.

$$w^{-1}.L = \{u \in A^* \mid wu \in L\} .$$

Nous définissons alors la congruence à droite $\sim_L$ sur A^*, appelée *congruence de Nérode*, par

$$\forall u, v \in A^*, \qquad u \sim_L v \Leftrightarrow u^{-1}.L = v^{-1}.L.$$

4. Et donc un ensemble fini de transitions, car l'alphabet A est fini.

En d'autres termes, $u \sim_L v$ si et seulement si pour tout mot $w \in A^*$, $uw \in L \Leftrightarrow vw \in L$ et nous notons $[w]_L$ la classe d'équivalence pour $\sim_L$ d'un mot w de A^*.

Nous définissons maintenant l'automate minimal $\mathcal{A}_L = (Q_L, A, \delta_L, q_{0,L}, T_L)$ d'un langage L comme suit :

i) $Q_L = \{w^{-1}.L \mid w \in A^*\}$;

ii) $q_{0,L} = \varepsilon^{-1}.L$;

iii) $T_L = \{u^{-1}.L \mid u \in L\}$;

iv) $\forall w \in A^*, \forall a \in A,\ \delta_L(w^{-1}.L, a) = (wa)^{-1}.L$.

Les automates que nous venons de définir ici sont des machines qui lisent des mots sur un alphabet donné et qui partitionnent l'ensemble des mots sur A en l'ensemble des mots qui sont acceptés par l'automate et l'ensemble de ceux qui ne le sont pas. Nous définissons maintenant un autre type de machine qui ne se contente plus de lire, mais qui peut également écrire. Il s'agit des *transducteurs*.

Un *transducteur* se définit exactement comme un automate, à savoir qu'il possède un ensemble d'états dont certains sont initiaux et d'autres finals et qu'il existe des transitions entre ces états. La différence avec les automates est la présence d'un second alphabet B et les étiquettes des transitions : celles-ci ne sont plus des chiffres de A, mais bien des éléments de $A^* \times B^*$. Le transducteur lit le mot sur la première composante et écrit le deuxième mot. La concaténation se fait composante à composante dans $A^* \times B^*$, ce qui en fait un monoïde libre également. Dans un transducteur, l'étiquette d'un chemin est donc un couple (u, v) de mots, où u et v appartiennent repsectivement à A^* et B^*. Un transducteur définit donc une relation de A^* dans B^*.

La figure suivante représente un transducteur $\mathcal{B}$ sur $\{0,1\} \times \{0,1\}$. Celui-ci effectue la division entière par 3 des entiers écrits en base binaire : un couple (u, v) est accepté par $\mathcal{B}$ si u est la représentation binaire d'un entier n multiple de 3 et si v est la représentation binaire de $\frac{n}{3}$.

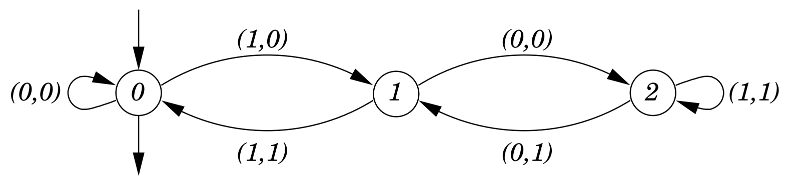

Figure 1.4 – Transducteur effectuant la division par 3.

Dans ce travail, nous ne considérons que des transducteurs *lettre-à-lettre*, i.e. des transducteurs dont les étiquettes sont des éléments de $A \times B$.

Si nous ne considérons que la première compsante d'un transducteur lettre-à-lettre $\mathcal{T}$, nous obtenons un automate sur A : l'*automate d'entrée* sous-jacent à $\mathcal{T}$. Nous disons qu'un transducteur est *séquentiel* (resp. *co-séquentiel*) si son automate d'entrée est déterministe (resp. co-déterministe).

Nous pourrions supposer que les flèches sortantes sont étiquetées par des couples de la forme[5] (ε, b), où b est une chiffre de B. Ainsi, si, dans un transducteur $\mathcal{T}$, le couple (u, v) de mots décrit un chemin débutant dans un état initial i et aboutissant dans un état final f, et si la flèche sortante partant de f est étiquetée par le couple (ε, b), alors le mot u de A^* est associé au mot vb de B^* par la relation définie par $\mathcal{T}$.

Il serait également possible de considérer des automates et des transducteurs qui lisent et écrivent de droite à gauche. Dans ce cas, nous parlons d'*automates à droite* et de *transducteurs à droite*.

La figure suivante représente un transducteur possédant ces deux dernières particularités : il lit et écrit de droite à gauche et ses flèches sortantes sont étiquetées. Celui-ci réalise la conversion entre l'alphabet $\{0, 1, 2\}$ et l'alphabet binaire $\{0, 1\}$, i.e. si w est un mot écrit sur $\{0, 1, 2\}$, alors ce transducteur fournit un mot écrit sur $\{0, 1\}$ ayant la même valeur numérique[6].

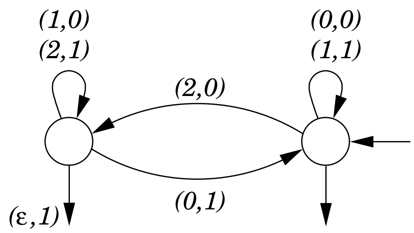

FIGURE 1.5 – Convertisseur de $\{0, 1, 2\}^*$ dans le système binaire.

5. i.e. le transducteur possède une fonction final qui, ici, écrit une lettre suplémentaitre.
6. La conversion entre alphabets est développée à la section 2.3.

1.3 Grammaires et langages algébriques

Soient A et V deux alphabets finis (supposés disjoints). Une *grammaire algébrique*, ou *grammaire hors contexte*, est la donnée d'un quatruple $G = (V, A, P, S)$, où

i) $P \subset V \times (V \cup A)^*$ est un ensemble fini appelé l'*ensemble des règles de dérivation* (ou *productions*) de G et

ii) $S \in V$ est le *symbole initial* de G.

Les éléments de l'alphabet V sont appelés *variables* (ou *symboles non terminaux*) et les éléments de l'alphabet A sont les *symboles terminaux*. Par convention, nous notons les variables par des lettres majuscules et les symboles terminaux par des minuscules.

Soient $X \in A$ une variable, $w \in (V \cup A)^*$ un mot et (X, w) une règle de dérivation de G. Nous disons que X (resp. w) est le *premier* (resp. *second*) *membre* de la production (X, w). Si $X \in V$ est une variable et si $(X, w_1), (X, w_2), \ldots, (X, w_n) \in P$ sont des productions ayant X comme premier membre et où $w_1, w_2, \ldots, w_n \in (V \cup A)^*$, alors nous notons

$$X \to w_1 \mid w_2 \mid \cdots \mid w_n.$$

Si w peut s'écrire uXv, avec $X \in V$ et $u, v \in (V \cup A)^*$, alors nous notons

$$w \Rightarrow z,$$

lorsque $z = uyv$ avec $(X, y) \in P$. Nous disons alors que z est obtenu grâce à une *dérivation de longueur* 1. En d'autres termes, nous avons remplacé dans w une occurence d'une variable X par le second membre y d'une production $X \to y$ de G ayant X comme premier membre. Nous notons $\Rightarrow^*$ la fermeture réflexive et transitive de $\Rightarrow$. Ainsi, $w \Rightarrow^* z$ si $w = z$ ou s'il existe des mots $w_1, w_2, \ldots, w_n \in (V \cup A)^*$, $n \geq 0$, tels que

$$w \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \cdots \Rightarrow w_n \Rightarrow z.$$

Dans ce cas, nous disons que z est obtenu à partir de w grâce à une *dérivation* de longueur $n + 1$.

Enfin, le *langage généré* par G est l'ensemble des mots de A^* qui s'obtiennent par dérivation à partir du symbole intial S de G, i.e.

$$L(G) = \{w \in A^* \mid S \Rightarrow^* w\}.$$

Un langage L est alors *algébrique*, ou *hors contexte*, s'il existe une grammaire algébrique G telle que $L = L(G)$.

Considérons la grammaire algébrique $G = (V, A, P, S)$, où $V = \{S, X\}$, $A = \{a, b\}$ et les productions de G sont données par

$$\begin{aligned} S &\to XX \\ X &\to XXX \mid bX \mid Xb \mid a. \end{aligned}$$

Le mot *ababaa* appartient au langage généré par G car

$$\begin{aligned}\underline{S} &\Rightarrow \underline{X}X \Rightarrow a\underline{X} \Rightarrow a\underline{X}XX \Rightarrow ab\underline{X}XX \\ &\Rightarrow aba\underline{X}X \Rightarrow abab\underline{X}X \Rightarrow ababa\underline{X} \Rightarrow ababaa.\end{aligned}$$

À chaque étape, nous avons souligné la variable substituée. Ainsi, le mot *ababaa* est obtnu à partir du symbole initial S par une dérivation de longueur 8. La suite des dérivations appliquées produisant un mot donné n'est pas forcément unique. En effet, le mot *ababaa* peut également être obtenu à partir du symbole S des manières suivantes :

$$\begin{array}{lll} S & \Rightarrow & \underline{X}X \\ & & \underline{X}XXX \\ & & a\underline{X}XX \\ & & ab\underline{X}XX \\ & & aba\underline{X}X \\ & & abab\underline{X}X \\ & & ababa\underline{X} \\ & & ababaa \end{array} \qquad \begin{array}{lll} S & \Rightarrow & X\underline{X} \\ & & \underline{X}a \\ & & XX\underline{X}a \\ & & XXb\underline{X}a \\ & & X\underline{X}baa \\ & & Xb\underline{X}baa \\ & & \underline{X}babaa \\ & & ababaa \end{array}$$

1.4 Représentation des nombres

Soit $U = (u_i)_{i\in\mathbb{Z}}$ une suite strictement croissante de nombres réels positifs telle que

$$\forall k \geq 0, \qquad \sum_{k\geq i\geq -\infty} u_i < +\infty.$$

Une *représentation* dans le système U d'un réel x positif sur un alphabet fini D de chiffres est une suite infinie $(d_i)_{-\infty\leq i\leq k}$, avec k un entier et d_i des éléments de D, telle que

$$x = \sum_{k\geq i\geq -\infty} d_i u_i.$$

Nous la notons

$$\langle x\rangle_U = d_k d_{k-1}\cdots d_0.d_{-1}d_{-2}\cdots ,$$

avec le chiffre le plus significatif en premier.

Lorsqu'une représentation se termine par une infinité de 0, elle est dite *finie* et la suite de 0 est généralement oubliée. Lorsque tous les d_i situés à droite du point (appelé *séparateur*) sont des 0, la représentation est dite *entière*.

Considérons l'exemple de la β-numération. Soit $\beta > 1$ un réel. La suite considérée est la suite des puissances de β :

$$U = \left(\beta^i\right)_{i\in\mathbb{Z}}$$

et l'alphabet D est $\{0, 1, \ldots \beta - 1\}$ si β est un entier et $\{0, 1, \ldots, \lfloor\beta\rfloor\}$ sinon, où $\lfloor\beta\rfloor$ désigne la partie entière de β.

Tout nombre réel $x \in [0, 1]$ peut être développé (pas nécessairement de manière unique) sous la forme

$$x = \sum_{i=-1}^{-\infty} x_i\beta^i,$$

avec $x_i \in D$ pour tout i. Le mot infini $x_{-1}x_{-2}x_{-3}\cdots$ est donc une représentation[7] de x dans le système U. Dans le cas de la β-numération, il est fréquent de considérer une représentation particulière : le *β-développement* dont les coéfficients peuvent être calculés comme ceci :

- $r_0 = x$
- pour tout $i \leq -1$,
  $x_i = \lfloor\beta r_{i+1}\rfloor$,
  $r_i = \{\beta r_{i+1}\}$.

7. La manière d'indicer les coefficients x_i (par des entiers négatifs décroissants) peut sembler étrange. Nous avons fait ce choix par souci de cohérence avec la définition d'une représentation dans un système U.

Nous notons $\langle x\rangle_\beta$ le β-développement de x. Nous avons souvent recourt à l'application suivante, appelée *β-transformation* :

$$T_\beta : [0,1] \to [0,1] : x \mapsto \{\beta x\}.$$

Ainsi, pour tout réel $x \in [0,1]$, nous avons $\langle x\rangle_\beta = (x_i)_{i\leq -1}$ si et seulement si $x_i = \left\lfloor \beta T_\beta^{-i-1}(x)\right\rfloor$ pour tout entier i strictement négatif.

L'algorithme utilisé pour le β-développement peut être qualifié de "glouton" : la suite $(x_i)_{i\leq -1}$ est le β-développement d'un réel $x \in [0,1[$ si et seulement si pour tout $j \leq -1$,

$$\sum_{i=j}^{-\infty} x_i\beta^i < \beta^{j+1}.$$

En effet, pour tout $j \leq -1$, nous montrons facilement

$$r_{j+1} < 1.$$

En effet, pour $j = -1$, nous avons $r_0 = x < 1$ et pour $j < -1$, nous obtenons $r_{j+1} = \{\beta r_{j+2}\} < 1$. De plus, nous avons pour tout $j \leq -1$

$$r_{j+1} = x_j\beta^{-1} + x_{j-1}\beta^{-2} + x_{j-2}\beta^{-3} + \cdots < 1$$

et en multipliant cette inégalité par β^{j+1}, nous obtenons le résultat annoncé.

Soit maintenant $x > 1$ un réel. Pour obtenir son β-développement, il suffit de le diviser par une puissance suffisante de β, notons β^k, afin d'obtenir un nouveau réel $y = \frac{x}{\beta^k} \in [0,1]$ dont nous pouvons calculer le β-développement

$$\langle y\rangle_\beta = y_{-1}y_{-2}y_{-3}\cdots .$$

Le β-développement de x se trouve alors en multipliant celui de y par β^k, ce qui a pour effet de décaler tous les coéfficient de $\langle y\rangle_\beta$ de k places vers la gauche. Ainsi, nous obtenons

$$\langle x\rangle_\beta = x_{k-1}x_{k-2}\cdots x_0.x_{-1}x_{-2}\cdots ,$$

avec pour tout entier $i \leq k-1$, $x_i = y_{i-k}$.

La valeur numérique dans le système U d'un mot sur un alphabet de chiffre D est donnée par l'application *évaluation* :

$$\pi_{U,D} : D^{\mathbb{Z}} \to \mathbb{R} : \boldsymbol{d} = (d_i)_{k\geq i\geq -\infty} \mapsto \pi_{U,D}(\boldsymbol{d}) = \sum_{k\geq i\geq -\infty} d_iu_i$$

En fait, cette application n'est pas définie sur tout l'ensemble $D^{\mathbb{Z}}$, mais uniquement sur les suites pour lesquelles il existe un entier N tel que $d_i = 0$ pour tout entier $i \geq N$.

Ainsi, le mot 100110 est la représentation binaire de l'entier 38. La suite U utilisée est la suite des puissances[8] de 2 et l'alphabet D est simplement $\{0,1\}$. Dans ce même système, le réel $2,75$ s'écrit 10.11.

8. Il s'agit donc de β-numération, avec $\beta = 2$.

Chapitre 2

Représentation des entiers

Nous introduisons dans ce chapitre le système de numération en base $\frac{p}{q}$ et nous montrons comment l'algorithme MD permet d'obtenir des représentations des entiers. Nous définissons le langage $L_{\frac{p}{q}}$ qui est l'ensemble des mots, écrits sur l'alphabet canonique du système, qui représentent les entiers positifs. Nous étudions alors les propriétés de ce langage : certains résultats connus pour la numération en base entière sont toujours valables et d'autres ne le sont plus. Enfin, nous montrons que la conversion entre deux alphabets peut être réalisée par un transducteur à droite fini, ce qui sera très utile lors de l'étude des représentations des réels au chapitre 4.

2.1 L'algorithme de Division Modifié et le système de numération en base $\frac{p}{q}$

Soient $p > q \geq 1$ deux entiers premiers entre eux et soit N un entier strictement positif. Nous notons $N_0 = N$ et pour tout entier positif i, nous notons

$$qN_i = pN_{i+1} + a_i,$$

où a_i est le reste de la division euclidienne de qN_i par p. Par conséquent, a_i appartient à l'ensemble $\{0, 1, \ldots, p-1\}$.

Puisque N_{i+1} est strictement inférieur à N_i, cette opération ne peut s'effectuer qu'un nombre fini de fois, jusqu'à ce que N_{k+1} soit égal à 0 pour un certain k. La suite de divisions successives pour $i = 0$ jusque $i = k$ décrit donc un algorithme, appelé *algorithme de Division Modifié*, ou *algorithme MD*. Celui-ci produit, à partir d'un entier N donné, $k+1$ nombres $a_0, a_1, \ldots, a_k$ vérifiant

$$N = \sum_{i=0}^{k} \frac{a_i}{q} \left(\frac{p}{q}\right)^i.$$

En effet, nous obtenons successivement

$$\begin{aligned} qN_0 &= pN_1 + a_0 \quad , \quad N_0 = \frac{p}{q}N_1 + \frac{a_0}{q} \\ qN_1 &= pN_2 + a_1 \quad , \quad N_0 = \left(\frac{p}{q}\right)^2 N_2 + \frac{p}{q}\frac{a_1}{q} + \frac{a_0}{q} \\ &\vdots \end{aligned}$$

Nous disons que le mot $a_k a_{k-1} \cdots a_0$ est la *représentation en base* $\frac{p}{q}$ de l'entier N. Remarquons que, contrairement à ce qui se fait usuellement en β-numération, ce mot est calculé de droite à gauche, i.e. le chiffre le moins siginificatif en premier.

Nous montrons au théorème 2.1.2 que cette représentation est unique. Nous l'appelons donc le $\frac{p}{q}$-*développement* de N et nous le notons $\langle N \rangle_{\frac{p}{q}}$. Par convention, le $\frac{p}{q}$-développement de 0 est le mot vide ε.

Illustrons ceci en considérant[1] $p = 3$ et $q = 2$; les représentations des entiers sont donc écrites sur l'alphabet $A = \{0, 1, 2\}$. Le tableau 2.1 fournit le $\frac{3}{2}$-développement des onze premiers entiers.

n	$\langle n \rangle_{\frac{3}{2}}$
0	ε
1	2
2	21
3	210
4	212
5	2101
6	2120
7	2122
8	21011
9	21200
10	21202

TABLE 2.1 – Tableau du $\frac{3}{2}$-développement des onze premiers entiers

Soit U la suite définie par

$$U = \left(\frac{1}{q}\left(\frac{p}{q}\right)^i\right)_{i \in \mathbb{Z}} .$$

Nous disons que U, associé à l'alphabet $A = \{0, 1, \ldots, p-1\}$ est le *système de numération* en base $\frac{p}{q}$. Dans le cas où $q = 1$, il s'agit exactement du système de numération en base p.

1. Nous utilisons la base $\frac{3}{2}$ comme exemple tout au long de ce travail.

Remarquons que cette définition n'est pas celle rencontrée usuellement pour les systèmes de β-numération : la suite U n'est pas la suite des puissances de $\frac{p}{q}$, mais bien la suite de ces puissances divisées par q. De plus, les chiffres a_i obtenus ne sont pas des entiers strictement inférieurs à $\frac{p}{q}$, mais des entiers dont le quotient par q est strictement inférieur à $\frac{p}{q}$. Cependant, la représentation étant unique, l'algorithme MD fournit la même développement que l'algorithme glouton classique pour le cas $q = 1$.

Le lemme suivant assure qu'une des plus importantes propriétés des systèmes de numération en base entière est également valable pour des bases rationnelles.

Lemme 2.1.1. *Soit $\pi : A^* \to \mathbb{R}$ l'application évaluation associée au système de numération en base $\frac{p}{q}$. Pour tout entier positif k, la restriction de π à A^k est injective.*

Démonstration. Soient $u = a_{k-1}a_{k-2}\cdots a_0$ et $v = b_{k-1}b_{k-2}\cdots b_0$ deux mots distincts sur A qui ont la même image par π. Il vient

$$\sum_{i=0}^{k-1} \frac{a_i}{q}\left(\frac{p}{q}\right)^i = \sum_{i=0}^{k-1} \frac{b_i}{q}\left(\frac{p}{q}\right)^i$$

et donc

$$\sum_{i=0}^{k-1} (a_i - b_i)\left(\frac{p}{q}\right)^i = 0.$$

Le polynôme de $\mathbb{Z}[X]$, $\sum_{i=0}^{k-1}(a_i - b_i)X^i$, s'annule donc en $\frac{p}{q}$ et est alors divisible par le polynôme minimum de $\frac{p}{q}$, $qX - p$, ce qui est impossible car le terme indépendant $a_0 - b_0$ est strictement inférieur à p en valeur absolue.

□

Il est évident que l'application π n'est pas injective sur tout A^*, car pour tout mot u et pour tout entier positif k, on a $\pi(0^k u) = \pi(u)$. Cependant, le lemme 2.1.1 implique que c'est la seule possibilité, i.e.

$$|u| > |v| \text{ et } \pi(u) = \pi(v) \;\Rightarrow\; u = 0^k v, \text{ avec } k = |u| - |v|.$$

Théorème 2.1.2. *Tout entier positif N admet un unique $\frac{p}{q}$-développement qui est une représentation entière. Il s'agit de l'unique représentation finie en base $\frac{p}{q}$ de N.*

Démonstration. Soit $a_{k-1}a_{k-2}\cdots a_0$ le $\frac{p}{q}$-développement de N fourni par l'algorithme MD. Supposons qu'il existe une autre représentation finie de N dans le système U de la forme

$$e_{l-1}e_{l-2}\cdots e_0.e_{-1}e_{-2}\cdots e_{-m},$$

avec $e_{-m} \neq 0$. Il vient :

$$N = \sum_{i=0}^{k-1} \frac{a_i}{q}\left(\frac{p}{q}\right)^i = \sum_{i=-m}^{l-1} \frac{e_i}{q}\left(\frac{p}{q}\right)^i.$$

En multipliant par $\left(\frac{p}{q}\right)^m$, nous obtenons

$$\sum_{i=0}^{k-1} a_i \left(\frac{p}{q}\right)^{m+i} = \sum_{i=-m}^{l-1} e_i \left(\frac{p}{q}\right)^{m+i},$$

et donc $\pi(e_{l-1}e_{l-2}\cdots e_0e_{-1}e_{-2}\cdots e_{-m}) = \pi(a_{k-1}a_{k-2}\cdots a_0 0^m) = \left(\frac{p}{q}\right)^m N$, d'où une contradiction car $e_{-m} \neq 0$ et π est injectif sur A^{m+p}, avec $p = \max\{l, k\}$.

Le mot $a_{k-1}a_{k-2}\cdots a_0$ est donc l'unique représentation en base $\frac{p}{q}$ de N (avec la condition $a_{k-1} \neq 0$). Il s'agit bien d'une représentation entière et nous notons

$$\langle N \rangle_{\frac{p}{q}} = a_{k-1}a_{k-2}\cdots a_0.$$

$\square$

2.2 L'ensemble des $\frac{p}{q}$-développements des entiers

Notons $L_{\frac{p}{q}}$ l'ensemble des $\frac{p}{q}$-développements des entiers positifs ; il s'agit d'un langage de A^*. Si $q = 1$, ce langage est l'ensemble des mots qui ne commencent pas par un 0. Si nous n'exigeons plus cette dernière condition [2], nous obtenons alors tous les mots de A^*. Nous étudions ici ce qu'il en est lorsque $q \neq 1$.

Vu le théorème 2.1.2, le langage $L_{\frac{p}{q}}$ est en bijection avec $\mathbb{N}$.

2.2.1 Contexte à droite

Par construction, le langage $L_{\frac{p}{q}}$ est préfixiel. En effet, si le mot $a_{k-1}a_{k-2}\cdots a_0$ est le $\frac{p}{q}$-développement d'un entier positif N_0, alors ce mot est obtenu par l'algorithme MD, c'est-à-dire par la suite de divisions

$$\begin{aligned}
qN_0 &= pN_1 + a_0 \\
qN_1 &= pN_2 + a_1 \\
&\vdots \\
qN_{k-2} &= pN_{k-1} + a_{k-2} \\
qN_{k-1} &= a_{k-1}
\end{aligned}$$

Soit l un entier apartenant à $\{1, 2, \dots, k-1\}$. Montrons que le mot $a_{k-1}a_{k-2}\cdots a_l$ est le $\frac{p}{q}$-développement d'un entier positif, i.e. montrons que $\pi(a_{k-1}a_{k-2}\cdots a_l) \in \mathbb{N}$. Par définition,

$$\pi(a_{k-1}a_{k-1}\cdots a_l) = \frac{a_l}{q} + \frac{a_{l+1}}{q}\frac{p}{q} + \frac{a_{l+2}}{q}\left(\frac{p}{q}\right)^2 + \cdots + \frac{a_{k-1}}{q}\left(\frac{p}{q}\right)^{k-l-1}.$$

Or, le système d'éqations ci-dessus nous donne

$$\begin{aligned}
N_{k-1} &= \frac{a_{k-1}}{q} \\
N_{k-2} &= \frac{a_{k-1}}{q}\frac{p}{q} + \frac{a_{k-2}}{q} \\
&\vdots \\
N_l &= \frac{a_{k-1}}{q}\left(\frac{p}{q}\right)^{k-l-1} + \frac{a_{k-2}}{q}\left(\frac{p}{q}\right)^{k-l-2} + \cdots + \frac{a_l}{q}
\end{aligned}$$

Nous avons donc $\pi(a_{k-1}a_{k-2}\cdots a_l) = N_l \in \mathbb{N}$.

2. C'est-à-dire si nous n'exigeons plus que les représentations ne commencent pas par un 0.

Si $q \neq 1$, l'observation du tableau 2.1 montre que le langage $L_{\frac{p}{q}}$ n'est en général pas suffixiel. En effet, le mot $\langle 8 \rangle_{\frac{3}{2}} = 21011$ appartient au langage $L_{\frac{3}{2}}$ alors que son suffixe 11 ne lui appartient pas car $\pi(11) = \frac{1}{2} + \frac{3}{2}\frac{1}{2} = 1,25$

Dans la suite et sauf mention explicite du contraire, nous supposons toujours $q \neq 1$.

Soient n et k deux entiers positifs. Nous notons $CD_k(n)$ l'ensemble des mots sur A de longueur au plus k dont la concaténation avec le $\frac{p}{q}$-développement de n est un mot de $L_{\frac{p}{q}}$, i.e.

$$CD_k(n) = \left\{ w \in A^{\leq k} \mid \langle n \rangle_{\frac{p}{q}} w \in L_{\frac{p}{q}} \right\}.$$

Lemme 2.2.1. *Soient m et n deux entiers positifs. Un mot w sur A de longueur k appartient à $CD_k(n) \cap CD_k(m)$ si et seulement si n et m sont congru modulo q^k et dans ce cas, $CD_k(n) = CD_k(m)$. Donc*

$$\left(\langle n \rangle_{\frac{p}{q}} w \in L_{\frac{p}{q}} \text{ et } \langle m \rangle_{\frac{p}{q}} w \in L_{\frac{p}{q}} \right) \Rightarrow CD_k(n) = CD_k(m).$$

Démonstration. Supposons d'abord que le mot w appartient aux deux langages $CD_k(n)$ et $CD_k(m)$. Les images respectives par π des mots $\langle n \rangle_{\frac{p}{q}} w$ et $\langle m \rangle_{\frac{p}{q}} w$ sont donc des entiers positifs, i.e.

$$\pi(\langle n \rangle_{\frac{p}{q}} w) \in \mathbb{N} \quad \text{et} \quad \pi(\langle m \rangle_{\frac{p}{q}} w) \in \mathbb{N}.$$

Or, $\pi(\langle n \rangle_{\frac{p}{q}} w) = n \left(\frac{p}{q}\right)^k + \pi(w)$ et $\pi(\langle m \rangle_{\frac{p}{q}} w) = m \left(\frac{p}{q}\right)^k + \pi(w)$. Il vient alors

$$\pi(\langle n \rangle_{\frac{p}{q}} w) - \pi(\langle m \rangle_{\frac{p}{q}} w) = \left(\frac{p}{q}\right)^k (n - m) \in \mathbb{Z}$$

et puisque p et q sont premier entre eux, m et n sont congrus modulo q^k.

Supposons maintenant que m et n sont congrus modulo q^k. Pour tout entier positif h inférieur à k, les entiers m et n sont donc congrus modulo q^h. Soit w un mot de longueur h dans $CD_k(m)$, nous avons

$$\pi(\langle m \rangle_{\frac{p}{q}} w) = m \left(\frac{p}{q}\right)^h + \pi(w) \in \mathbb{N}.$$

Il suffit de vérifier que le mot w appartient à $CD_k(n)$, i.e. que $\pi(\langle n \rangle_{\frac{p}{q}} w)$ est un entier positif. Il est clair que $n \left(\frac{p}{q}\right)^h + \pi(w)$ est positif, il suffit donc de vérifier que c'est un entier. Or, nous avons

$$\pi(\langle n \rangle_{\frac{p}{q}} w) - \pi(\langle m \rangle_{\frac{p}{q}} w) = (n - m) \left(\frac{p}{q}\right)^h$$

et $(n-m)\left(\frac{p}{q}\right)^h$ appartient à $\mathbb{Z}$ car n et m sont congru modulo q^h, ce qui suffit car $\pi(\langle m\rangle_{\frac{p}{q}} w)$ appartient à $\mathbb{N}$.

$\square$

Le lemme 2.2.1 implique que le langage $L_{\frac{p}{q}}$ n'est pas régulier.

Corollaire 2.2.2. *Si $q \neq 1$, alors le langage $L_{\frac{p}{q}}$ n'est pas régulier.*

Démonstration. Il suffit de montrer que l'automate minimal $\mathcal{A}_{\frac{p}{q}}$ du langage $L_{\frac{p}{q}}$ possède un nombre infini d'états. Cet automate est accessible et l'ensemble de ses états est

$$\left\{[w] = w^{-1}.L_{\frac{p}{q}} \mid w \in A^*\right\}.$$

L'état initial est $[\varepsilon]$ et l'ensemble des états finals est

$$\left\{[w] \mid w \in L_{\frac{p}{q}}\right\}.$$

À tout mot du langage $L_{\frac{p}{q}}$, il correspond donc un état final de l'automate. Par conséquent, il suffit de montrer que cet état est unique, car dans ce cas, l'automate possède un nombre infini [3] d'états finals et donc d'états.

Soient m et n deux entiers positifs et soient u et v leurs $\frac{p}{q}$-développements respectifs. Les mots u et v correspondent à un même état final de l'automate minimal si et seulement s'ils sont équivalent pour la congruence de Nérode, i.e. si et seulement si

$$u^{-1}.L_{\frac{p}{q}} = v^{-1}.L_{\frac{p}{q}}.$$

Or, pour tout mot $\langle x\rangle_{\frac{p}{q}}$ appartenant à $L_{\frac{p}{q}}$, nous avons

$$\langle x\rangle_{\frac{p}{q}}^{-1}.L_{\frac{p}{q}} = \bigcup_{k\in\mathbb{N}} CD_k(x).$$

Par conséquent, les mots u et v correspondent au même état de $\mathcal{A}_{\frac{p}{q}}$ si et seulement si pour tout entier positif k, nous avons

$$CD_k(n) = CD_k(m)$$

et, par le lemme 2.2.1, si et seulement si n et m sont congru modulo q^k pour tout k. Puisque l'entier q est strictement supérieur à 1, les entiers n et m sont congru modulo q^k pour tout k si et seulement s'ils sont égaux, d'où la conclusion.

$\square$

Dans la même lignée que le lemme 2.2.1, le lemme suivant donne une précision sur les suffixes qui sont une puissance d'un mot donné.

3. En effet, le langage $L_{\frac{p}{q}}$ est en bijection avec $\mathbb{N}$ et est donc infini.

Lemme 2.2.3. *Soient $w \in L_{\frac{p}{q}}$ et $w = uv$ une factorisation propre*[4] *de w. Le mot uv^k appartient au langage $L_{\frac{p}{q}}$ seulement si $q^{(k-1)|v|}$ divise $\pi(w) - \pi(u)$.*

Démonstration. Le mot uv^k appartient au langage $L_{\frac{p}{q}}$ qui est préfixiel. Le mot uv^{k-1} lui appartient donc aussi et nous avons

$$\pi(uv^k) - \pi(uv^{k-1}) \in \mathbb{Z}.$$

Or,

$$\begin{aligned}
\pi(uv^k) - \pi(uv^{k-1}) &= \left(\frac{p}{q}\right)^{|v|} \left(\pi(uv^{k-1}) - \pi(uv^{k-2})\right) \\
&= \left(\frac{p}{q}\right)^{2|v|} \left(\pi(uv^{k-2}) - \pi(uv^{k-3})\right) \\
&\vdots \\
&= \left(\frac{p}{q}\right)^{(k-1)|v|} (\pi(uv) - \pi(u))
\end{aligned}$$

Puisque p et q sont premiers entre eux, $q^{(k-1)|v|}$ divise bien $\pi(w) - \pi(u)$.

□

Ce lemme trouve son utilité dans la suite pour montrer que la fermeture[5] de $L_{\frac{p}{q}}$ ne contient pas de mots infinis ultimement périodiques. Combiné au *lemme de la pompe*[6], il permet également d'établir le corollaire 2.2.5.

Commençons par rappeler le lemme de la pompe[7].

Lemme 2.2.4 (Lemme de la pompe)**.** *Soit L un langage algébrique. Il existe un entier positif n tel que pour tout mot z dans L de longueur au moins n, nous pouvons écrire $z = uvwxy$, avec*

1. $|vx| \neq 0$ *;*
2. $|vwx| \leq n$ *;*
3. $\forall i \neq 0$, $uv^iwx^iy \in L$.

Corollaire 2.2.5. *Si $q \neq 1$, alors le langage $L_{\frac{p}{q}}$ n'est pas algébrique.*

---

4. i.e. les mots u et v sont non vide.
5. Par "fermeture de $L_{\frac{p}{q}}$", nous entendons l'ensemble des mots infinis sur A dont les préfixes finis appartiennent à $0^*L_{\frac{p}{q}}$; cette notion sera expliquée en temps voulu.
6. Il s'agit de la version du lemme concernant les langages algébriques.
7. La démonstration de ce lemme se trouve dans [7].

Démonstration. Supposons le langage $L_{\frac{p}{q}}$ algébrique et notons n l'entier fourni par le lemme de la pompe. Soit z un mot de $L_{\frac{p}{q}}$ de longueur au moins[8] n. Nous pouvons écrire $z = uvwxy$ avec $|vx| \neq 1$, $|vwx| \leq n$ et $uv^iwx^iy \in L_{\frac{p}{q}}$ pour tout entier positif i.

Supposons v non vide[9]. Pour tout entier positif i, le mot uv^iwx^iy appartient au langage $L_{\frac{p}{q}}$ qui est préfixiel ; le mot uv^i lui appartient donc également. Par le lemme 2.2.3, $q^{(i-1)|v|}$ divise $\pi(uv) - \pi(u)$ pour tout i, d'où une contradiction car v est non vide et q strictement supérieur à 1.

□

2.2.2 Suffixes

Nous savons déjà que le langage $L_{\frac{p}{q}}$ n'est en général pas suffixiel. Le résultat de cette sous-section stipule que tout mot sur A est suffixe d'un mot de $L_{\frac{p}{q}}$.

Proposition 2.2.6. *Pour tout entier positif k, pour tout mot w sur A de longueur k, il existe un unique entier positif n strictement inférieur à p^k tel que le mot w est un suffixe de longueur k du $\frac{p}{q}$-développement de tous les entiers positifs m congrus à n modulo p^k.*

Démonstration. Soit n un entier positif. Notons $n_0 = n$ et appliquons k fois la division de l'algorithme MD, il vient

$$\begin{aligned} qn_0 &= pn_1 + a_0 \\ qn_1 &= pn_2 + a_1 \\ &\vdots \\ qn_{k-1} &= pn_k + a_{k-1} \end{aligned}$$

Il découle de cette suite de divisions que

$$q^k n_0 = p^k n_k + q^k \pi(a_{k-1}a_{k-2}\cdots a_0). \tag{2.1}$$

Si nous appliquons le même procédé à un autre entier positif m et que nous soustrayons les deux équations obtenues, nous obtenons

$$q^k(n_0 - m_0) = p^k(n_k - m_k) + q^k\left(\pi(a_{k-1}a_{k-2}\cdots a_0) - \pi(b_{k-1}b_{k-2}\cdots b_0)\right). \tag{2.2}$$

8. L'existence d'un tel mot est évidente car le langage $L_{\frac{p}{q}}$ est en bijection avec $\mathbb{N}$ et le nombre de mots sur A de longueur au plus n est fini.

9. Puisque $|vx| \neq 0$, au moins un des deux mots v et x est non vide ; le même raisonnement s'applique si nous supposons x non vide.

Montrons que les deux entiers m et n sont congrus modulo p^k si et seulement si $a_{k-1}a_{k-2}\cdots a_0 = b_{k-1}b_{k-2}\cdots b_0$.

Tout d'abord, si les deux mots sont identiques, ils ont la même image par l'application π et l'équation 2.2 devient simplement

$$q^k(n-m) = p^k(n_k - m_k),$$

ce qui suffit puisque p et q sont premiers entre eux.

Ensuite, si m et n sont congru modulo p^k, alors $q^k(n_0 - m_0) = ip^k$ pour un entier i. Il découle alors l'équation 2.2 que $\pi(a_{k-1}\cdots a_0) - \pi(b_{k-1}\cdots b_0)$ est un multiple de p^k. Or,

$$\pi(a_{k-1}a_{k-2}\cdots a_0) - \pi(b_{k-1}b_{k-2}\cdots b_0) = \sum_{j=0}^{k-1} \frac{a_j - b_j}{q}\left(\frac{p}{q}\right)^j.$$

Puisque $a_j, b_j \in \{0, 1\ldots, p-1\}$ pour tout j, nous obtenons $a_j - b_j = 0$ pour tout j.

Il suffit maintenant de conclure. Pour tout entier positif n, nous pouvons construire un mot w de longueur k. Ce mot est bien un suffixe de $\langle n \rangle_{\frac{p}{q}}$: soit la longueur de $\langle n \rangle_{\frac{p}{q}}$ est strictement supérieure à k et cela suffit vu que les premiers chiffres[10] se calculent en itérant à nouveau la division de l'algorithme MD, soit sa longueur est inférieure et dans ce cas, le mot obtenu est simplement $\langle n \rangle_{\frac{p}{q}}$ auquel nous avons rajouté un préfixe composé uniquement de 0, préfixe que nous pouvons donc supprimer[11]. De plus, vu ce qui précède, ce mot trouvé est également suffixe de $\langle m \rangle_{\frac{p}{q}}$ pour tout entier m congru à n modulo p^k. En particulier, si les deux entiers m et n sont strictement inférieurs à p^k, alors ils sont égaux. Ainsi, un mot w sur A de longueur k n'est suffixe de $\langle n \rangle_{\frac{p}{q}}$ que pour un seul entier n inférieur à p^k. Puisque les deux ensembles A^k et $\{0, 1, \ldots, p^k - 1\}$ ont même cardinal, chaque mot de A^k est préfixe d'un unique $\langle n \rangle_{\frac{p}{q}}$, avec $n \in \{0, 1, \ldots, p^k - 1\}$.

□

2.2.3 L'odomètre

La proposition 2.2.6 peut être interprétée en parallèle avec la construction de la célèbre *Pascaline*, machine inventée par B. Pascal. Il s'agit d'une machine composée d'un ensemble de roues dentées agencées en un engrenage comme suit : lorsqu'une roue a effectué une rotation complète, elle envoit une impulsion à la roue située à sa gauche qui fait tourner cette dernière d'un cran.

Imaginons la Pascaline étendue à gauche jusqu'à l'infini, chaque roue étant indexée par p chiffres[12], à savoir les chiffres $0, 1, \ldots, p-1$. B. Pascal comptait en base 10,

10. Sous-entendu les chiffres les plus à gauche.
11. Remarquons que dans ce cas, l'équation 2.1 reste vraie car nous avons alors $n_k = 0$.
12. Un chiffre pour chaque cran de la roue.

la Pascaline originale comptait donc 10 chiffres par roue, mais n'importe quel entier positif p aurait fait l'affaire.

Considérons l'exemple $p = 3$ pour la suite de l'explication ; les chiffres indexant les roues sont donc 0 ,1 et 2. Chaque roue de la Pascaline est raccordée à un bras, celui-ci indiquant la position de la roue. Nous pouvons donc voir la Pascaline comme une horloge possédant autant d'aiguilles que de roues. Supposons qu'à chaque seconde, la roue la plus à droite tourne d'un cran. Le bras correspondant passe successivement par $0, 1, 2, 0, 1, 2, 0, 1, \ldots$. À chaque passage de la première aiguille devant le 0, la seconde avance également et passe donc à son tour par $0, 1, 2, 0, 1, \ldots$. Après n secondes, nous lisont un certain mot sur l'horloge, écrit sur l'alphabet $A = \{0, 1, 2\}$. Ce mot est exactement l'écriture en base 3 de l'entier n. Réciproquement, tout mot w sur A apparaît exactement une fois, au temps $\pi(w)$, où $\pi(w)$ est le nombre dont l'écriture en base 3 est w.

Imaginons maintenant que la machine est "boostée" en ce sens qu'une unité de rotation fait passer une aiguille de deux crans à la place d'un cran. La première aiguille passe donc successivement par $0, 2, 1, 0, 2, 1, \ldots$ et à chaque passage devant le 0, que l'aiguille s'arrête ou non, une impulsion est envoyée à la deuxième aiguille qui se décale alors également de 2 crans. La suite des mots que nous lisons alors sur l'horloge est exactement la suite des $\frac{p}{q}$-développements des entiers, i.e. le langage $L_{\frac{p}{q}}$, ordonnés par l'ordre généalogique.

Nous pouvons maintenant interpréter la proposition 2.2.6. Si nous construisions une telle machine finie [13], avec par exemple k roues, la proposition 2.2.6 stipule que le comportement de cette horloge serait périodique de période p^k et que chaque configuration possible des k roues apparaîtrait exactement une fois durant un cycle.

L'odomètre d'un système de numération est une machine qui réalise la fonction *successeur*, i.e. qui reçoit un mot représentant un entier n et qui en fournit un mot représentant l'entier $n + 1$. Vu la description de la Pascaline "boostée", il est facile de construire un transducteur à droite lettre-à-lettre qui réalise l'odomètre pour le système de numération en base $\frac{p}{q}$. La figure suivante représente l'odomètre pour la base $\frac{3}{2}$.

2.2.4 Ordre sur les nombres, ordre sur les mots

Dans une base entière, l'ordre sur les nombres et l'ordre généalogique sur les mots (écrits sur l'alphabet canonique) qui représentent ces nombres coïncident, et il en est de même pour l'ordre lexicographique sur les mots de même longueur (auxquels nous avons parfois rajouté un préfixe composé de 0 uniquement). La même propriété est valable dans le système de numération en base $\frac{p}{q}$, du moment que nous ne considèrons que les mots de $L_{\frac{p}{q}}$.

13. Ce qui est plus réaliste qu'une infinie.

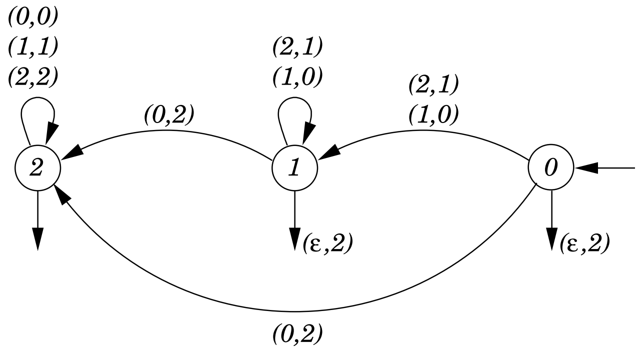

FIGURE 2.1 – L'odomètre pour le sytème de numération en base $\frac{3}{2}$

Proposition 2.2.7. *Soient u et v deux mots appartenant à $L_{\frac{p}{q}}$. Nous avons*

$$u \preceq v \Leftrightarrow \pi(u) \leq \pi(v).$$

Démonstration. Notons $u = a_{k-1}a_{k-2}\cdots a_0$, $m = \pi(u)$, $v = b_{l-1}b_{l-2}\cdots b_0$ et $n = \pi(v)$. Vu le théorème 2.1.2, nous avons déjà $u = v$ si et seulement si $\pi(u) = \pi(v)$. Il suffit donc de montrer le résultat pour une inégalité stricte.

Supposons l supérieur à k et procédons par récurrence sur l. Le cas $l = 1$ est évident. Supposons le résultat acquis pour une longueur égale à $l - 1$ et démontrons le pour l. Posons $u' = a_{k-1}a_{k-2}\cdots a_1$, $m' = \pi(u')$, $v' = b_{l-1}b_{l-2}\cdots b_1$ et $n' = \pi(v')$. Il vient

$$n - m = \frac{p}{q}(n' - m') + \frac{1}{q}(b_o - a_0).$$

Par définition de l'ordre généalogique, nous avons

$$u \prec v \Leftrightarrow ((u' \prec v') \text{ ou } (u' = v' \text{ et } a_0 < b_0))$$

Or, par hypothèse de récurrence, nous avons $u' \prec v'$ si et seulement si $n' - m' \geq 1$. Comme a_0 et b_0 appartiennent à $\{0, 1, \ldots, p-1\}$, nous avons $b_0 - a_0 \geq -(p-1)$ et donc

$$\begin{aligned} n - m &= \frac{1}{q}\left(p\,(n' - m') + b_0 - a_0\right) \\ &\geq \frac{1}{q}\left(p - (p-1)\right) \\ &\geq \frac{1}{q}. \end{aligned}$$

De plus, si $u' = v'$, nous obtenons $n - m = \frac{b_0 - a_0}{q} > 0$, d'où la conclusion.

□

Corollaire 2.2.8. *Soient u et v deux mots de $0^*L_{\frac{p}{q}}$ de même longueur. Nous avons $u \sqsubseteq v$ si et seulement si $\pi(u) \leq \pi(v)$.*

Démonstration. Si u et v sont des mots de $L_{\frac{p}{q}}$, il suffit de remarquer que l'ordre généalogique et l'ordre lexicographique coïncident sur les mots de même longueur. Sinon, c'est évident.

$\square$

Remarquons qu'il est vraiment nécessaire que les mots que nous considérons appartiennent à $L_{\frac{p}{q}}$ (resp. à $0^*L_{\frac{p}{q}}$). En effet, si nous considérons notre exemple de base, à savoir le système de numération en base $\frac{3}{2}$, nous avons

$$\pi(10) = \frac{3}{4} < \pi(2) = 1$$

alors que

$$10 \succ 2$$

et nous avons

$$\pi(2000) = \frac{27}{8} < \pi(0212) = 4$$

alors que

$$2000 \sqsupset 0212.$$

2.3 Conversion entre alphabets

Une autre propriété des sytèmes de numération en base entière [14] qui s'étend aux systèmes de numération en base rationnelle est que la conversion entre des alphabets peut être réalisée par un transducteur à droite fini. Plus précisément, si D est un alphabet contenant A et si w est un mot écrit sur D, alors il existe un transducteur à droite fini qui fournit un mot w' sur A tel que [15] $\pi(w) = \pi(w')$. Pour rappel, le transducteur représenté à la figure 2.2 réalise la conversion entre l'alphabet $\{0,1,2\}$ et l'alphabet binaire.

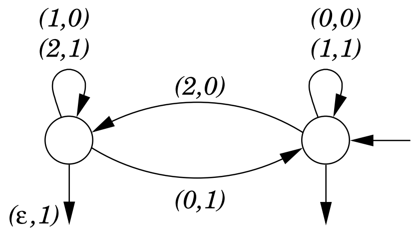

FIGURE 2.2 – Convertisseur de $\{0,1,2\}^*$ dans le système binaire.

Ainsi, nous considérons un alphabet D qui contient A ; rien n'empêche que D contienne des nombres négatifs. La *conversion* est une application χ_D définie sur D^* et à valeurs dans A^* qui commute avec π, i.e. qui préserve la valeur numérique :

$$\forall w \in D^*, \qquad \pi(\chi_D(w)) = \pi(w).$$

Considérons le mot $w = 1210$ représentant l'entier $\pi(w) = 0.1 + 1.2 + 2.4 + 1.8 = 18$ dans le système binaire. Si ce mot constitue le mot d'entrée fourni au transducteur ci-dessus, le mot de sortie obtenu est alors $w' = 10010$ et sa valeur numérique est $\pi(w') = 0.1 + 1.2 + 0.4 + 0.8 + 1.16 = 18$.

Proposition 2.3.1. *Pour tout alphabet D contenant A, la conversion χ_D peut être réalisée par un transducteur à droite, lettre-à-lettre, séquentiel et fini, noté $\mathcal{C}_D$.*

Démonstration. Soit $\mathcal{U}_D = (\mathbb{Z}, D \times A, E, \{0\}, \omega)$ le transducteur lettre-à-lettre infini dont l'ensemble des transitions est défini par

$$(z, (d,a), z') \in E \Leftrightarrow qz + d = pz' + a.$$

Le couple (z', a) de $\mathbb{Z} \times A$ est univoquement déterminé par le couple (z, d) ; ce transducteur est donc séquentiel.

14. Voir [2].

15. $\pi(w)$ est défini exactement de la même façon que $\pi(w')$, i.e. la fonction évaluation est étendue à l'ensemble des mots écrits sur l'alphabet D.

Si nous définissons la fonction finale ω par

$$\omega(z) = \begin{cases} \langle z \rangle_{\frac{p}{q}} & \text{si } z \in \mathbb{N}, \\ \emptyset & \text{sinon,} \end{cases}$$

le transducteur $\mathcal{U}_D$, vu comme un transducteur à droite, réalise la conversion des mots w sur D dont la valeur numérique est positive. En effet, soit $w = d_{k-1}d_{k-2}\cdots d_0$ un mot sur D de valeur numérique positive. Si ce mot décrit un chemin dans $\mathcal{U}_D$ débutant en 0 et arrivant en $z \in \mathbb{N}$, alors le mot sur A qui lui est associé par le transducteur est $\langle z \rangle_{\frac{p}{q}} w'$. Il suffit de vérifier que ces deux mots ont la même valeur numérique. Procédons par récurrence sur la longueur k de w.

Si $k = 0$, c'est évident. Si $k = 1$, alors $w = d_0 \in D$ et $\pi(w) = \frac{d_0}{q}$. Il existe une transition dans $\mathcal{U}_D$ partant de 0, aboutissant en z et d'étiquette (d_0, a_0). Il vient alors $d_0 = pz + a_0$ et en divisant par q, nous obtenons

$$\pi(w) = \frac{d_0}{q} = \frac{p}{q}z + \frac{a_0}{q} = \pi(\langle z \rangle_{\frac{p}{q}} a_0).$$

Supposons le résultat acquis pour les mots de longueur $k-1$ et montrons-le pour ceux de longueur k. Nous considérons le suffixe $v = d_{k-2}d_{k-3}\cdots d_0$ de w. Ce mot décrit un chemin dans $\mathcal{U}_D$ débutant en 0 et aboutissant en z'. Par l'hypothèse de récurrence, il vient $\pi(d_{k-2}d_{k-3}\cdots d_0) = \pi(\langle z' \rangle_{\frac{p}{q}} a_{k-2}a_{k-3}\cdots a_0)$ et donc

$$\sum_{i=0}^{k-2} \frac{d_i - a_i}{q} \left(\frac{p}{q}\right)^i = \left(\frac{p}{q}\right)^{k-1} z'.$$

Or, il existe une transition $(z', (d_{k-1}, a_{k-1}), z)$ dans E ; il en découle que

$$z' = \frac{p}{q}z + \frac{a_{k-1} - d_{k-1}}{q}.$$

Ainsi, nous avons

$$\sum_{i=0}^{k-1} \frac{d_i - a_i}{q} \left(\frac{p}{q}\right)^i = \left(\frac{p}{q}\right)^{k} z,$$

ou encore

$$\pi(w) = \sum_{i=0}^{k-1} \frac{d_i}{q} \left(\frac{p}{q}\right)^i = \left(\frac{p}{q}\right)^{k} z + \sum_{i=0}^{k-1} \frac{a_i}{q} \left(\frac{p}{q}\right)^i = \pi(\langle z \rangle_{\frac{p}{q}} a_{k-1}a_{k-2}\cdots a_0).$$

Il reste maintenant à montrer que la partie accessible $\mathcal{C}_D$ de $\mathcal{U}_D$ est finie. Nous pouvons supposer sans perte de généralité que l'alphabet D est un intervalle de $\mathbb{Z}$ contenant A. Notons e (resp. f) le plus grand (resp. le plus petit) chiffre dans D. Nous obtenons

$$e \geq p - 1 \quad \text{et} \quad f \leq 0,$$

où au moins une de ces deux inégalités est sticte. Vu la manière dont sont définies les transitions dans $\mathcal{U}_D$, il est possible, à partir d'un état z, d'atteindre l'état $z+1$ (resp. l'état $z-1$) si et seulement s'il exite des chiffres d et a dans D et A respectivement tels que

$$qz + d = p(z+1) + a \qquad (\text{resp. } qz + d = p(z-1) + a).$$

En remaniant cette égalité, nous parvenons à écrire z en fonction de a, d, p et q sous la forme

$$z = \frac{d-a-p}{p-q} \qquad (\text{resp. } z = \frac{d-a+p}{p-q}).$$

Ainsi, l'état positif accessible maximal et l'état accessible minimal sont respectivement

$$z_{\max} = \left\lfloor \frac{e-p}{p-q} \right\rfloor + 1 \quad \text{et} \quad z_{\min} = \left\lceil \frac{f-(p-1)+p}{p-q} \right\rceil - 1 = \left\lceil \frac{f+1}{p-q} \right\rceil - 1.$$

La partie accessible $\mathcal{C}_D$ est donc finie, ce qui conclut.

□

Cette proposition permet de voir l'addition entière comme un cas particulier de la conversion χ_D. En effet, pour additionner deux entiers m et n, nous considérons leurs $\frac{p}{q}$-développements respectifs $\langle m \rangle_{\frac{p}{q}}$ et $\langle n \rangle_{\frac{p}{q}}$, et nous additionnons les a_i et b_i correspondants. Le mot obtenu est alors un mot écrit sur l'alphabet $\{0, 1, \ldots, 2(p-1)\}$ et nous pouvons dès lors réaliser la conversion de ce mot sur l'alphabet A par un transducteur à droite fini.

Les deux figures suivantes représentent respectivement le transducteur réalisant l'addition dans le sytème de numération en base $\frac{3}{2}$ et le transducteur qui convertit les mots écrits sur l'alphabet $D = \{, -1, 0, 1, 2\}$ dans le système en base $\frac{3}{2}$. L'intérêt de présenter ce dernier transducteur apparaîtra à la section 4.2. Pour éviter toute confusion, nous notons $\bar{1}$ le chiffre -1 de l'alphabet D.

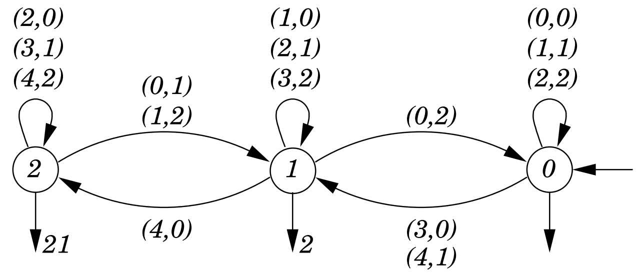

FIGURE 2.3 – Le convertisseur pour l'addition.

Remarquons que la conversion χ_D est définie pour tout mot w sur D, même si la valeur numérique de ce mot n'est pas entière. Par contre, si cette valeur est un entier positif, alors la conversion de w dans l'alphabet A est l'unique $\frac{p}{q}$-développement de $\pi(w)$.

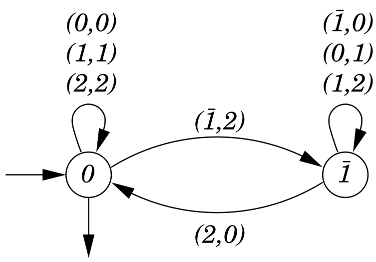

FIGURE 2.4 – Le convertisseur pour $D = \{-1, 0, 1, 2\}$.

Chapitre 3

L'arbre $T_{\frac{p}{q}}$

Le graphe de Cayley du monoïde libre A^* est un arbre infini d'arité p. Chaque noeud est étiqueté par un mot sur A et possède p fils. De plus, chaque branche entre un père et un de ses fils est étiquetée par une chiffre de A et l'étiquette d'un noeud est précisément la concaténation de l'unique chemin qui part de la racine de l'arbre et qui arrive à ce noeud.

Puisque le langage $L_{\frac{p}{q}}$ est préfixiel, nous pouvons le représenter dans le graphe de Cayley de A^* comme un sous-arbre obtenu en supprimant certaines branches. Nous remplaçons également les étiquettes des noeuds par leurs valeurs numériques respectives et nous obtenons alors ce que nous appellons l'*arbre (infini)* $T_{\frac{p}{q}}$.

Cet arbre sera la base de la représentation des réels dans le système de numération en base $\frac{p}{q}$. Nous présentons ici une description "entière" de l'arbre, basée qur la définition d'une famille d'applications de $\mathbb{N}$ dans $\mathbb{N}$ qui se révelera très utile dans l'étude des chemins infinis.

3.1 Construction de l'arbre $T_{\frac{p}{q}}$

Pour tout chiffre a dans A, nous considérons l'application partielle τ_a définie sur $\mathbb{N}$ par

$$\tau_a(n) = \frac{1}{q}(pn + a) \qquad \text{si}\ [1]\ \frac{1}{q}(pn + a) \in \mathbb{N}.$$

Nous notons alors $\mathsf{d}(n)$ l'ensemble des chiffres a dans A pour lesquels $\tau_a(n)$ est défini. Nous notons également $\mathsf{Md}(n)$ (resp. $\mathsf{md}(n)$) le plus grand (resp. le plus petit) chiffre

1. Si ce n'est pas le cas, $\tau_a(n)$ n'est pas défini.

dans[2] $\mathsf{d}(n)$, i.e.

$$\begin{aligned}\mathsf{Md}(n) &= \max\{\mathsf{d}(n)\}\\ \mathsf{md}(n) &= \min\{\mathsf{d}(n)\}\end{aligned}$$

L'arbre $T_{\frac{p}{q}}$ est l'arbre étiqueté[3] infini construit comme suit. Les noeuds sont étiquetés par des entiers positifs et les branches par des chiffres de A ; la racine de l'arbre est étiquetée par 0. Les fils d'un noeud étiqueté par un entier positif n sont étiquetés par $\tau_a(n)$ pour des chiffres a dans $\mathsf{d}(n)$ et la branche entre deux noeuds d'étiquettes respectives n et $\tau_a(n)$ est étiquetée par a. Il est clair que l'arbre ainsi construit représente le langage[4] $L_{\frac{p}{q}}$. En effet, si, au départ d'un noeud s étiqueté par n, une branche étiquetée par a nous amène dans l'état t étiqueté par $\tau_a(n)$, alors $\tau_a(n)$ est un entier positif et on a bien $\tau_a(n) = \frac{1}{q}(pn + a) = \pi(\langle n \rangle_{\frac{p}{q}}\, a)$.

Nous appellons *étiquette de chemin* d'un noeud s l'étiquette du chemin partant de la racine de l'arbre et arrivant en s ; nous la notons $\mathsf{p}(s)$. Un noeud v est un *embranchement* s'il a au moins deux fils, i.e. si le cardinal de $\mathsf{d}(\pi(\mathsf{p}(v)))$ est supérieur à 2. Enfin, nous notons $I_{\frac{p}{q}}$ le sous-arbre de $T_{\frac{p}{q}}$ obtenu en supprimant les noeuds dont l'étiquette de chemin commence par un 0. La figure 3.1 représente l'arbre $T_{\frac{3}{2}}$ et son sous-arbre $I_{\frac{3}{2}}$ en est la partie grise.

La manière dont l'arbre a été construit implique que si deux noeuds ont la même étiquette, alors ils sont racines de deux sous-arbres isomorphes de $T_{\frac{p}{q}}$ et il découle directement du lemme 2.2.1 que la réciproque est également vraie, i.e. deux noeuds qui ont des étiquettes différentes sont racines de deux sous-arbres distincts de $T_{\frac{p}{q}}$.

Puisque deux noeud de $I_{\frac{p}{q}}$ n'ont jamais la même étiquette, il vient :

Proposition 3.1.1. *Si $q \neq 1$, alors deux sous-arbres de $I_{\frac{p}{q}}$ ne sont jamais isomorphes.*

Le lemme suivant découle de l'algorithme MD et de la construction de l'arbre. Il sera fréquemment utilisé dans les démonstrations à venir, le plus souvent sans référence explicite.

Lemme 3.1.2. *Pour tout entier positif n, nous avons*

1. $\mathsf{md}(n) = \mathsf{d}(n) \cap \{0, 1, \ldots, q-1\}$ *et* $\mathsf{Md}(n) = \mathsf{d}(n) \cap \{p-q, p-q+1, \ldots, p-1\}$.
2. $a \in \mathsf{d}(n)$ *et* $a + q \in A \;\Rightarrow\; a + q \in \mathsf{d}(n)$.
3. $a, a+q \in \mathsf{d}(n) \;\Rightarrow\; \tau_{a+q}(n) = \tau_a(n) + 1$.
4. $\mathsf{md}(n+1) = \mathsf{Md}(n) + q - p$ *et* $\tau_{\mathsf{md}(n+1)}(n+1) = \tau_{\mathsf{Md}(n)}(n) + 1$.

2. Il s'agit donc respectivement du plus grand et du plus petit chiffre dans A tels que $\tau_a(n)$ est défini.

3. Les noeuds et les branches sont étiquetés.

4. En fait, ce langage est représenté par le sous-arbre $I_{\frac{p}{q}}$ de $T_{\frac{p}{q}}$ défini ci-après. L'arbre $T_{\frac{p}{q}}$ représente quant à lui le langage $0^* L_{\frac{p}{q}}$.

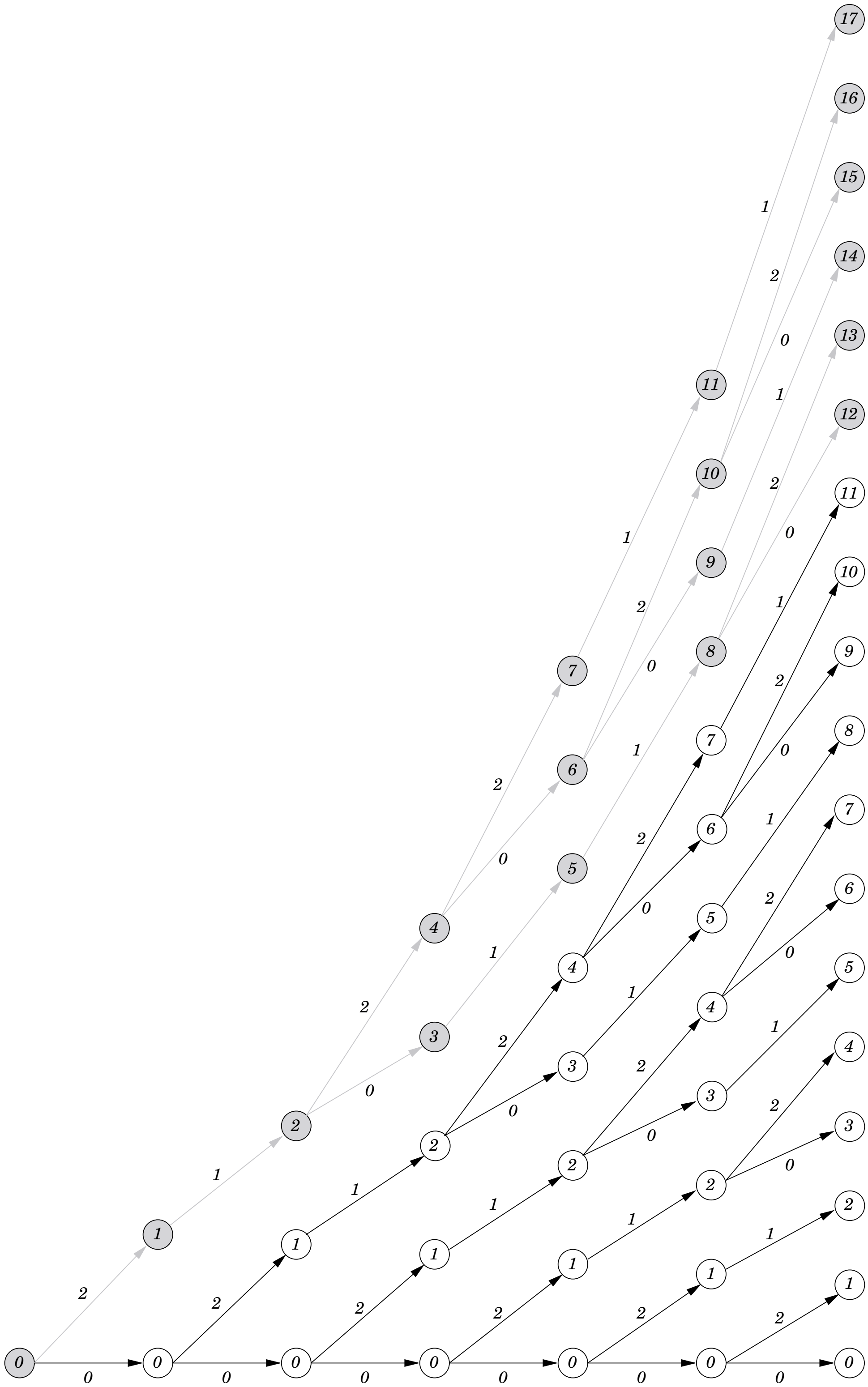

FIGURE 3.1 – L'arbre $T_{\frac{3}{2}}$ et son sous-arbre $I_{\frac{3}{2}}$.

5. *L'étiquette d'un noeud s est $\pi(\mathsf{p}(s))$.*

Démonstration. Le premier point découle de la longueur des sous-intervalles choisis. Un chiffre a de A appartient à $\mathsf{d}(n)$ si et seulement si $pn+a$ est un multiple de q. Or, les ensembles $\{0,1,\ldots,q-1\}$ et $\{p-q,p-q+1,\ldots,p-1\}$ contiennent respectivement les q plus petits et les q plus grands chiffres dans A. Dans chacun de ces sous-intervalles, il existe donc un unique chiffre vérifiant la condition.

Les deuxième et troisième points sont évident : par définition, nous avons

$$\begin{aligned}\tau_{a+q}(n) &= \frac{1}{q}(pn+a+q)\\ &= \tau_a(n)+1\end{aligned}$$

Passons au quatrième point. Si $a=\mathsf{Md}(n)\in\{p-q,p-q+1,\ldots,p-1\}$, alors $a-p+q\in\{0,1,\ldots,q-1\}$ et il suffit donc de montrer que $\tau_{a-p+q}(n+1)=\tau_a(n)+1$. Or,

$$\begin{aligned}\tau_{a-p+q}(n+1) &= \frac{1}{q}\left(p(n+1)+a-p+q\right)\\ &= \frac{1}{q}\left(pn+a\right)+1\\ &= \tau_a(n)+1\end{aligned}$$

Le dernier point découle directement de la construction de l'arbre $T_{\frac{p}{q}}$.

□

En particulier, nous avons directement

Corollaire 3.1.3. *Pour tout entier positif n, $\mathsf{md}(n)=d \Leftrightarrow \tau_d(n)=\left\lceil \frac{p}{q}n\right\rceil$.*

Suite à ce corollaire, nous définissons la suite d'entiers $(G_k)_{k\in\mathbb{N}}$ par

$$G_0=1 \quad \text{et} \quad \forall k\in\mathbb{N},\ G_{k+1}=\left\lceil \left(\frac{p}{q}\right) G_k\right\rceil.$$

Pour tout entier k strictement positif, l'entier G_k est donc l'élément minimum parmis les étiquettes des noeuds accessibles à partir d'un noeud d'étiquette G_{k-1}, i.e.

$$G_k=\tau_{\mathsf{md}(G_{k-1})}(G_{k-1}).$$

Par induction sur k, il vient alors

Proposition 3.1.4. *Les noeuds de niveau k dans $T_{\frac{p}{q}}$, ordonnés par étiquette de chemin dans l'ordre lexicographique (ou généalogique), sont étiquetés par les entiers de 0 à G_k-1.*

Démonstration. Procédons par récurrence sur k. Le résultat est évident pour $k = 0$. Supposons-le vrai pour k et démontrons-le pour $k + 1$.

Les noeuds de niveau k dans l'arbre sont étiquetés (après avoir été ordonné par étiquette de chemin dans l'ordre lexicographique) par les entiers de 0 à $G_k - 1$. Au niveau $k + 1$, il est clair que les chemins qui passent par un noeud de niveau k et d'étiquette $i \in \{0, 1, \dots G_k - 1\}$ sont plus petits que ceux qui passent par un noeud de niveau k et d'étiquette $j > i$. Il suffit donc, pour ordonner tous les chemins, d'ordonner ceux qui passent par les mêmes noeuds au niveau k, ce qui se fait sans peine.

Montrons que ces noeuds sont bien étiquetés par les entiers de 0 à $G_{k+1} - 1$.

Tout d'abord, les points *3* et *4* du lemme 3.1.2 assurent que deux noeuds du niveau $k + 1$ n'auront jamais la même étiquette. En effet, si un noeud d'étiquette n est un embranchement, alors il existe deux chiffres a et $a+q$ dans $\mathsf{d}(n)$ et nous avons $\tau_{a+q}(n) = \tau_a(n)+1$ et deux noeuds de niveau $k+1$ provenant d'un même noeud de niveau k n'ont donc jamais la même étiquette. De plus, puisque $\tau_{\mathsf{md}(n+1)}(n + 1) = \tau_{\mathsf{Md}(n)}(n) + q - p$, deux noeuds de niveau $k + 1$ provenant de noeuds "consécutifs" de niveau k n'ont pas la même étiquette non plus. Par transitivité, deux noeuds de niveau $k + 1$ provenant de noeuds différents de niveau k n'ont pas la même étiquette.

Ensuite, il est clair qu'il existe un noeud de niveau $k+1$ ayant 0 comme étiquette : $\tau_0(0) = 0$. De plus, l'étiquette maximum des noeuds de niveau $k + 1$ est $G_{k+1} - 1$. En effet, G_{k+1} est par définition l'élément minimum parmi les étiquettes de noeuds accessibles à partir d'un noeud d'étiquette G_k, i.e.

$$G_{k+1} = \tau_{\mathsf{md}(G_k)}(G_k).$$

Or, vu le lemme 3.1.2, nous avons $\mathsf{md}(G_k) = \mathsf{Md}(G_k - 1) + q - p$ et

$$\tau_{\mathsf{Md}(G_k-1)}(G_k - 1) = \tau_{\mathsf{md}(G_k)}(G_k) - 1 = G_{k+1} - 1.$$

Le noeud d'étiquette maximum au niveau $k + 1$ est donc celui ayant pour étiquette $G_{k+1} - 1$.

Enfin, par un raisonnement analogue à ce qui précède, il est clair que tout entier dans $\{0, 1, \dots, G_{k+1} - 1\}$ est l'étiquette d'un noeud de niveau $k + 1$. □

Nous étudions la suite des G_k plus en détail à la section 3.4.

3.2 Mots minimaux et maximaux

Les chemins infinis dans l'arbre $T_{\frac{p}{q}}$ sont exploités au chapitre 4 pour définir la représentation des nombres réels. Nous ne considérons ici que certains chemins infinis particuliers.

Ainsi, nous notons $\mathsf{W}(n)$ (resp. $\mathsf{w}(n)$) l'étiquette d'un chemin infini qui débute dans un noeud d'étiquette n et qui emprunte toujours la branche d'étiquette maximale (resp. minimale). Un tel mot est dit *maximal* (resp. *minimal*) dans $T_{\frac{p}{q}}$.

La proposition suivante découle alors directement du lemme 3.1.2.

Proposition 3.2.1. *1. Pour tout entier positif n, $\mathsf{W}(n) \in \{p-q, p-q+1, \ldots, p-1\}^{\mathbb{N}}$ et $\mathsf{w}(n) \in \{0, 1, \ldots, q-1\}^{\mathbb{N}}$.*

2. Respectivement, soit $\boldsymbol{u}$ l'étiquette d'un chemin infini dans $T_{\frac{p}{q}}$.
Si $\boldsymbol{u} \in \{p-q, p-q+1, \ldots, p-1\}^{\mathbb{N}}$ (resp. si $\boldsymbol{u} \in \{0, 1, \ldots, q-1\}^{\mathbb{N}}$), alors[5] il existe un entier positif n tel que $\boldsymbol{u} = \mathsf{W}(n)$ (resp. $\boldsymbol{u} = \mathsf{w}(n)$).

3. Pour tout eniter positif n, la différence composante à composante entre $\mathsf{W}(n)$ et $\mathsf{w}(n+1)$ est le mot infini $(p-q)^{\omega}$.

Nous introduisons des notations pour deux mots particuliers :

$$\boldsymbol{t}_{\frac{p}{q}} = \mathsf{W}(0) \quad \text{et} \quad \boldsymbol{g}_{\frac{p}{q}} = \mathsf{w}(1).$$

Le mot infini $\boldsymbol{t}_{\frac{p}{q}}$ est l'élément maximal[6] parmi les étiquettes de chemins infinis dans $T_{\frac{p}{q}}$ qui commencent à la racine. De plus, puisque $\tau_q(0) = 1$, le mot infini $q\boldsymbol{g}_{\frac{p}{q}}$ est l'élément minimal parmi les étiquettes de chemins infinis de $I_{\frac{p}{q}}$ qui commencent à la racine. Remarquons également que, pour toute base $\frac{p}{q}$, le mot infini 0^{ω} est l'élément minimal parmi les étiquettes de chemins infinis dans $T_{\frac{p}{q}}$. Enfin, si $q = 1$, i.e. si nous travaillons dans une base entière, alors

$$\mathsf{W}(n) = (p-1)^{\omega} \quad \text{et} \quad \mathsf{w}(n) = 0^{\omega} \ \forall n \in \mathbb{N}.$$

Dans notre exemple de base, nous avons

$$\begin{aligned} \boldsymbol{t}_{\frac{3}{2}} &= 212211122121122121211221\cdots \\ \boldsymbol{g}_{\frac{3}{2}} &= 101100011010011010100110\cdots \end{aligned}$$

Dans le chapitre 2, les mots de $L_{\frac{p}{q}}$, et donc les étiquettes des chemins finis dans $T_{\frac{p}{q}}$, étaient indicés de droite à gauche, où si l'on préfère, de gauche à droite par des entiers positifs décroissants jusqu'à 0. Nous avons mentionné à la section 1.4 la possibilité

5. Remarquons que cette propriété n'est pas vraie pour tout mot infini de $A^{\mathbb{N}}$. Par exemple, le mot infini $\boldsymbol{a} = 1^{\omega}$ n'est l'étiquette d'aucun chemin dans $T_{\frac{3}{2}}$.

6. L'ordre utilisé ici est évidemment l'ordre lexicographique, l'ordre généalogique n'étant pas défini sur l'ensemble des mots infinis.

de prolonger les indices vers la droite, après le séparateur, en utilisant des entiers négatifs, éventuellement jusque $-\infty$, dans le but de représenter la partie fractionnaire des nombres [7]. Lorsque nous travaillons avec des nombres ayant uniquement une partie fractionnaire, comme c'est le cas à la section suivante, il paraît naturel de changer notre manière d'indicer. Ainsi, des tels nombres seront représentés par des mots indicés par des eniers positifs de gauche à droite dans l'ordre croissant, en commençant par 1. En particulier, nous notons

$$\boldsymbol{g}_{\frac{p}{q}} = g_1 g_2 g_3 \cdots .$$

Vu la définition de la suite $(G_k)_{k \in \mathbb{N}}$, il vient alors facilement

Corollaire 3.2.2. *$G_0 = \pi(q) = 1$ et $\forall k \in \mathbb{N}$, $G_k = \pi(q g_1 g_2 \cdots g_k)$.*

7. Dans le cas où nous ne travaillons pas avec des entiers.

3.3 Evaluation des mots infinis dans $T_{\frac{p}{q}}$

Vu les conventions prises pour les indices, l'application évaluation devient

$$\forall \boldsymbol{a} = a_1 a_2 a_3 \cdots \in A^{\mathbb{N}}, \qquad \pi(.\boldsymbol{a}) = \frac{1}{q} \sum_{i \geq 1} a_i \left(\frac{q}{p}\right)^i.$$

L'utilisation du séparateur à la gauche du mot infini a pour but d'indiquer la position de l'indice 0 ; il permet également de bien faire la distinction entre le cas qui nous occupe à présent et tout ce qui précède.

Soient $\boldsymbol{a} \in A^{\mathbb{N}}$ un mot infini sur A et $x = \pi(.\boldsymbol{a})$ sa valeur numérique. Il vient

$$\begin{array}{rcl} \forall h \in \mathbb{N}, \quad \left(\frac{p}{q}\right)^h x & = & \pi(a_1 a_2 \cdots a_h . a_{h+1} a_{h+2} \cdots) \\ x & = & \lim_{h \to +\infty} \left(\frac{q}{p}\right)^h \pi(a_1 a_2 \cdots a_h) \\ & = & \lim_{h \to +\infty} \left(\frac{q}{p}\right)^h \pi(\boldsymbol{a}_{[h]}) \end{array}$$

Tout comme dans le cas des bases entières, nous avons

Proposition 3.3.1. *L'application $\pi : A^{\mathbb{N}} \to \mathbb{R}$ est continue.*

Démonstration. Soit $\boldsymbol{a} \in A^{\mathbb{N}}$ un mot infini. Pour tout $\varepsilon > 0$, il nous faut trouver un $\eta > 0$ tel que pour tout mot infini $\boldsymbol{b} \in A^{\mathbb{N}}$ vérifiant $\delta(\boldsymbol{a}, \boldsymbol{b}) < \eta$, nous avons $|\pi(.\boldsymbol{a}) - \pi(.\boldsymbol{b})| < \varepsilon$.

Un mot infini $\boldsymbol{b} \in A^{\mathbb{N}}$ est tel que $\delta(\boldsymbol{a}, \boldsymbol{b}) < 2^{-n}$ si et seulement s'il peut s'écrire $\boldsymbol{b} = \boldsymbol{a}_{[n]} \boldsymbol{c}$, avec $\boldsymbol{c} \in A^{\mathbb{N}}$. De plus, la différence composante à composante entre $\boldsymbol{a}$ et $\boldsymbol{b}$ est au plus le mot infini $\boldsymbol{x} = 0^n (p-1)^\omega$ dont la valeur numérique est

$$\pi(.\boldsymbol{x}) = \left(\frac{q}{p}\right)^n \frac{p-1}{p-q}.$$

En effet, nous avons

$$\begin{array}{rcl} \left(\frac{p}{q}\right)^n \pi(.\boldsymbol{x}) & = & \pi\left(.(p-1)^\omega\right) \\ & = & \frac{p-1}{q} \sum_{i \geq 1} \left(\frac{q}{p}\right)^i \\ & = & \frac{p-1}{q} \left(\frac{1}{1 - \frac{q}{p}} - 1\right) \\ & = & \frac{p-1}{p-q}. \end{array}$$

Ainsi, pour tout $\varepsilon > 0$, il existe un entier n tel que $\left(\frac{q}{p}\right)^n \frac{p-1}{p-q} < \varepsilon$ et il suffit alors de choisir $\eta = 2^{-n}$.

□

Le choix du réel η dans la démonstration précédente ne dépendant pas du mot infini $\boldsymbol{a}$ de départ, la fonction évaluation π est même uniformément continue.

Nous notons $W_{\frac{p}{q}}$ le sous-ensemble de $A^{\mathbb{N}}$ composé des étiquettes de chemins infinis commençant à la racine de $T_{\frac{p}{q}}$. Par définition, les préfixes finis des mots de $W_{\frac{p}{q}}$ sont donc des éléments de $0^*L_{\frac{p}{q}}$ et il vient

Proposition 3.3.2. *Si $q > 1$, alors il n'existe pas de mot ultimement périodique dans $W_{\frac{p}{q}}$, hormis 0^ω.*

Démonstration. Procédons par l'absurde. Si le mot infini $\boldsymbol{a} \neq 0^\omega$ appartient au langage $W_{\frac{p}{q}}$ et est ultimement périodique, il peut s'écrire

$$\boldsymbol{a} = uv^\omega, \qquad \text{avec } v \neq \varepsilon.$$

Puisque $\boldsymbol{a}$ est un mot de $W_{\frac{p}{q}}$, tout préfixe fini $\boldsymbol{a}_{[n]}$ de $\boldsymbol{a}$ appartient au langage $0^*L_{\frac{p}{q}}$. Considérons les facteurs de $\boldsymbol{a}$ du type[8] $u'v^k \in L_{\frac{p}{q}}$. Pour tout entier positif k, le mot $u'v^k$ appartient au langage $L_{\frac{p}{q}}$ et par le lemme 2.2.3, $q^{(k-1)|v|}$ divise alors $\pi(u'v) - \pi(u)$, d'où une contradiction car v est non vide et q est strictement supérieur à 1.

□

Les lemmes suivants donnent des informations sur l'évaluation des mots de $W_{\frac{p}{q}}$. Le lemme 3.3.3 fournit par exemple une approximation de la valeur numérique d'un mot de $W_{\frac{p}{q}}$.

Lemme 3.3.3. *Soient $\boldsymbol{a} = a_1a_2a_3\cdots$ un mot de $W_{\frac{p}{q}}$ et x sa valeur numérique. Pour tout entier positif k, nous avons*

$$\left\lfloor \left(\frac{p}{q}\right)^k x \right\rfloor = \pi(a_1a_2\cdots a_k) + \rho_k(x),$$

avec $\rho_k(x) = \lfloor \pi(.a_{k+1}a_{k+2}\cdots)\rfloor < \frac{p-1}{p-q}$.

Démonstration. Il est clair que, pour tout entier positif k,

$$\begin{aligned}\left(\frac{p}{q}\right)^k x &= \pi(a_1a_2\cdots a_k.a_{k+1}a_{k+2}\cdots)\\ &= \pi(a_1a_2\cdots a_k) + \pi(.a_{k+1}a_{k+2}\cdots)\end{aligned}$$

8. Ces facteurs sont obtenus en supprimant toutes les occurences de 0 situées au début du mot.

Or, $a_1a_2\cdots a_k$ est un préfixe fini de $\boldsymbol{a}$ et sa valeur numérique $\pi(a_1a_2\cdots a_k)$ est donc entière. Il vient alors

$$\left\lfloor \left(\frac{p}{q}\right)^k x \right\rfloor = \pi(a_1a_2\cdots a_k) + \lfloor \pi(.a_{k+1}a_{k+2}\cdots)\rfloor$$

Il suffit donc de montrer que $\lfloor \pi(.a_{k+1}a_{k+2}\cdots)\rfloor = \rho_k(x)$ est strictement inférieur à $\frac{p-1}{p-q}$. Remarquons que

$$\begin{aligned}
\pi(.(p-1)^\omega) &= \frac{p-1}{q}\sum_{i\geq 1}\left(\frac{q}{p}\right)^i \\
&= \frac{p-1}{q}\left(\frac{1}{1-\frac{p}{q}}-1\right) \\
&= \frac{p-1}{p-q}
\end{aligned}$$

Or, le mot infini $\boldsymbol{a}$ appartenant à $W_{\frac{p}{q}}$, il n'est pas ultimement périodique et est donc strictement inférieur au mot infini $(p-1)^\omega$, d'où la conclusion.

□

Le lemme suivant stipule qu'à tout embranchement de l'arbre $T_{\frac{p}{q}}$, il correspond un réel admettant au moins deux représentations dans le système de numération.

Lemme 3.3.4. *Soient v un embranchement de $T_{\frac{p}{q}}$ et $n = \pi(\mathsf{p}(v))$ son étiquette. Soient également a_1 et $b_1 = a_1 + q$ des chiffres dans*[9] *$\mathsf{d}(n)$. Soient enfin*[10] *$m_1 = \tau_{a_1}(n)$ et $m_2 = \tau_{b_1}(n) = m_1 + 1$.*
Si nous notons $\mathsf{W}(m_1) = a_2a_3\cdots$ et $\mathsf{w}(m_2) = b_2b_3\cdots$, alors

$$\pi(.a_1a_2a_3\cdots) = \pi(.b_1b_2b_3\cdots).$$

Démonstration. Par la proposition 3.2.1, la différence composante à composante entre $\mathsf{W}(m_1)$ et $\mathsf{w}(m_2)$ est le mot infini $(p-q)^\omega$. Il vient alors

$$\begin{aligned}
\pi(.a_1a_2\cdots) - \pi(.b_1b_2\cdots) &= \sum_{i\geq 1}\frac{a_i}{q}\left(\frac{q}{p}\right)^i - \sum_{i\geq 1}\frac{b_i}{q}\left(\frac{q}{p}\right)^i \\
&= \frac{1}{q}\left(a_1\frac{q}{p} - b_1\frac{q}{p} + \sum_{i\geq 2}(a_i-b_i)\left(\frac{q}{p}\right)^i\right) \\
&= \frac{1}{q}\left(-q\frac{q}{p} + (p-q)\sum_{i\geq 2}\left(\frac{q}{p}\right)^i\right) \\
&= \frac{1}{q}\left(-q\frac{q}{p} + (p-q)\frac{q^2}{p(p-q)}\right) \\
&= 0,
\end{aligned}$$

9. L'existence de tels chiffres est assurée par le fait que le noeud v est un embranchement.
10. La deuxième égalité découle du lemme 3.1.2.

ce qui conclut. $\square$

Nous définissons maintenant deux réels $\omega_{\frac{p}{q}}$ et $\gamma_{\frac{p}{q}}$ par

$$\omega_{\frac{p}{q}} = \pi(.\boldsymbol{t}_{\frac{p}{q}}) \quad \text{et} \quad \gamma_{\frac{p}{q}} = \pi(.q\boldsymbol{g}_{\frac{p}{q}}).$$

La racine de l'arbre étant un embranchement [11], il vient directement

$$\boldsymbol{t}_{\frac{p}{q}} = \mathsf{W}(0) \quad \text{et} \quad \boldsymbol{g}_{\frac{p}{q}} = \mathsf{w}(1)$$

et par le lemme 3.3.4, nous obtenons alors

$$\frac{q}{p}\omega_{\frac{p}{q}} = \pi(.0\boldsymbol{t}_{\frac{p}{q}}) = \pi(.q\boldsymbol{g}_{\frac{p}{q}}) = \gamma_{\frac{p}{q}}.$$

Nous avons par exemple

$$\omega_{\frac{3}{2}} = 1,62227050288476731595695098 2\cdots.$$

Le lemme suivant est en quelque sorte la réciproque du lemme 3.3.4, mais sa démonstration se révèle un peu plus technique.

Lemme 3.3.5. *Supposons $q > 1$. Soient un entier $k > \frac{q-1}{p-q}$ et $r = \left\lceil \frac{q-2}{p-q} \right\rceil$. Pour tout entier positif n et pour tous mots u et v sur A de même longueur l vérifiant $\pi(u) = n$ et $\pi(v) = n + k$, nous avons*

$$\pi(.v\mathsf{w}(n+k)) - \pi(.u\mathsf{w}(n)) \geq \left(\frac{q}{p}\right)^{l+r} \omega_{\frac{p}{q}}.$$

Démonstration. Commençons par introduire les notations suivantes, inspirées par le corollaire 3.1.3 :

$$\begin{aligned} \forall n \in \mathbb{N}, \qquad & \mu(n) = \tau_{\mathsf{md}(n)}(n) = \left\lceil \frac{p}{q} n \right\rceil \\ \text{et} \qquad & \mu^{i+1}(n) = \mu(\mu^i(n)) \end{aligned}$$

Pour tout entier n, $\mu(n)$ est donc l'élément minimum parmi les étiquettes des noeuds accessibles à partir d'un noeud d'étiquette n dans l'arbre $T_{\frac{p}{q}}$. De même, $\mu^i(n)$ est l'élément minimum parmi les étiquettes des noeuds accessibles à partir d'un noeud d'étiquette n en empuntant un chemin de longueur [12] i.

Il nous faut d'abord montrer quelques inégalités avant de pouvoir établir le lemme.

11. En effet, deux de ses fils sont $\tau_0(0) = 0$ et $\tau_q(0) = 1$.
12. La longueur d'un chemin étant le nombre de branches qui composent ce chemin.

Premièrement, il est clair que

$$\left\lceil \frac{p}{q}n \right\rceil - \frac{q-1}{q} \leq \frac{p}{q}n$$

et donc que

$$\left\lceil \frac{p}{q}(n+k) \right\rceil - \left\lceil \frac{p}{q}n \right\rceil \geq \frac{p}{q}(n+k) - \left(\frac{p}{q}n + \frac{q-1}{q} \right).$$

Vu le choix de k, il vient alors facilement

$$\left\lceil \frac{p}{q}(n+k) \right\rceil - \left\lceil \frac{p}{q}n \right\rceil > k.$$

Puisque les deux membres de cette inégalité sont des entiers, nous obtenons, en utilisant les notations introduites ci-dessus,

$$\forall n \in \mathbb{N}, \qquad \mu(n+k) \geq \mu(n) + k + 1.$$

Par récurrence sur i, il vient alors

$$\mu^i(n+k) \quad \geq \quad \mu^i(n) + k + i. \tag{3.1}$$

En effet, nous avons

$$\begin{aligned} \mu^i(n+k) &= \mu(\mu^{i-1}(n+k)) \\ &\geq \mu\left(\mu^{i-1}(n) + k + i - 1\right) \\ &\geq \mu^i(n) + k + i. \end{aligned}$$

Ensuite, pour tous n, m et z dans $\mathbb{N}$, Nous avons

$$\left\lceil \frac{p}{q}(n+m+z) \right\rceil \geq \left\lceil \frac{p}{q}n \right\rceil + \left\lceil \frac{p}{q}m \right\rceil + \frac{p}{q}z - 2\frac{q-1}{q}. \tag{3.2}$$

Montrons maintenant que

$$\frac{p}{q}(k+r) - 2\frac{q-1}{q} \geq k + r. \tag{3.3}$$

Nous avons

$$\frac{p}{q}k \geq k + 1 \quad \Leftrightarrow \quad k \geq \frac{q}{p-q}.$$

Or, par hypothèse, k est un entier strictement supérieur à $\frac{q-1}{p-q}$ et l'inégalité ci-dessus est donc vérifiée. Il en découle que

$$\frac{p}{q}(k+r) - 2\frac{q-1}{q} \geq k + 1 + \frac{p}{q}r - 2\frac{q-1}{q}.$$

En modifiant le membre de droite, nous obtenons

$$\frac{p}{q}(k+r) - 2\frac{q-1}{q} \geq k + r + \frac{(p-q)r-(q-2)}{q}$$

et, vu le choix de r, le terme $\frac{(p-q)r-(q-2)}{q}$ est positif et l'inégalité est donc bien démontrée.

Par induction sur j, il s'ensuit que

$$\mu^j(n+k+r) \geq \mu^j(n) + G_{j-1} + k + r. \tag{3.4}$$

En effet, l'inégalité est vérifiée pour $j=1$ par l'inégalité 3.1 et les inégalités 3.2 et 3.3 nous donnent alors

$$\begin{aligned} \mu\left(\mu^j(n) + G_{j-1} + k + r\right) &= \left\lceil \frac{p}{q}\mu^j(n) + \frac{p}{q}G_{j-1} + \frac{p}{q}(k+r) \right\rceil \\ &\geq \mu^{j+1}(n) + G_j + \frac{p}{q}(k+r) - 2\frac{q-1}{q} \\ &\geq \mu^{j+1}(n) + G_j + k + r \end{aligned}$$

Nous pouvons maintenant démontrer le résultat attendu.

Pour simplifier les notations, nous posons $\boldsymbol{a} = u\mathsf{w}(n)$ et $\boldsymbol{b} = v\mathsf{w}(n+k)$. Par définition de $\mathsf{w}(n)$ et de $\mu^i(n)$, nous avons $\pi(\boldsymbol{a}_{[h]}.) = \mu^i(n)$ (resp. $\pi(\boldsymbol{b}_{[h]}.) = \mu^i(n+k)$) pour $h = l+i$. Par l'inégalité 3.4, il vient alors, $\forall j \in \mathbb{N}$ et $h = l + r + (j-1)$,

$$\begin{aligned} \pi(\boldsymbol{b}_{[h]}.) - \pi(\boldsymbol{a}_{[h]}.) &= \mu^j\left(\mu^{r-1}(n+k)\right) - \mu^j\left(\mu^{r-1}(n)\right) \\ &\geq \mu^j\left(\mu^{r-1}(n) + k + r - 1\right) - \mu^j\left(\mu^{r-1}(n)\right) \\ &\geq \mu^j\left(\mu^{r-1}(n) + k + r\right) - \mu^j\left(\mu^{r-1}(n)\right) \\ &\geq G_{j-1} + k + r - 1 \end{aligned}$$

Il découle alors de la section 3.3 que

$$\pi(.\boldsymbol{b}) - \pi(.\boldsymbol{a}) \geq \lim_{j\to+\infty} \left(\frac{q}{p}\right)^{l+r+j-1} (G_{j-1} + k + r) = \left(\frac{q}{p}\right)^{l+r} \left(\frac{p}{q} \lim_{j\to+\infty} \left(\frac{q}{p}\right)^j G_{j-1}\right).$$

Or, par le corollaire 3.2.2, nous avons $G_{j-1} = \pi(qg_1g_2\cdots g_{j-1})$ et donc

$$\frac{p}{q} \lim_{j\to+\infty} \left(\frac{q}{p}\right)^j G_{j-1} = \frac{p}{q}\pi(.q\boldsymbol{g}) = \frac{p}{q}\gamma_{\frac{p}{q}} = \omega_{\frac{p}{q}},$$

ce qui termine la démonstration.

□

La figure suivante représente à nouveau l'arbre $T_{\frac{3}{2}}$, mais cette fois, chaque noeud s possède une ordonnée précise, à savoir $\pi(.\mathsf{p}(s))$. Les noeuds d'un même niveau dans l'arbre ont la même abscisse (comme à la figure 3.1) et la distance entre deux noeuds consécutifs de même niveau est multipliée par $\frac{q}{p}$ à chaque niveau, ce qui donne un aspect fractal.

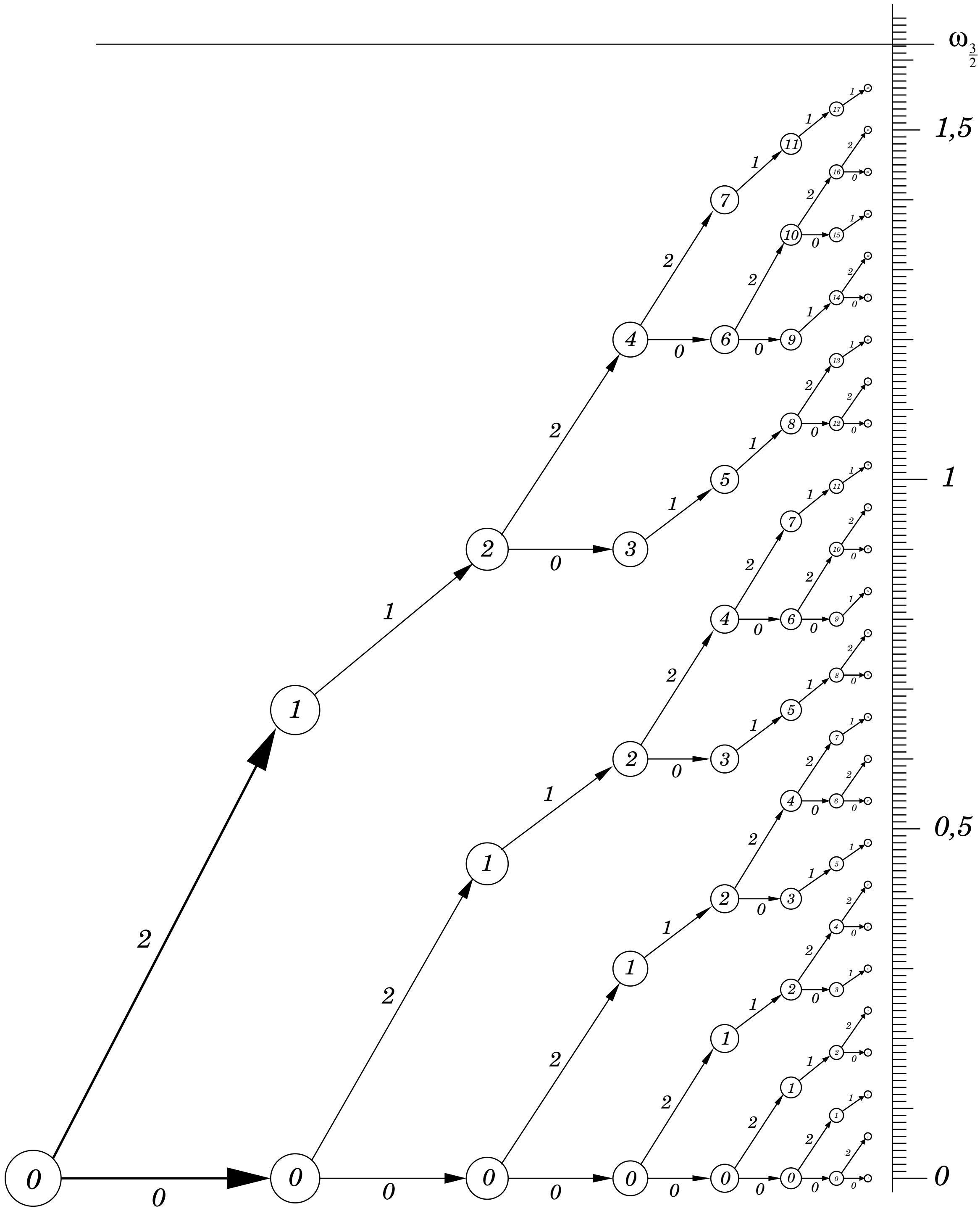

FIGURE 3.2 – L'arbre $T_{\frac{3}{2}}$.

3.4 Le problème de Josephus

L'histoire de Josephus est la suivante.

Josephus faisait partie d'un groupe de rebelles juifs durant la guerre qui les opposait aux romains. Lorsque ce groupe fût sur le point d'être capturé, les membres du groupe décidèrent de se donner la mort tour à tour plutôt que d'être faits prisonniers par leurs ennemis. Ils avaient décidé de se placer en cercle et d'exécuter un membre sur deux. Josephus, n'approuvant pas cette décision, se devait de trouver la place à occuper pour être le dernier survivant.

Par exemple, si le groupe comporte dix personnes, numérotées $1, 2, \ldots, 10$, placées en cercle et si la première personne exécutée est celle qui porte le numéro 2, alors la liste des exécutions est

$$2, 4, 6, 8, 10, 3, 7, 1, 9$$

et la position que doit occuper Josephus est donc la numéro 5. Nous notons $J(10) = 5$.

Nous pouvons généraliser ce problème en considérant, pour un entier q, qu'une personne sur q est exécutée. Si le groupe est composé de n personnes, la position à occuper par Josephus est notée $J_q(n)$.

Une approche de ce problème est décrite dans [1]. L'auteur donne une procédure pour trouver $J_q(n)$:

1. Définir une suite $\left(D_n^{(q)}\right)_{n\in\mathbb{N}}$ par
$$D_n^{(q)} = \left\lceil \frac{q}{q-1} D_{n-1}^{(q)} \right\rceil \qquad (n \geq 1; D_0^{(q)} = 1);$$
2. Déterminer le plus petit entier k tel que $D_k^{(q)} > (q-1)n$;
3. Nous obtenons $J_q(n) = qn + 1 - D_k^{(q)}$.

Inspirés par ce problème, M. Odlyzko et S. Wilf se sont intéressés [13], pour un réel α strictement supérieur à 1, aux itérations de la fonction $f(x) = \lceil \alpha x \rceil$: $f_0 = 1$ et $f_{n+1} = \lceil \alpha f_n \rceil$ pour tout entier positif n. Ils ont démontré que, pour les cas $\alpha \geq 2$ et $\alpha = 2 - \frac{1}{q}$, avec q un entier supérieur à 2, il existe une constante $H(\alpha)$ telle que $f_n = \lfloor H(\alpha)\alpha^n \rfloor$ pour tout n. En calculant cette constante pour des valeurs de α, ils se sont apperçu que la constante $H(\alpha)$, vue comme une fonction de α, se comporte de manière irrégulière. Ils ont alors demandé s'il existait une manière "indépendante" de calculer cette constante, i.e. si cette constante satisfaisait par exemple une équation fonctionnelle.

Dans cette section, nous démontrons un résultat analogue pour tout rationnel $\alpha = \frac{p}{q}$, avec $p \geq 2q-1$, et nous obtenons comme constante $H\left(\frac{p}{q}\right) = \frac{p}{q}\gamma_{\frac{p}{q}} = \omega_{\frac{p}{q}}$. Cette méthode ne fournit pas une manière indépendante pour calculer $H(\alpha)$ comme demandé, mais le développement de $\omega_{\frac{p}{q}}$ fournit au moins un algorithme facile.

13. Voir [4].

Proposition 3.4.1. *Pour tout entier positif k, il existe un entier e_k, $0 \leq e_k < \frac{q-1}{p-q}$, tel que*
$$G_k = \left\lfloor \gamma_{\frac{p}{q}} \left(\frac{p}{q}\right)^{k+1} \right\rfloor - e_k.$$

Démonstration. Par le lemme 3.3.3, nous avons
$$\begin{aligned} \left\lfloor \gamma_{\frac{p}{q}} \left(\frac{p}{q}\right)^{k+1} \right\rfloor &= \pi(qg_1g_2\cdots g_k.) + \lfloor \pi(.g_{k+1}g_{k+2}\cdots) \rfloor \\ &= G_k + \lfloor \pi(.g_{k+1}g_{k+2}\cdots) \rfloor \end{aligned}$$
Or, pour tout entier positif i, le chiffre g_i appartient à $\{0, 1, \ldots, q-1\}$ et puisque $\pi(.(p-1)^\omega) = \frac{q-1}{p-q}$, l'entier
$$e_k = \lfloor \pi(.g_{k+1}g_{k+2}\cdots) \rfloor$$
convient.

□

Le corollaire suivant est une conséquence directe de la proposition 3.4.1.

Corollaire 3.4.2. *Si $p \geq 2q - 1$, alors pour tout entier positif k, nous avons*
$$G_k = \left\lfloor \gamma_{\frac{p}{q}} \left(\frac{p}{q}\right)^{k+1} \right\rfloor.$$

Toujours dans le cas où $p \geq 2q - 1$, pour tout entier k supérieur à 1, le chiffre g_k du $\frac{p}{q}$-développement de $\gamma_{\frac{p}{q}}$ est obtenu comme suit :

1. calculer $G_{k+1} = \left\lceil \frac{p}{q} G_k \right\rceil$,
2. $g_{k+1} = qG_{k+1} \pmod p$

En effet, par définition, nous avons
$$g_k = \mathsf{md}(G_{k-1}),$$
i.e. g_k est le plus petit chiffre qu'il est possible de lire dans l'arbre $T_{\frac{p}{q}}$ à partir d'un noeud d'étiquette G_{k-1}. Par le corollaire 3.1.3, nous avons alors
$$\tau_{g_k}(G_{k-1}) = G_k,$$
i.e.
$$\frac{1}{q}\left(pG_{k-1} + g_k\right) = G_k.$$
Nous obtenons donc
$$g_k = qG_k - pG_{k-1}$$
et puisque g_k appartient à l'alphabet $A = \{0, 1, \ldots, p-1\}$, cela revient à dire
$$g_k = qG_k \pmod p.$$

Chapitre 4

Représentation des réels

Tout mot infini $\boldsymbol{a}$ dans $A^{\mathbb{N}}$ est associé à un réel x par la fonction évaluation π :

$$x = \pi(.\boldsymbol{a});$$

le mot $\boldsymbol{a}$ est donc une *représentation en base* $\frac{p}{q}$ de x. Le but premier de ce chapitre est d'associer à tout nombre réel une représentation en base $\frac{p}{q}$ *aussi canonique que possible.* Contrairement à ce qui se fait dans les systèmes de β-numération (dans lesquels la représentation est calculée algorithmiquement pour tout réel) nous posons *a priori* ce que sont les $\frac{p}{q}$-développements canoniques. Nous devons donc d'abord prouver que ceux-ci représentent en effet les réels et ensuite, montrer en quoi ils sont canoniques.

Dans la seconde partie, nous présentons un algorithme qui calcule une représentation en base $\frac{p}{q}$ que nous appellons la *représentation compagnon.* Cette représentation est écrite sur un alphabet plus large que A contenant des chiffres négatifs. Nous montrons alors comment obtenir un $\frac{p}{q}$-développement à partir de la représentation compagnon et c'est de la relation entre ces deux repésentations que nous obtenons de nouveaux résultats sur les puissances des nombres rationnels.

4.1 Les $\frac{p}{q}$-développements des nombres réels

Comme anoncé, nous définissons *a priori*, et non algorithmiquement, l'ensemble des développements dans le système de numération en base $\frac{p}{q}$ par $W_{\frac{p}{q}}$. En d'autres termes, un mot infini $\boldsymbol{a}$ de $W_{\frac{p}{q}}$ est le $\frac{p}{q}$-développement d'un réel $x = \pi(.\boldsymbol{a})$ et tout mot de $A^{\mathbb{N}}$ qui n'appartient pas à $W_{\frac{p}{q}}$ n'est pas un $\frac{p}{q}$-développement. Nous démontrons dans la suite que ce choix est cohérent : tout réel admet au moins un $\frac{p}{q}$-développement et les réels admettant plus d'un développement n'en possèdent qu'un nombre fini.

Lemme 4.1.1. *L'application* $\pi : W_{\frac{p}{q}} \to \mathbb{R}$ *préserve l'ordre.*

Remarquons que ce résultat est moins fort que le corollaire 2.2.8. En effet, si un mot infini $\boldsymbol{a}$ est strictement inférieur à un mot infini $\boldsymbol{b}$ pour l'ordre lexicographique, alors la valeur numérque de $\boldsymbol{a}$ ne peut être strictement supérieure à celle de $\boldsymbol{b}$, mais elle peut lui être égale [1].

Démonstration. Soient $\boldsymbol{a}$ et $\boldsymbol{b}$ des mots infinis dans $W_{\frac{p}{q}}$ tels que $\boldsymbol{a} \sqsubseteq \boldsymbol{b}$. Pour tout entier positif k, nous avons $a_1a_2\cdots a_k \sqsubseteq b_1b_2\cdots b_k$ et vu le corollaire 2.2.8, il vient $\pi(a_1a_2\cdots a_k) \leq \pi(b_1b_2\cdots b_k)$. En passant à la limite, nous obtenons $\pi(.\boldsymbol{a}) \leq \pi(.\boldsymbol{b})$. □

Par contraste, les exemples cités après le corollaire 2.2.8 montrent que l'application $\pi : A^{\mathbb{N}} \to \mathbb{R}$ ne préserve pas l'ordre.

Posons $X_{\frac{p}{q}} = \pi(.W_{\frac{p}{q}})$. Vu le lemme précédent, les éléments de $X_{\frac{p}{q}}$ sont des réels positifs inférieurs à $\omega_{\frac{p}{q}}$, i.e. $X_{\frac{p}{q}} \subseteq [0, \omega_{\frac{p}{q}}]$, avec $\omega_{\frac{p}{q}} < \frac{p-1}{p-q}$.

Théorème 4.1.2. *Tout nombre réel dans* $[0, \omega_{\frac{p}{q}}]$ *admet au moins un* $\frac{p}{q}$*-développement, i.e.* $X_{\frac{p}{q}} = [0, \omega_{\frac{p}{q}}]$.

Démonstration. Par définition, l'ensemble $W_{\frac{p}{q}}$ est l'ensemble des mots infinis sur A dont les préfixes finis appartiennent à $0^*L_{\frac{p}{q}}$. Puisque le langage $L_{\frac{p}{q}}$ et préfixiel et contient le mot vide, le langage $0^*L_{\frac{p}{q}}$ est préfixiel et l'ensemble $W_{\frac{p}{q}}$ est par conséquent fermé [2] dans le compact $A^{\mathbb{N}}$, donc compact. Puisque l'application π est définie sur un espace métrique compact, à valeurs dans un espace séparé et est continue, l'ensemble $X_{\frac{p}{q}}$ est fermé dans $[0, \omega_{\frac{p}{q}}]$.

Supposons qu'il existe un réel u dans $[0, \omega_{\frac{p}{q}}] \backslash X_{\frac{p}{q}}$ et posons [3]

$$y = \sup\{x \in X_{\frac{p}{q}} \mid x < u\} \quad \text{et} \quad z = \inf\{x \in X_{\frac{p}{q}} \mid x > u\}.$$

Puisque $X_{\frac{p}{q}}$ est fermé, les réels y et z lui appartiennent et par définition de $X_{\frac{p}{q}}$, il existe des mots infinis dans $W_{\frac{p}{q}}$ dont y et z sont les valeurs numériques respectives, i.e. les réels y et z possèdent un $\frac{p}{q}$-développement. Soient $\boldsymbol{a} = a_1a_2\cdots$ le plus grand $\frac{p}{q}$-développement [4] de y et $\boldsymbol{b}$ le plus petit $\frac{p}{q}$-développement[4] de z. Nous avons bien entendu $\boldsymbol{a} \sqsubset \boldsymbol{b}$. En effet, par construction nous avons $y \in X_{\frac{p}{q}}$ et $y \leq u$. Or, comme u n'appartient pas à $X_{\frac{p}{q}}$, nous avons $y < u$. De même, nous avons $z > u$.

1. Vu le lemme 3.3.4.
2. Cela découle de la proposition 7 dans l'annexe.
3. L'existence de y et de z est assurée par le fait que les réels 0 et $\omega_{\frac{p}{q}}$ appartiennent à l'ensemble $X_{\frac{p}{q}}$.
4. La représentation des réels n'étant pas nécessairement unique, nous classons les représentations dans l'ordre lexicographiqe.

Soit $a_1a_2\cdots a_N$ le plus grand préfixe commun[5] de $\boldsymbol{a}$ et de $\boldsymbol{b}$. Si nous notons $m = \pi(a_1a_2\cdots a_N.)$, $n = \pi(a_1a_2\cdots a_N a_{N+1}.)$ et $r = \pi(a_1a_2\cdots a_N b_{N+1}.)$, nous obtenons

$$\boldsymbol{a} \sqsubseteq a_1a_2\cdots a_N a_{N+1}\mathsf{W}(n) \sqsubset a_1a_2\cdots a_N b_{N+1}\mathsf{w}(r) \sqsubseteq \boldsymbol{b}.$$

Vu le lemme 4.1.1, nous avons $\pi(.a_1a_2\cdots a_N a_{N+1}\mathsf{W}(n)) < z$. Or, nous savons que le mot infini $a_1a_2\cdots a_{N+1}\mathsf{W}(n)$ appartient à $W_{\frac{p}{q}}$ et sa valeur numérique appartient donc à $X_{\frac{p}{q}}$. Puisque celle-ci est strictement inférieure à z, elle est également strictement inférieure à u et vu le choix de $\boldsymbol{a}$, nous avons alors $\boldsymbol{a} = a_1a_2\cdots a_N a_{N+1}\mathsf{W}(n)$. Nous montrons de la même manière que $\boldsymbol{b} = a_1a_2\cdots a_N b_{N+1}\mathsf{w}(r)$.

Si $a_{N+1}+q < b_{N+1}$, alors il existe un chiffre c dans $\mathsf{d}(m)$ tel que $a_{N+1}+q \leq c < b_{N+1}$. Dans ce cas, pour tout mot infini $\boldsymbol{c'}$ dans $A^{\mathbb{N}}$ tel que le mot infini $\boldsymbol{c} = a_1a_2\cdots a_N c\boldsymbol{c'}$ est dans $W_{\frac{p}{q}}$, nous avons

$$\boldsymbol{a} \sqsubset \boldsymbol{c} \sqsubset \boldsymbol{b}.$$

Ainsi, que la valeur numérique de $\boldsymbol{c}$ soit y ou z, nous avons une contradiction avec le choix de $\boldsymbol{a}$ et de $\boldsymbol{b}$.

Si maintenant $a_{N+1} + q = b_{N+1}$, alors nous avons $r = n + 1$ et par le lemme 3.3.4, $y = z$, d'où une contradiction.

□

Un mot de $W_{\frac{p}{q}}$ est dit *ultimement maximal* (resp. *ultimement minimal*) s'il a un suffixe qui est un mot maximal (resp. un mot minimal). La proposition suivante découle alors directement du lemme 3.3.4.

Proposition 4.1.3. *Si le mot infini* $\boldsymbol{a}$ *de* $W_{\frac{p}{q}}$ *est ultimement maximal, alors le réel* $x = \pi(.\boldsymbol{a})$ *admet un autre* $\frac{p}{q}$*-développement qui est un mot infini ultimement minimal et réciproquement.*

Théorème 4.1.4. *L'ensemble des nombres réels dans* $X_{\frac{p}{q}}$ *qui admettent plus d'un* $\frac{p}{q}$*-développement est infini dénombrable. Les* $\frac{p}{q}$*-développements de ces réels sont des mots infinis ultimement maximaux ou ultimement minimaux.*

Nous avons vu, dans le lemme 3.3.4, qu'à chaque embranchement de l'arbre $T_{\frac{p}{q}}$, il correspond un réel admettant au moins deux $\frac{p}{q}$-développements. Il suffit donc de démontrer la proposition suivante pour établir le théorème 4.1.4.

Proposition 4.1.5. *Soit* x *un réel dans* $[0, \omega_{\frac{p}{q}}]$ *admettant plus d'un* $\frac{p}{q}$*-développement. Alors* x *admet au plus* $k+1$ $\frac{p}{q}$*-développements, avec* $k \in \mathbb{Z}$, $k > \frac{q-1}{p-q}$*, et peut être associé à au plus* k *embranchements de l'arbre* $T_{\frac{p}{q}}$*. Le plus petit de ces développements est un mot infini ultimement maximal et le plus grand est un mot infini ultimement minimal. Les autres, s'il en existe, sont des mots infinis ultimement maximaux et ultimement minimaux.*

5. Avec la convention que N peut être égal à 0.

Démonstration. Posons $R = \pi^{-1}(x) \cap W_{\frac{p}{q}}$ l'ensemble (fermé) des $\frac{p}{q}$-développements de x. Nous notons $\boldsymbol{a} = a_1 a_2 \cdots$ le plus petit $\frac{p}{q}$-développement de x, $\boldsymbol{b} = b_1 b_2 \cdots$ son plus grand et $a_1 a_2 \cdots a_N$ le plus long préfixe commun de $\boldsymbol{a}$ et de $\boldsymbol{b}$ (où N peut être égale à 0).

Soit $\boldsymbol{c}$ un mot infini dans R différent (et donc plus petit que) $\boldsymbol{b}$. Nous commençons par montrer qu'il n'existe pas d'entier h tel que $\pi(\boldsymbol{b}_{[h]}.) - \pi(\boldsymbol{c}_{[h]}.) \geq k + 1$. Procédons par l'absurde et notons $\pi(\boldsymbol{c}_{[h]}.) = n - 1$, $\pi(\boldsymbol{b}_{[h]}.) = m \geq n + k$ et d le mot de $0^* L_{\frac{p}{q}}$ de longueur h tel que $\pi(d) = n$. Il vient

$$\boldsymbol{c} \sqsubset d\mathsf{w}(n) \sqsubset \boldsymbol{b}_{[h]}\mathsf{w}(m) \sqsubseteq \boldsymbol{b},$$

et donc, par les lemmes 3.3.5 et 4.1.1,

$$\pi(.\boldsymbol{c}) \leq \pi(.d\mathsf{w}(n)) < \pi(.\boldsymbol{b}_{[h]}\mathsf{w}(m)) \leq \pi(.\boldsymbol{b}),$$

ce qui contredit l'égalité $\pi(.\boldsymbol{c}) = \pi(.\boldsymbol{b}) = x$. Cela implique directement que l'ensemble R, qui consiste en l'ensemble des mots infinis $\boldsymbol{c}$ de $W_{\frac{p}{q}}$ tels que $\boldsymbol{a} \sqsubseteq \boldsymbol{c} \sqsubseteq \boldsymbol{b}$, contient au plus $k + 1$ éléments et que le sous-arbre de $T_{\frac{p}{q}}$ correspondant a au plus k embranchements.

Il est clair que le mot infini $\boldsymbol{a}$ est utlimemement maximal et que le mot infini $\boldsymbol{b}$ est ultimement minimal. Il reste donc à montrer que les autres $\frac{p}{q}$-développements de x sont des mots infinis ultimement maximaux et ultimement minimaux. Intuitivement, puisque le sous-arbre correspondant à R contient au plus k embranchements, les $\frac{p}{q}$-développements finiront par ne plus avoir le choix du chemin à emprunter : à partir d'un noeud, ils ne pourront emprunter qu'une branche menant vers un noeud et l'étiquette de cette branche sera le chiffre le plus grand et le plus petit possible, puisqu'il sera le seul. Plus formellement, supposons que pour un entier h supérieur à N, le chiffre c_{h+1} n'est pas le plus grand, i.e. c_{h+1} est strictement inférieur à $\mathsf{Md}(\pi(\boldsymbol{c}_{[h]}))$. Posons $\boldsymbol{c'}_h = \mathsf{W}(\pi(\boldsymbol{c}_{[h]}))$, il vient

$$\boldsymbol{c} \sqsubset \boldsymbol{c}_{[h]}\boldsymbol{c'}_h \sqsubseteq \boldsymbol{b}.$$

Puisqu'il n'y a que k embranchements, la suite des entiers h tels que c_{h+1} n'est pas maximal est finie et le mot infini $\boldsymbol{c}$ est donc ultimement maximal. Nous montrons de la même façon que tout $\frac{p}{q}$-développement de x différent de $\boldsymbol{a}$ est ultimement minimal. □

En particulier, si $p \geq 2q$, alors il n'existe pas de nombre réel admettant plus de deux $\frac{p}{q}$-développements. En effet, il suffit de prendre[6] $k = \left\lceil \frac{q-1}{p-q} \right\rceil = 1$ car $\frac{q-1}{p-q} \leq \frac{q-1}{q}$.

Un simple argument de combinatoire permet alors d'élargir la condition afin de satisfaire au cas $\frac{p}{q} = \frac{3}{2}$.

6. Ce choix pourrait poser problème si $\frac{q-1}{p-q}$ était un nombre entier car dans ce cas, $k = \left\lceil \frac{q-1}{p-q} \right\rceil = \frac{q-1}{p-q}$ et k n'est donc pas strictement supérieur à $\frac{q-1}{p-q}$. Cependant, pour le cas $p \geq 2q$, la fraction $\frac{q-1}{p-q}$ n'est jamais un nombre entier.

Corollaire 4.1.6. *Si $p \geq 2q-1$, alors il n'existe pas de nombre réel admettant plus de deux $\frac{p}{q}$-développements.*

Démonstration. Il suffit de démontrer le résultat pour $p = 2q - 1$. Si le nombre réel x admet plus de deux $\frac{p}{q}$-développements, alors, au moins l'un d'entre eux est à la fois ultimement maximal et ultimement minimal. Par la proposition 3.2.1, le suffixe maximal et minimal de celui-ci s'écrit donc sur l'alphabet unaire [7]

$$\{0, 1, \ldots, q-1\} \cap \{p-q, p-q+1, \ldots, p-1\} = \{q-1\}.$$

Ce mot est donc ultimement périodique, d'où une contradiction car il appartient à $W_{\frac{p}{q}}$. $\square$

Remarquons que, vu les conventions prises et contrairement aux systèmes de numération standards, les préfixes finis des $\frac{p}{q}$-développements d'un nombre réel, auxquels nous rajoutons le suffixe 0^ω, ne sont pas des $\frac{p}{q}$-développements (bien que l'application π leur associe une valeur numérique réelle). En d'autres termes, si un mot non vide w appartient au langage $L_{\frac{p}{q}}$, alors le mot $w0^\omega$ n'appartient pas à $W_{\frac{p}{q}}$.

7. Car $p - q = q - 1$.

4.2 La représentation compagnon en base $\frac{p}{q}$ et le co-convertisseur

Une des particularités de la représentation des entiers en base $\frac{p}{q}$ est que les $\frac{p}{q}$-développements sont calculés avec "le chiffre le moins significatif en premier", i.e. de droite à gauche. Cette méthode est raisonnable tant qu'il ne s'agit que d'entiers, mais elle commence à poser problème pour les réels, lorsque la représentation est un mot infini à droite. Cette difficulté peut être surmontée en adoptant la définition d'une autre représentation en base $\frac{p}{q}$ des réels qui peut être calculée avec n'importe quelle précision (en supposant que nous pouvons calculer dans $\mathbb{Q}$ avec cette même précision) et de gauche à droite. Le prix à payer pour cela est de devoir utiliser un alphabet plus large contenant des chiffres négatifs.

Soit $\psi : \mathbb{R}^+ \to \mathbb{Z}$ la fonction définie par

$$\psi(x) = q \left\lfloor \left(\frac{p}{q}\right) x \right\rfloor - p \lfloor x \rfloor .$$

Lemme 4.2.1. *La fonction ψ est périodique de période q et pour tout réel positif x, $\psi(x)$ appartient à l'alphabet*

$$C = \{-(q-1), \dots, 0, 1, \dots, p-1\}.$$

Démonstration. La fonction ψ est clairement périodique de période q : pour tout entier positif i,

$$\begin{aligned} \psi(x+iq) &= q \left\lfloor \left(\frac{p}{q}\right)(x+iq) \right\rfloor - p \lfloor x+iq \rfloor \\ &= q \left(\left\lfloor \left(\frac{p}{q}\right) x \right\rfloor + ip \right) - p \lfloor x \rfloor - ipq \\ &= \psi(x) \end{aligned}$$

De plus, nous avons

$$\left(\frac{p}{q}\right) x - \frac{p}{q} < \left(\frac{p}{q}\right) \lfloor x \rfloor \leq \left(\frac{p}{q}\right) x < \left\lfloor \left(\frac{p}{q}\right) x \right\rfloor + 1.$$

En multipliant chaque membre de ces inégalités par q, nous obtenons

$$\begin{aligned} & q \left(\frac{p}{q}\right) \lfloor x \rfloor < q \left\lfloor \left(\frac{p}{q}\right) x \right\rfloor + q \\ \Leftrightarrow \quad & \psi(x) > -q \end{aligned}$$

et

$$\begin{aligned}
& \left(\frac{p}{q}\right) x - \frac{p}{q} < \left(\frac{p}{q}\right) \lfloor x \rfloor \\
\Rightarrow \quad & \left\lfloor \left(\frac{p}{q}\right) x \right\rfloor - \frac{p}{q} < \left(\frac{p}{q}\right) \lfloor x \rfloor \\
\Leftrightarrow \quad & q \left\lfloor \left(\frac{p}{q}\right) x \right\rfloor - p < p \lfloor x \rfloor \\
\Leftrightarrow \quad & \psi(x) < p
\end{aligned}$$

□

Les graphiques suivants représentent la fonction ψ pour les cas $\frac{3}{2}$ et $\frac{7}{3}$.

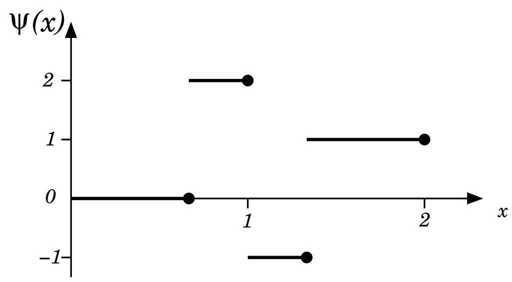

FIGURE 4.1 – La fonction ψ pour le cas $\frac{3}{2}$.

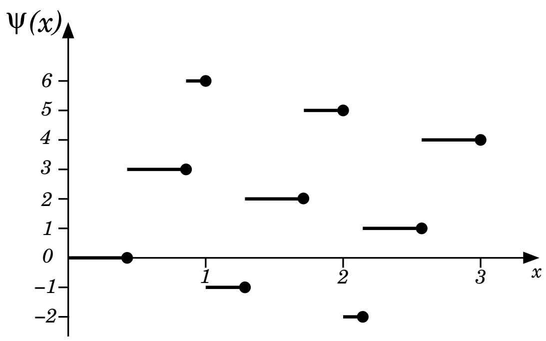

FIGURE 4.2 – La fonction ψ pour le cas $\frac{7}{3}$.

Nous définissons maintenant la *représentation compagnon* en base $\frac{p}{q}$ d'un réel positif x par la suite $\varphi(x)$ définie par

$$\varphi(x) = \boldsymbol{c} = c_1 c_2 \cdots \qquad \text{avec } c_n = \psi\left(\left(\frac{p}{q}\right)^{n-1} x\right) \ \forall n \geq 1.$$

Dans le cas $q = 1$, le chiffre c_n est exactement le n-ième chiffre après le séparateur dans le[8] p-développement de x. En effet, l'application ψ dans le cas des bases entières devient

$$\psi(x) = \lfloor px \rfloor - p \lfloor x \rfloor .$$

Considérons l'exemple[9] de la base 10 et choisissons un réel $x = 0,1658916\cdots$. Nous obtenons successivement

$$\begin{aligned}
c_1 = &\quad \psi(x) = \lfloor 10x \rfloor - 10 \lfloor x \rfloor &= 1 \\
c_2 = &\quad \psi(10x) = \lfloor 100x \rfloor - 10 \lfloor 10x \rfloor &= 6 \\
c_3 = &\quad \psi(100x) = \lfloor 1000x \rfloor - 10 \lfloor 100x \rfloor &= 5 \\
&\quad \vdots
\end{aligned}$$

L'application π devient alors

$$\begin{aligned}
\pi(c_1 c_2 \cdots c_n.) &= \frac{1}{q} \sum_{i=0}^{n-1} c_{n-i} \left(\frac{p}{q}\right)^i \\
&= \frac{1}{q} \sum_{i=0}^{n-1} \psi\left(\left(\frac{p}{q}\right)^{n-i-1} x\right) \left(\frac{p}{q}\right)^i \\
&= \frac{1}{q} \sum_{i=0}^{n-1} \left(q \left\lfloor \left(\frac{p}{q}\right)^{n-i} x \right\rfloor - p \left\lfloor \left(\frac{p}{q}\right)^{n-i-1} x \right\rfloor\right) \left(\frac{p}{q}\right)^i \\
&= \sum_{i=0}^{n-1} \left(\left\lfloor \left(\frac{p}{q}\right)^{n-i} x \right\rfloor \left(\frac{p}{q}\right)^i - \left\lfloor \left(\frac{p}{q}\right)^{n-i-1} x \right\rfloor \left(\frac{p}{q}\right)^{i+1}\right) \\
&= \left\lfloor \left(\frac{p}{q}\right)^n x \right\rfloor - \left(\frac{p}{q}\right)^n \lfloor x \rfloor .
\end{aligned}$$

Proposition 4.2.2. *Pour tout nombre réel positif x, $\varphi(x)$ est une représentation en base $\frac{p}{q}$ de*[10] $\{x\}$.

Démonstration. De la section 3.3, il vient

$$\pi(.\varphi(x)) = \lim_{n\to+\infty} \left(\frac{q}{p}\right)^n \pi(c_1 c_2 \cdots c_n.) = \lim_{n\to+\infty} \left(\frac{q}{p}\right)^n \left\lfloor \left(\frac{p}{q}\right)^n x \right\rfloor - \lfloor x \rfloor = x - \lfloor x \rfloor ,$$

où la dernière égalité vient du fait que la limite de $\left(\frac{q}{p}\right)^n \left\lfloor \left(\frac{p}{q}\right)^n x \right\rfloor$ lorsque n tend vers l'infini est x.

$\square$

8. Le p-développement dont il est question ici est celui fournit par l'algorithme glouton.
9. N'importe quelle base entière se traite de la même façon.
10. Rappellons que $\{x\}$ désigne la partie fractionnaire de x, i.e. $\{x\} = x - [x]$.

Soient x un réel dans $[0, \omega_{\frac{p}{q}}]$, $\boldsymbol{a} = a_1 a_2 \cdots$ son $\frac{p}{q}$-développement (ou un de ses $\frac{p}{q}$-développements s'il en a plusieurs) et $\varphi(x) = \boldsymbol{c} = c_1 c_2 \cdots$ sa représentation compagnon. Comme au lemme 3.3.3, nous notons $\rho_n(x) = \lfloor \pi(.a_{n+1}a_{n+2}\cdots) \rfloor$ et nous avons $0 \leq \rho_n(x) < \frac{p-1}{p-q}$. Vu la définition de c_n,

$$c_n = \psi\left(\left(\frac{p}{q}\right)^{n-1} x\right) = q\left\lfloor\left(\frac{p}{q}\right)^{n} x\right\rfloor - p\left\lfloor\left(\frac{p}{q}\right)^{n-1} x\right\rfloor,$$

il découle du lemme[11] 3.3.3 et de l'équation $q\pi(a_1a_2\cdots a_n) = p\pi(a_1a_2\cdots a_{n-1}) + a_n$ que

$$c_n + p\rho_{n-1}(x) = a_n + q\rho_n(x). \tag{4.1}$$

Nous introduisons maintenant un transducteur lettre-à-lettre qui convertit la représentation compagnon d'un réel x en un de ses $\frac{p}{q}$-développements.

Soit $\mathcal{A}_{\frac{p}{q}} = (H, C \times A, F, H, H)$ le transducteur lettre-à-lettre (à gauche) dont l'ensemble d'états est $H = \{h \in \mathbb{N} \mid 0 \leq h < \frac{p-1}{p-q}\}$ et dont l'ensemble des transitions F est défini par

$$(h, (c, a), h') \in F \Leftrightarrow ph + c = qh' + a.$$

Remarquons que ce transducteur est en fait le transposé du transducteur à droite $\mathcal{C}_C$ qui réalise la conversion[12] χ_C (une fois que les étiquettes des états ont été remplacées par leurs opposés respectifs). En effet, l'ensemble E des transitions de ce transducteur est défini par

$$(z, (d, a), z') \in E \Leftrightarrow qz + d = pz' + a.$$

Ainsi, si nous remplaçons z et z' respectivement par $-z$ et $-z'$, nous obtenons

$$\begin{aligned}(-z, (d, a), -z') \in E &\Leftrightarrow -qz + d = -pz' + a \\ &\Leftrightarrow pz' + d = qz + a \\ &\Leftrightarrow (z', (d, a), z) \in F\end{aligned}$$

De plus, l'état minimum accessible dans $\mathcal{C}_C$ est

$$z_{\min} = \left\lceil \frac{f+1}{p-q} \right\rceil - 1 = \left\lceil \frac{2-q}{p-q} \right\rceil - 1,$$

car le plus petit chiffre dans C est $f = -(q-1)$. Ainsi, l'état maximal accessible dans $\mathcal{A}_{\frac{p}{q}}$ est $-z_{\min}$; il faut donc montrer que $z_{\min}$ est le plus petit entier strictement supérieur à $\frac{-(p-1)}{p-q}$. Supposons l'existence d'un entier a vérifiant

$$\frac{-(p-1)}{p-q} < a < \left\lceil \frac{2-q}{p-q} \right\rceil - 1;$$

11. En prenant $k = n$ et $k = n - 1$.
12. cf. la section 2.3.

nous obtenons

$$a \leq \frac{2-q}{p-q} - 1 = \frac{2-p}{p-q}$$

et donc

$$\frac{1-p}{p-q} < a \leq \frac{2-p}{p-q},$$

d'où une contradiction car $p > q \geq 1$.

Le théorème suivant découle directement de la définition des transitions de $\mathcal{A}_{\frac{p}{q}}$ et de l'équation 4.1.

Théorème 4.2.3. *Soient x un nombre réel dans $[0, \omega_{\frac{p}{q}}]$, $\boldsymbol{c}$ sa représentation compagnon et $\boldsymbol{a}$ un $\frac{p}{q}$-développement de x. Alors le couple $(\boldsymbol{c}, \boldsymbol{a})$ est l'étiquette d'un chemin infini dans $\mathcal{A}_{\frac{p}{q}}$ qui débute dans l'état $\rho_0(x) = \lfloor x \rfloor$.*

Ecrivons l'alphabet $C = \{-(q-1), \ldots, 0, 1, \ldots, p-1\}$ comme une union disjointe $C = C_1 \cup C_2 \cup C_3$, avec

$$\begin{aligned} C_1 &= \{-(q-1), \ldots, -1\}; \\ C_2 &= \{0, \ldots, q-1\}; \\ C_3 &= \{q, \ldots, p-1\}. \end{aligned}$$

Si nous considérons le cas $p \geq 2q-1$, alors le transducteur $\mathcal{A}_{\frac{p}{q}}$ a uniquement deux états : 0 et 1. En effet, nous avons dans ce cas

$$1 < \frac{p-1}{p-q} \leq \frac{2(q-1)}{q-1} = 2.$$

De plus, vu la manière dont sont définies les transitions, les chiffres appartenant respectivement aux trois alphabets C_1, C_2 et C_3 impliquent des comportements différents du transducteur. Ainsi, il est impossible, partant de l'état 0 (resp. de 1), de lire un couple dont la première composante est un chiffre appartenant à l'alphabet C_1 (resp. à C_3). De même, toute transition dont l'étiquette est un couple dont la première composante appartient à l'alphabet C_2 aboutit dans son état d'origine. Par contre, si à partir de l'état 0 (resp. de 1), nous lisons un couple dont la première composante est dans C_3 (resp. dans C_1), alors nous pouvons aboutir dans chacun des deux états ; ce transducteur n'est donc pas séquentiel.

Les images suivantes représentent respectivement le transducteur $\mathcal{A}_{\frac{p}{q}}$ pour $p \geq 2q-1$ et le transducteur $\mathcal{A}_{\frac{3}{2}}$.

Remarquons que le théorème 4.2.3 assure l'existence d'un chemin d'étiquette $(\boldsymbol{c}, \boldsymbol{a})$ dans le transducteur $\mathcal{A}_{\frac{p}{q}}$, mais ne suffit pas pour calculer un $\frac{p}{q}$-développement d'un

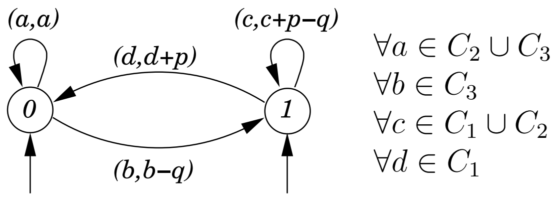

FIGURE 4.3 – Le transducteur $A_{\frac{p}{q}}$ pour $p \geq 2q - 1$.

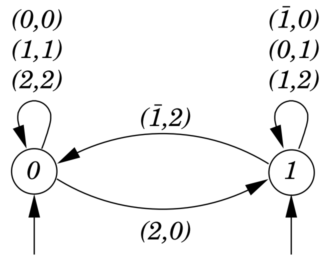

FIGURE 4.4 – Le transducteur $A_{\frac{3}{2}}$.

réel à partir de sa représentation compagnon. En effet, le transducteur $\mathcal{A}_{\frac{p}{q}}$ n'étant en général pas séquentiel, il est impossible de construire un algorithme sur base de celui-ci qui fournirait un $\frac{p}{q}$-développement à partir de la représentation compagnon. Pour ce faire, nous allons alors utiliser le convertisseur $\mathcal{C}_C$ qui est le transposé de $\mathcal{A}_{\frac{p}{q}}$ et qui est séquentiel.

Le calcul de la représentation compagnon des nombres réels est la première étape de l'"algorithme" pour le calcul de leurs $\frac{p}{q}$-développements. Soit x un nombre réel dans $[0, \omega_{\frac{p}{q}}]$ et soit $\boldsymbol{c}$ sa représentation compagnon. Soient aussi n un entier positif fixé (supposé grand) et $v = c_1 c_2 \cdots c_n$ le préfixe de longueur n de $\boldsymbol{c}$. Lorsque v est lu de droite à gauche par le transducteur $\mathcal{C}_C$ à partir d'un état initial s, ce qui en ressort est un mot $f^{(s)}$ sur A qui dépend de s. Le préfixe de longueur maximum qui est commun à tous les mots $f^{(s)}$ est un préfixe de tous les $\frac{p}{q}$-développements du réel x.

Pour obtenir des préfixes plus long, il faut evidemment refaire cette opération avec un préfixe de la représentation compagnon de longueur n' plus grande que n, mais il est impossible de savoir à l'avance quelle taille minimum doit faire ce préfixe pour obtenir une meilleure approximation des développements.

Nous présentons maintenant une caractérisation de la représentation compagnon des nombres réels ayant plus d'un $\frac{p}{q}$-développement.

Proposition 4.2.4. *Supposons $p \geq 2q-1$. Un nombre réel x a deux $\frac{p}{q}$-développements si et seulement si sa représentation compagnon fini par s'écrire sur l'alphabet C_2, i.e. s'il existe un suffixe de sa représentation compagnon appartenant à $C_2^{\mathbb{N}}$.*

Démonstration. Vu les hypothèses, le transducteur $\mathcal{A}_{\frac{p}{q}}$ n'a que deux états : 0 et 1 et nous avons déjà vu ci-dessus comment il se comporte.

Soient x un nombre réel et $\boldsymbol{c}$ sa représentation compagnon. Par le théorème 4.1.2, ce réel a au moins un $\frac{p}{q}$-développement $\boldsymbol{a}$ et par le théorème 4.2.3, le couple $(\boldsymbol{c}, \boldsymbol{a})$ décrit alors un chemin infini dans le transducteur $\mathcal{A}_{\frac{p}{q}}$.

Supposons que x admet deux $\frac{p}{q}$-développements. Si le mot infini $\boldsymbol{c}$ ne finit pas par s'écrire sur l'alphabet C_2, alors il existe une suite croissante d'indices $(n_i)_{i\in\mathbb{N}}$ tels que c_{n_i} appartient à $C_1 \cup C_3$ pour tout i. Or, vu le comportement du transducteur, l'état à partir duquel part une transition $(c_{n_i}, .)$ est univoquement déterminé : si le chiffre c_{n_i} appartient à l'alphabet C_1 (resp. à C_3), alors nous savons que la transition part de l'état 0 (resp. de 1) car il est impossible, à partir de l'état 1 (resp. de 0), de lire un couple dont la première composante appartient à C_1 (resp. à C_3). Ensuite, puisque le transducteur est co-séquentiel, nous pouvons lire à l'envers, i.e. de droite à gauche, le couple $(\boldsymbol{c}_{[n_i]}, \boldsymbol{a}_{[n_i]})$ des préfixes respectifs de longueur n_i des mots $\boldsymbol{c}$ et $\boldsymbol{a}$ et le réel x n'admet donc qu'un unique développement, d'où une contradiction.

Supposons maintenant que le mot infini $\boldsymbol{c}$ finit par s'écrire sur l'alphabet C_2. Il existe donc un entier positif N tel que pour tout entier n supérieur à N, le chiffre c_n appartient à l'alphabet C_2. Vu le comportement du transducteur, le chemin d'étiquette $(\boldsymbol{c}, \boldsymbol{a})$ finit donc par boucler dans le même état. Si cet état est 0 (resp. 1), alors les chiffres de $\boldsymbol{a}$ finissent par appartenir à l'ensemble $\{0, 1, \ldots, q-1\}$ (resp. $\{p-q, p-q+1, \ldots, p-1\}$) et le mot $\boldsymbol{a}$ est donc ultimement minimal (resp. ultimement maximal). Vu la proposition 4.1.3, le réel x admet un autre $\frac{p}{q}$-développement qui est un mot ultimement maximal (resp. ultimement minimal).

$\square$

Chapitre 5

À propos de la partie fractionnaire des puissances de nombres rationnels

G. H. Hardy et J. E. Littlewood ont démontré que la suite $(\{\theta^n\})_{n\in\mathbb{N}}$ est équirépartie pour presque tout réel $\theta > 1$. Des résultats expérimentaux [1] ont montré que la distribution de la suite $\left(\left\{\left(\frac{p}{q}\right)^n\right\}\right)_{n\in\mathbb{N}}$, pour p et q des nombres entiers premiers entre eux, $p > q \geq 2$, était plus chaotique que pour les puissances d'un nombre transcendant comme π ou e.

Nous nous intéressons dans ce chapitre à un problème bien connu de K. Malher : la démonstration de l'existence, ou de l'inexistence, des *Z-nombres*. Un *Z-nombre* est un réel z non nul tel que la partie fractionnaire de $z\left(\frac{3}{2}\right)^n$ appartient à l'intervalle $[0, \frac{1}{2}[$ pour tout entier $n \geq 0$. Nous ignorons si de tels nombres existent, mais K. Malher a démontré [2] que l'ensemble de ces nombres était au plus dénombrable.

Une généralisation de ce problème est la suivante : soit I un sous-interval strict de $[0,1[$, nous notons

$$\boldsymbol{Z}_{\frac{p}{q}}(I) = \left\{z > 0 \mid \exists N \in \mathbb{N}\ :\ \forall n \geq N,\ \left\{z\left(\frac{p}{q}\right)^n\right\} \in I\right\}.$$

Nous démontrons dans ce chapitre que la recherche de Z-nombre peut être transposée à la recherche de mots infinis ultimement minimaux vérifiant certaines propriétés.

Lemme 5.0.5. *Supposons $p \geq 2q-1$. Pour tout chiffre c dans $C_2 = \{0, 1, \ldots, q-1\}$, soit k_c l'unique entier dans $A = \{0, 1, \ldots, p-1\}$ tel que $qk_c = c \pmod p$. Alors $\psi(qx) = c$ si et seulement si $\{x\}$ appartient à l'intervalle $\left[\frac{1}{p}k_c, \frac{1}{p}(k_c+1)\right[$.*

Démonstration. Puisque la fonction ψ est périodique de période q, nous avons

$$\psi(qx) = \psi\left(q\left(\lfloor x\rfloor + \{x\}\right)\right) = \psi\left(q\{x\}\right)$$

1. Voir [10].
2. Voir [3].

et nous pouvons donc supposer x dans $[0,1[$.

L'unicité de l'entier k_c vient du fait que les entiers p et q sont premiers entre eux :

$$q(k-k') = 0 \pmod p \;\Rightarrow\; k = k' \pmod p \;\Rightarrow\; k = k'$$

Pour la même raison, $\psi(qx) = q\lfloor px \rfloor - p\lfloor qx \rfloor = c$ implique l'existence d'un unique couple (k_c, j_c) de $A \times C_2$ tel que $\lfloor px \rfloor = k_c$ et $\lfloor qx \rfloor = j_c$. En effet, nous avons

$$\begin{aligned}\psi(qx) = q\lfloor px \rfloor - p\lfloor qx \rfloor = c = qk_c \pmod p\\ \Leftrightarrow \lfloor px \rfloor = k_c \in A \quad \text{ et } \quad \lfloor qx \rfloor = j_c.\end{aligned}$$

Or, comme $qk_c - pj_c = c \in C_2$, nous avons $j_c \in C_2$.

Donc, $\psi(qx) = c$ si et seulement si x appartient à $\left[\frac{1}{q}j_c, \frac{1}{q}(j_c+1)\right[\cap \left[\frac{1}{p}k_c, \frac{1}{p}(k_c+1)\right[$. Par hypothèse sur c et sur p et q, il vient alors

$$0 \leq qk_c - pj_c \leq q - 1 \leq p - q.$$

En divisant les membres de ces inégalités par pq, nous obtenons enfin

$$\frac{1}{q}j_c \leq \frac{1}{p}k_c \quad \text{et} \quad \frac{1}{p}(k_c+1) \leq \frac{1}{q}(j_c+1),$$

d'où la conclusion.

□

Pour tout nombre rationnel fixé $\frac{p}{q}$, nous notons $Y_{\frac{p}{q}}$ l'ensemble

$$Y_{\frac{p}{q}} = \bigcup_{0 \leq c \leq q-1} \left[\frac{1}{p}k_c, \frac{1}{p}(k_c+1)\right[,$$

où les entiers k_c sont définis par le lemme 5.0.5.

L'ensemble $Y_{\frac{p}{q}}$ est un sous-ensemble de $[0,1[$ qui consiste en l'union de q semi-intervalles de longueur $\frac{1}{p}$. Nous avons par exemple

$$Y_{\frac{3}{2}} = \left[0, \frac{1}{3}\right[\cup \left[\frac{2}{3}, 1\right[.$$

Vu les notations introduites, le lemme 5.0.5 peut être reformulé comme suit :

$$\psi(qx) \in C_2 \textit{ si et seulement si } \{x\} \in Y_{\frac{p}{q}}.$$

La notation de Mahler généralisée est alors

$$\boldsymbol{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right) = \left\{z > 0 \mid \exists N \in \mathbb{N} \;:\; \forall n \geq N,\; \left\{z\left(\frac{p}{q}\right)^n\right\} \in Y_{\frac{p}{q}}\right\}.$$

Théorème 5.0.6. *Si $p \geq 2q-1$, alors $\boldsymbol{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right)$ est infini dénombrable.*

Ce théorème est en fait une conséquence directe du théorème suivant et du théorème 4.1.4.

Théorème 5.0.7. *Supposons $p \geq 2q-1$. Un réel z appartient à $\boldsymbol{Z}_{\frac{p}{q}}\left(Y_{\frac{p}{q}}\right)$ si et seulement s'il a deux $\frac{p}{q}$-développements.*

Démonstration. Par la proposition 4.2.4, nous savons qu'un réel z possède deux $\frac{p}{q}$-développements si et seulement s'il existe un entier strictement positif N tel que pour tout entier n strictement supérieur à N, le chiffre $c_n = \psi\left(z\left(\frac{p}{q}\right)^{n-1}\right)$ de sa représentation compagnon appartient à l'alphabet C_2 et par le lemme 5.0.5, cette condition devient

$$\left\{\left(\frac{p}{q}\right)^{n-1}\frac{z}{q}\right\} \in \bigcup_{0\leq c\leq q-1}\left[\frac{1}{p}k_c, \frac{1}{p}(k_c+1)\right[= Y_{\frac{p}{q}},$$

d'où la conclusion.

$\square$

Puisque nous pouvons considérer des rationnels arbitrairement grand qui satisfont la condition $p \geq 2q-1$, il vient

Corollaire 5.0.8. *Pour tout $\varepsilon > 0$, il existe un rationnel $\frac{p}{q}$ et un sous-ensemble $Y_{\frac{p}{q}}$ dans $[0, 1[$ de mesure de Lebesgue inférieure à ε tels que $\boldsymbol{Z}\left(Y_{\frac{p}{q}}\right)$ est infini dénombrable.*

Revenons maintenant au papier original de K. Malher. Un *Z-nombre* est un nombre réel z tel que pour tout entier positif n, la partie fractionnaire de $z\left(\frac{3}{2}\right)^n$ appartient au semi-intervalle $\left[0, \frac{1}{2}\right[$. Si nous décomposons le réel $z\left(\frac{3}{2}\right)^n$ en sa partie entière et sa partie fractionnaire, nous obtenons

$$z\left(\frac{3}{2}\right)^n = h_n + r_n.$$

Puisque r_{n+1} appartient à $\left[0, \frac{1}{2}\right[$ pour tout n, nous avons $r_n \in \left[0, \frac{1}{3}\right[$ si h_n est paire et $r_n \in \left[\frac{1}{3}, \frac{1}{2}\right[$ si h_n est impaire. Il s'ensuit que, suivant la parité de h_n, soit

$$\frac{z}{2}\left(\frac{3}{2}\right)^n = \frac{h_n}{2} + \frac{r_n}{2}, \qquad \text{avec } \frac{r_n}{2} \in \left[0, \frac{1}{6}\right[,$$

soit

$$\frac{z}{2}\left(\frac{3}{2}\right)^n = \frac{h_n-1}{2} + \frac{r_n+1}{2}, \qquad \text{avec } \frac{r_n+1}{2} \in \left[\frac{2}{3}, \frac{3}{4}\right[.$$

En d'autres termes, pour tout entier positif n, nous avons

$$\left\{\frac{z}{2}\left(\frac{3}{2}\right)^n\right\} \in \left[0, \frac{1}{6}\right[\cup \left[\frac{2}{3}, \frac{3}{4}\right[$$

et par le théorème 5.0.7, cela implique que z admet deux $\frac{3}{2}$-développements.

De plus, en supposant que z soit un Z-nombre, K. MALHER calcule un "développement" en base $\frac{3}{2}$ de $r_0 = \{z\}$, qui est une suite de 0 et de 1 et montre que ce développement est unique pour un entier donné $h_0 = \lfloor z \rfloor$. Avec nos notations, cette suite est simplement $\mathsf{w}(h_0)$, i.e. le mot infini minimal dans $T_{\frac{3}{2}}$ qui débute dans un noeud d'étiquette h_0. K. MALHER exprime certaines conditions concernant cette suite, notament le fait qu'elle ne peut pas contenir de facteur 11. La démonstration de l'inexistance des Z-nombres est maintenant transférée à l'étude des mots minimaux (qui existent) et à la démonstration de l'inexistance d'entiers n tels que le mot infini $\mathsf{w}(n)$ satisfait aux conditions mentionnées.

Annexe

Plusieurs démonstrations dans ce travail nécessitent certaines notions de topologie sur l'espace $A^{\mathbb{N}}$. Nous présentons ici quelques résultats suffisants à une bonne compréhension de ces preuves. Notons que, dans cette partie, nous ne supposons plus l'alphabet A fini, mais seulement dénombrable.

Il existe plusieurs manières de définir une topologie naturelle sur $A^{\mathbb{N}}$. La première consiste à associer la *topologie discrète*[3] à l'alphabet A ; l'espace $A^{\mathbb{N}}$ est alors muni de la *topologie produit.* Nous donnons ici deux autres définitions de cette même topologie : l'une obtenue en décrivant directement les ouverts de la topologie, l'autre définie par une distance. Nous caractérisons enfin les ensembles fermés ainsi que les ensembles relativement compacts de cette topologie.

La topologie de $A^{\mathbb{N}}$

Commençons par décrire les ouverts de la topologie.

Nous définissons tout d'abord une nouvelle classe de langages qui est l'ensemble des *langages préfixes.* Un langage L sur A est *préfixe* si pour tout mot w dans L, il n'existe pas de préfixe propre de w dans L, i.e. si L peut s'écrire sous la forme

$$L = L \backslash LA^+.$$

Par exemple, sur l'alphabet $A = \{a, b\}$, le langage $\{a^n b \mid n \geq 0\}$ est un langage préfixe alors que le langage $L = \{ba^n \mid n \geq 0\}$ n'est pas préfixe car pour tout entier n strictement positif, le mot ba^n de L peut également s'écrire sous la forme $ba^{n-1}a$, avec $ba^{n-1} \in L$ et $a \in A$.

Proposition 1. *Les ouverts de $A^{\mathbb{N}}$ sont les ensembles de la forme $LA^{\mathbb{N}}$ avec $L \subset A^*$. Ils sont aussi les ensembles de la forme $LA^{\mathbb{N}}$, où L est un langage préfixe de A^*.*

Démonstration. Par définition, les ensembles de la forme

$$A^{n_1} a_1 A^{n_2} a_2 \cdots A^{n_k} a_k A^{\mathbb{N}}, \qquad \text{avec } k \geq 0, n_1, \ldots, n_k \geq 0,$$

3. Dans laquelle tout sous-ensemble de A est ouvert, et donc aussi fermé.

forment un base d'ouverts de la topologie produit. Ces ensembles sont bien de la forme $LA^{\mathbb{N}}$ et, réciproquement, tout ensemble de la forme $LA^{\mathbb{N}}$ est une union de tels ensembles car

$$LA^{\mathbb{N}} = \bigcup_{a_1 \cdots a_k \in L} a_1 \cdots a_k A^{\mathbb{N}}.$$

Les ensembles de la forme $LA^{\mathbb{N}}$ forment donc une base d'ouverts de la topologie et comme cette base est fermée pour l'union, elle contient tous les ouverts de la topologie.

Un langage de la forme $L = L \backslash LA^+$ étant un langage préfixe, la seconde partie de la proposition découle de la formule

$$LA^{\mathbb{N}} = \left(L \backslash LA^+\right) A^{\mathbb{N}}.$$

Tout d'abord, il est clair que tout mot appartenant à $(L \backslash LA^+)\, A^{\mathbb{N}}$ appartient également à $LA^{\mathbb{N}}$. Supposons maintenant que le mot $w\boldsymbol{a}$ appartient à $LA^{\mathbb{N}}$. Si w appartient à LA^+, il suffit de considérer le plus grand suffixe v de w tel que le préfixe u de w vérifiant $w = uv$ appartienne au langage L. Nous obtenons alors $w\boldsymbol{a} = u\boldsymbol{a'}$, avec $\boldsymbol{a'} = v\boldsymbol{a} \in A^{\mathbb{N}}$, et le mot u n'appartient pas à LA^+ vu la manière dont nous l'avons choisi, ce qui conclut.

□

Passons maintenant à la description de la topologie par une distance.

Puisque l'alphabet A est muni de la topologie discrète, il est métrisable : il suffit de considérer la distance δ' sur A définie par

$$\forall a, b \in A, \qquad \delta'(a, b) = \begin{cases} 0 & \text{si } a = b \\ 1 & \text{sinon.} \end{cases}$$

Le résultat suivant est purement topologique et ne sera donc pas démontré ici. Nous revoyons le lecteur dans [5] pour plus de détails.

Proposition 2. *Le produit de toute famille dénombrable d'espaces métrisables*[4] *(resp. d'espaces complétement métrisables*[5]*) est un espace métrisable (resp. un espace complétement métrisable).*

Vu le résultat précédent, l'espace $A^{\mathbb{N}}$ est métrisable et la proposition suivante fournit une distance pour celui-ci.

Proposition 3. *La topologie sur $A^{\mathbb{N}}$ est donnée par la distance δ définie par*

$$\delta(\boldsymbol{a}, \boldsymbol{b}) = 2^{-r}, \qquad \textit{avec } r = \min\{n \mid \boldsymbol{a}_n \neq \boldsymbol{b}_n\},$$

avec les conventions habituelles $\min \emptyset = +\infty$ *et* $2^{-\infty} = 0$.

4. Rappelons qu'un espace topologique est *métrisable* si sa topologie peut être définie par une distance.

5. Un espace *complétement métrisable* est un espace dont la topologie peut être définie par une distance complète.

Démonstration. Il suffit de montrer que les ouverts de la topologie induite par la distance et ceux de la topologie produit coïncident.

D'une part, la base d'ouverts de la topologie induite par δ est l'ensemble des boules ouvertes

$$B(\boldsymbol{a}, \varepsilon) = \{\boldsymbol{b} \in A^{\mathbb{N}} \mid \delta(\boldsymbol{a}, \boldsymbol{b}) < \varepsilon\}, \qquad \boldsymbol{a} \in A^{\mathbb{N}},$$

et la formule

$$B\left(\boldsymbol{a}, 2^{-k}\right) = \{\boldsymbol{b} \in A^{\mathbb{N}} \mid \delta(\boldsymbol{a}, \boldsymbol{b}) < 2^{-k}\} = \boldsymbol{a}_{[k+1]} A^{\mathbb{N}}$$

montre que pour tout mot infini $\boldsymbol{a}$ sur A et pour tout entier positif k, la boule ouverte centrée en $\boldsymbol{a}$ et de rayon 2^{-k} est ouverte dans la topologie produit.

D'autre part, les formules

$$LA^{\mathbb{N}} = \bigcup_{u \in L} uA^{\mathbb{N}} \quad \text{et}^{6} \quad uA^{\mathbb{N}} = B\left(u, 2^{-|u|+1}\right)$$

montrent que tout ouvert de la topologie produit est une union de boules ouvertes pour la distance δ, d'où la conclusion.

□

Remarquons que pour deux mots infinis $\boldsymbol{a}$ et $\boldsymbol{b}$ sur A, nous avons $\delta(\boldsymbol{a}, \boldsymbol{b}) \leq 2^{-k}$ si et seulement si les k premières lettres des deux mots coïncident. Par conséquent, plus deux mots infinis ont un long préfixe en commun, plus ils sont proches pour la distance δ. Nous obtenons par exemple

$$\begin{aligned} \delta(aba\underline{a}b^{\omega}, aba\underline{b}ba^{\omega}) &= 2^{-3} \\ \delta(\underline{a}ba^{\omega}, (\underline{b}a)^{\omega}) &= 2^{-0} \\ \delta(aba^{\omega}, aba^{\omega}) &= 2^{-\infty} \end{aligned}$$

où la lettre soulignée est la première que les deux mots n'ont pas en commun.

Proposition 4. *Pour tout alphabet A, l'espace métrique $\left(A^{\mathbb{N}}, \delta\right)$ est complet.*

Démonstration. Cela découle directement de la proposition 2, sachant que l'espace métrique (A, δ') est complet. □

Ensembles fermés

La proposition 1 fournit une descritpion combinatoire des ouverts de $A^{\mathbb{N}}$. Dans la même idée, nous pouvons donner une caractérisation de l'adhérence des sous-ensembles de $A^{\mathbb{N}}$.

6. Nous étendons la distance δ à l'ensemble des mots finis et infinis sur A en considérant un nouveau symbole $\xi \notin A$ et en ajoutant à tout mot fini u sur A le suffixe ξ^{ω}.

Proposition 5. *Soit X un sous-ensemble de $A^{\mathbb{N}}$ et soit $\boldsymbol{a}$ un mot infini sur A. Le mot infini $\boldsymbol{a}$ appartient à l'adhérence de X si et seulement si tout préfixe de $\boldsymbol{a}$ est préfixe d'un mot de X.*

Démonstration. Par définition de l'adhérence d'un ensemble, le mot infini $\boldsymbol{a}$ appartient à $\overline{X}$ si et seulement si pour tout entier positif n, il existe un mot $\boldsymbol{b}$ dans X tel que $\delta(\boldsymbol{a}, \boldsymbol{b}) < 2^{-n}$. Cela revient à dire que pour tout entier positif n, le mot fini $\boldsymbol{a}_{[n+1]}$ est un préfixe d'un mot de X.

□

Considérons quelques exemples.

1. Si B est un sous-ensemble de A, alors $B^{\mathbb{N}}$ est un sous-ensemble fermé de $A^{\mathbb{N}}$. En effet, si $B^{\mathbb{N}}$ n'est pas fermé, alors il existe une mot infini $\boldsymbol{a}$ sur A appartenant à $\overline{B^{\mathbb{N}}} \backslash B^{\mathbb{N}}$. Si $\boldsymbol{a}$ n'appartient pas à $B^{\mathbb{N}}$, alors il existe un entier positif n tel que la lettre $\boldsymbol{a}_n$ de $\boldsymbol{a}$ n'appartient pas à B et donc le préfixe $\boldsymbol{a}_{[n+1]} = \boldsymbol{a}_0 \boldsymbol{a}_1 \cdots \boldsymbol{a}_n$ de $\boldsymbol{a}$ n'est pas préfixe d'un mot de $B^{\mathbb{N}}$ et n'appartient donc pas à $\overline{B^{\mathbb{N}}}$, d'où une contradiction.
2. Si L est un ensemble de mots de longueur n de A^+, alors l'ensemble $LA^{\mathbb{N}}$ est fermé. En effet, l'égalité
$$A^{\mathbb{N}} \backslash LA^{\mathbb{N}} = (A^n \backslash L)\, A^{\mathbb{N}}$$
montre que $LA^{\mathbb{N}}$ est le complémentaire d'un ouvert de $A^{\mathbb{N}}$.

Nous allons maintenant pouvoir donner un caractérisation des ensembles fermés de $A^{\mathbb{N}}$. Nous avons pour cela besoin d'un lemme, mais nous devons introduire auparavant une nouvelle notation :

$$\forall L \subseteq A^*,\ \vec{L} = \left\{ \boldsymbol{a} \in A^{\mathbb{N}} \mid \boldsymbol{a} \text{ possède une infinité de préfixes dans } L \right\}.$$

Par exemple, si $L = \{a^n b \mid n \geq 0\}$, alors il est clair que $\vec{L} = \emptyset$. Par contre, pour $L = \{(ab)^{2n} \mid n \geq 0\}$, nous avons $\vec{L} = \{(ab)^{\omega}\}$.

Lemme 6. *Soit X un sous-ensemble de $A^{\mathbb{N}}$. Les assertions suivantes sont équivalentes.*

1. *Il existe un langage préfixiel L sur A tel que X est l'ensemble des mots infinis sur A dont tous les préfixes appartiennent à L ;*
2. *il existe un langage M sur A tel que X est l'ensemble des mots infinis sur A n'ayant aucun préfixe dans M ;*
3. *il existe un langage préfixiel L sur A tel que $X = \vec{L}$;*
4. *$X = \overrightarrow{Pref(X)}$.*

Ce lemme pourrait paraître trivial en ce sens que pour tout sous-ensemble X de $A^{\mathbb{N}}$, nous aimerions simplement prendre comme langage L (par exemple pour le point *1*), le langage formé de tous les préfixes des mots de X. L'exemple suivant montre en quoi cela est une erreur. En effet, sur l'alphabet $A = \{a, b\}$, considérons l'ensemble $X = A^{\mathbb{N}} \backslash \{a^\omega\}$. Si un langage L satisfaisant le point *1* existe, il doit nécessairement contenir tous les mots sur A, i.e. $L = A^*$. Or, dans ce cas, l'ensemble des mots infinis sur A dont tous les préfixes sont dans L est l'ensemble $A^{\mathbb{N}} \supsetneq X$, d'où une contradiction.

Démonstration. Montrons que *1* implique *2*. Il suffit de prendre $M = A^* \backslash L$. En effet, si un mot infini $\boldsymbol{a}$ appartient à X, alors tous ses préfixes sont dans L et, réciproquement, un mot infini n'ayant aucun préfixe dans M a tous ses préfixes dans L et appartient donc à X.

Montrons que *2* implique *1*. Il suffit d'appliquer le même raisonnement que pour le cas précédent en considérant $L = A^* \setminus MA^*$.

Montrons que les points *1* et *3* sont équivalents. Il faut montrer que si L est un langage préfixiel de A^*, alors tout mot infini $\boldsymbol{a}$ sur A ayant une infinité de préfixes dans L a tous ses préfixes dans L. Par hypothèse, la longueur des préfixes de $\boldsymbol{a}$ qui sont dans L n'est pas bornée. Donc, si le mot u est un préfixe de $\boldsymbol{a}$, il est préfixe d'un autre préfixe v de $\boldsymbol{a}$ qui est dans L. Puisque le langage L est préfixiel, le mot u appartient également à L.

Montrons que *4* implique *3*. En effet, le langage $\mathrm{Pref}(X)$ est préfixiel dans A^*.

Montrons que *3* implique *4*. Nous avons $X = \vec{L}$, où L est un langage préfixiel de A^*. Si un mot u appartient au langage $\mathrm{Pref}(X)$, alors il existe un mot infini $\boldsymbol{a}$ sur A tel que $u\boldsymbol{a}$ appartient à $X = \vec{L}$. Puisque $u\boldsymbol{a}$ possède une infinité de préfixes dans L, u est préfixe d'un mot de L et puisque ce langage est préfixiel, le mot u lui appartient. Nous obtenons donc $\mathrm{Pref}(X) \subset L$. Par conséquence, il vient $X \subseteq \overrightarrow{\mathrm{Pref}(X)} \subseteq \vec{L} = X$ et donc $X = \overrightarrow{\mathrm{Pref}(X)}$.

□

Proposition 7. *Soit X un sous-ensemble de $A^{\mathbb{N}}$. Les assertions suivantes sont équivalentes.*

1. *X est fermé ;*
2. *tout mot infini sur A dont chaque préfixe est préfixe d'un mot de X appartient à X ;*
3. *$X = \overrightarrow{Pref(X)}$.*

Démonstration. L'équivalence entre les points *1* et *2* découle de la proposition 5. Montrons que *2* implique *3*. Nous avons bien entendu $X \subset \overrightarrow{\mathrm{Pref}(X)}$. Soit alors $\boldsymbol{a}$ un mot infini appartenant à $\overrightarrow{\mathrm{Pref}(X)}$. Tout préfixe de $\boldsymbol{a}$ est un préfixe d'un mot de X et donc $\boldsymbol{a} \in X$.

Montrons que *3* implique *1*. Par le lemme 6, il existe un langage M de A^* tel que l'ensemble X est l'ensemble des mots infinis sur A dont aucun préfixe n'appartient à M. Cela revient à dire que le complémentaire de X dans $A^{\mathbb{N}}$ est l'ensemble $MA^{\mathbb{N}}$, c'est-à-dire un ouvert de $A^{\mathbb{N}}$ et X est donc fermé.

□

Par exemple, pour tout mot infini $\boldsymbol{a}$ sur A, le singleton $\{\boldsymbol{a}\}$ est un sous-ensemble fermé de $A^{\mathbb{N}}$.

Compacité

Nous caractérisons ici les sous-ensembles *relativement compacts*[7] de $A^{\mathbb{N}}$.

Nous notons p_n la projection de $A^{\mathbb{N}}$ sur A définie par $p_n(\boldsymbol{a}) = \boldsymbol{a}_n$.

Proposition 8. *Un sous-ensemble X de $A^{\mathbb{N}}$ est relativement compact si et seulement si pour tout entier positif n, l'ensemble $p_n(X)$ est fini.*

Démonstration. Si, pour tout entier positif n, l'ensemble $p_n(X)$ est fini, donc compact dans A, alors l'ensemble X est inclus dans le produit de compacts $\prod_{n\in\mathbb{N}} p_n(X)$ et est donc compact.

Réciproquement, s'il existe un entier positif k tel que $p_k(X)$ est infini, alors il existe une suite infinie $(a_n)_{n\in\mathbb{N}}$ de lettres de A distinctes deux-à-deux et une suite infinie $(\boldsymbol{a}(n))_{n\in\mathbb{N}}$ de mots infinis dans X tels que $p_k(\boldsymbol{a}(n)) = a_n$ pour tout entier positif. Considérons la famille de boules

$$\left\{B\left(\boldsymbol{u}, 2^{-(k+1)}\right) \mid \boldsymbol{u} \in X\right\}.$$

Il s'agit d'un recouvrement ouvert de X dont nous ne pouvons extraire de recouvrement fini. L'ensemble X n'est donc pas compact et, par conséquences, pas relativement compact non plus.

□

Vu la proposition précédente, nous avons directement :

Corollaire 9. *L'espace $A^{\mathbb{N}}$ est compact si et seulement si l'alphabet A est fini.*

7. Rappelons qu'un ensemble est *relativement compact* si son adhérence est compact.

Bibliographie

[1] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company, 1994.

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

[3] K. Mahler. An unsolved problem on the powers of 3/2. *J. Austral. Math. Soc.*, 8 :313–321, 1968.

[4] A. M. Odlyzko and H. S. Wilf. Functional iteration and the Josephus problem. *Glasgow Math. J.*, 33(2) :235–240, 1991.

[5] D. Perrin and J.-É. Pin. *Infinite Words*, volume 141 of *Pure and Applied Mathematics*, chapter 3, pages 133–152. Elsevier, 2004.

[6] M. Rigo. *Compléments en théorie des langages formels*. ULg, 2005. Notes de cours.

[7] M. Rigo. *Théorie des automates et langages formels*. ULg, 2005. Notes de cours.

[8] M. Rigo. *Mathématiques discrètes*. ULg, 2007. Notes de cours.

[9] J. Sakarovitch, S. Akiyama, and C. Frougny. Powers of rationals modulo 1 and rational base number systems. *Israël J. Math.*, 2007. à paraître.

[10] É. W. Wesstein. Power fractional parts. from MathWorld, http ://mathworld.wolfram.com/PowerFractionalParts.html.

Table des matières