N° d'ordre : 276-2010 Année 2010

THÈSE

présentée devant

l'**UNIVERSITÉ CLAUDE BERNARD-LYON 1**

pour l'obtention

du **DIPLÔME DE DOCTORAT**

(arrêté du 7 août 2006)

présentée et soutenue publiquement le 8 Décembre 2010

par

Alina FIRICEL

Quelques contributions à l'étude des séries formelles à coefficients dans un corps fini

---

Après avis des rapporteurs :

M. Jason BELL	Professeur à Simon Fraser University
M. Christian MAUDUIT	Professeur à l'Université de Méditerranée

Devant le jury composé de :

M. Boris ADAMCZEWSKI	Chargé de Recherche au CNRS
M. Jean-Paul ALLOUCHE	Directeur de Recherche au CNRS
M^{me}. Valérie BERTHÉ	Directrice de Recherche au CNRS
M. Christian MAUDUIT	Professeur à l'Université de Méditerranée
M. Federico PELLARIN	Professeur à l'Université Jean Monnet
M. Luca ZAMBONI	Professeur à l'Université Lyon 1

Thése de doctorat
Université Claude Bernard Lyon 1
Institut Camille Jordan

Quelques contributions à l'étude des séries formelles à coefficients dans un corps fini

Alina FIRICEL
sous la direction de **Boris ADAMCZEWSKI**

Remerciements

Une thèse, c'est trois ans de sa vie, de travail, de rencontres diverses, de moments partagés avec les autres thésards, de hauts et de bas, de passages difficiles ou bien de grande satisfaction. J'ai eu l'opportunité d'être entourée pendant ces trois ans par des personnes qui ont participé, à leur manière, à la réussite de cette thèse. Je tiens très sincèrement à les en remercier !

Tout d'abord, je veux exprimer mon immense gratitude envers mon directeur de thèse, Boris Adamczewski, pour avoir dirigé mes premiers pas dans le monde de la recherche au travers de sujets intéressants. Il a su être très présent au cours de ces années tout en me laissant une grande liberté dans mes choix. Ses idées, sa motivation, ainsi que sa grande curiosité et culture mathématique m'ont servi de modèle. Grâce à ses conseils, sa patience et sa gentillesse, j'ai pu mener à bien ce projet de thèse. C'est pour tout cela que je le remercie infiniment.

Je souhaite exprimer toute ma reconnaissance envers mes rapporteurs, Jason Bell et Christian Mauduit, qui m'ont fait l'honneur de lire et de commenter mon mémoire de thèse. L'intérêt qu'ils ont porté à ce travail, leurs remarques et suggestions m'ont permis d'améliorer la qualité de mon manuscrit et je les en remercie.

Je remercie également Jean-Paul Allouche, Valérie Berthé, Federico Pellarin et Luca Zamboni pour avoir accepté d'examiner mon mémoire et de faire partie de mon jury de thèse. Pendant ces années, j'ai eu la chance d'avoir des discussions avec chacun d'entre eux qui m'ont beaucoup apporté. En particulier, certains travaux de Jean-Paul Allouche et de Valérie Berthé ont beaucoup influencé le départ de ma thèse.

Cette thèse a été effectuée à l'Institut Camille Jordan, au sein de l'équipe de théorie des nombres et combinatoire. Je tiens à remercier tous ses membres. Un grand merci à Christophe Delaunay pour son aide ainsi que pour son constant soutien. Merci aussi à Gaelle Dejou, ma collègue de bureau et de séminaires ; nos thèses respectives n'ont pas toujours été des moments de plaisir, mais nous avons su nous encourager mutuellement pour pouvoir avancer.

Les différents chercheurs ou doctorants que j'ai pu rencontrer lors des séminaires ou conférences, et avec qui j'ai pu échanger des discussions intéressantes, ont eu leur part dans la réussite de cette thèse. Qu'ils en soient tous remerciés. Merci tout particulièrement à Alain Lasjaunias pour les discussions mathématiques que nous avons eues ; il m'a fait découvrir de nombreux travaux autour

du développement en fraction continue des séries formelles et je lui en suis reconnaissante. Merci également à Vincent Delecroix pour son aide et pour sa grande disponibilité.

Je voudrais aussi exprimer ma reconnaissance envers mes professeurs qui, tout au long de mon cursus, ont su me transmettre leur passion pour les mathématiques et m'ont ainsi donné le désir de travailler dans ce domaine. Parmi eux, je voudrais mentionner Iuliana Coravu, Georges Grekos et Dragos Iftimie.

Mes pensées se tournent naturellement vers mes collègues doctorants ou exdoctorants avec qui j'ai pu partager mes réussites et mes doutes au quotidien. En particulier, je remercie Amélie et Elodie qui ont toujours su mettre une bonne ambiance dans les pauses café-cigarette ; merci aussi pour leur patience pendant la correction du manuscrit (un grand merci à Elodie pour son temps consacré à l'orthographe, mais surtout à la ponctuation de mon manuscrit !). Je remercie également Yoann, mon cher collègue et ami ; son optimisme et sa volonté d'avancer ont toujours été des plus encourageants et je lui exprime ma profonde sympathie. Merci aussi à tous les autres collègues avec qui j'ai partagé de très bons moments pendant ces années : Fred, Ioana, Laurent, Marianne, Mickaël, Nicolas, Onu, Pierre, Romain, Thomas, ...

Je tiens bien sûr à remercier mes amis qui ont toujours été là pour partager des moments de ma vie en dehors du labo. Merci à Adi, Alex, Alin, Amira, Cristina, Daiana, Damien, Elena, John, Lilia, Robert, Victor (et j'en oublie sûrement) pour toutes les belles soirées et vacances passées ensembles ! Je voudrais aussi mentionner mes amis de Roumanie : Carmen, Cornel, Lori, Roxana, Sorin qui me donnent à chaque coup de fil ou chaque retrouvaille une bouffée d'oxygène. Dédicace toute particulière à Mara, mon amie de longue date, qui a été tout le temps à mes côtés ; elle a toujours su avoir les bons mots au bon moment et je l'en remercie. Malgré la distance, notre amitié est restée intacte et c'est très important pour moi de pouvoir continuer à partager de bons moments avec vous tous !

Sans doute, mes remerciements les plus profonds s'adressent à ma famille à qui je dédie cette thèse. Je dois à mes parents tout mon parcours et mes réussites obtenues grâce à leur sacrifice, à la confiance et à l'amour qu'ils ont toujours su m'accorder. Avec des parents comme vous, j'ai souvent l'impression d'être la personne la plus chanceuse du monde !

Ma dernière pensée sera pour Alex, celui qui m'a supporté, dans tous les sens du terme, tout au long de cette thèse et avec qui je partage ma vie depuis près de sept ans maintenant. Tu as été à mes côtés dans les meilleurs moments comme dans les pires, tu as suivi mon parcours et tu m'as encouragé à continuer chaque fois que je voulais baisser les bras. Pour tout cela et pour ces années passionantes, du fond de mon coeur : MERCI !

Résumé

Cette thèse se situe à l'interface de trois grands domaines : la combinatoire des mots, la théorie des automates et la théorie des nombres. Plus précisément, nous montrons comment des outils provenant de la combinatoire des mots et de la théorie des automates interviennent dans l'étude de problèmes arithmétiques concernant les séries formelles à coefficients dans un corps fini.

Le point de départ de cette thèse est un célèbre théorème de Christol qui caractérise les séries de Laurent algébriques sur le corps $\mathbb{F}_q(T)$, l'entier q désignant une puissance d'un nombre premier p, en termes d'automates finis[1], et dont l'énoncé est : « Une série de Laurent à coefficients dans le corps fini $\mathbb{F}_q$ est algébrique si et seulement si la suite de ses coefficients est engendrée par un p-automate fini ». Ce résultat, qui révèle dans un certain sens la simplicité de ces séries de Laurent, a donné naissance à des travaux importants parmi lesquels de nombreuses applications et généralisations. La théorie des automates et la combinatoire des mots interviennent naturellement et s'avèrent, parfois, indispensables pour établir des résultats arithmétiques importants. Citons par exemple les travaux d'Allouche, Berthé et Thakur [12, 29, 27, 28, 121] sur la transcendance de certains analogues de Carlitz ou bien ceux de Thakur sur la transcendance de la période de Tate [119] (voir aussi [21]) ; plus récemment, l'article de Kedlaya [75] dans lequel est décrite en termes d'automates la clôture algébrique du corps des fractions rationnelles à coefficients dans un corps fini, ou encore le travail de Derksen [56] sur l'analogue du théorème de Skolem-Mahler-Lech en caractéristique non nulle[2], qui a été ensuite étendu au cas de plusieurs variables par Adamczewski et Bell [6] .

L'objet principal de cette thèse est, dans un premier temps, d'exploiter la simplicité des séries de Laurent algébriques à coefficients dans un corps fini afin d'obtenir des résultats diophantiens, puis d'essayer d'étendre cette étude à des fonctions transcendantes arithmétiquement intéressantes. Nous nous concentrons tout d'abord sur une classe de séries de Laurent algébriques particulières qui généralisent la fameuse cubique de Baum et Sweet[3]. Le résultat principal

[1]Notons que les automates finis sont les machines de calcul les plus basiques parmi la hiérarchie induite par les travaux fondateurs de Turing [130].

[2]Voir aussi le travail récent de Derksen et Masser[57].

[3]La série de Baum et Sweet est le premier exemple montrant qu'il existe des séries algébriques de degré strictement supérieur à 2 dont les quotients partiels sont bornés.

obtenu pour ces dernières est une description explicite de leur développement en fraction continue, généralisant ainsi certains travaux de Mills et Robbins [95] et de Lasjaunias [82]. Rappelons que le développement en fraction continue permet généralement d'obtenir des informations très précises sur l'approximation rationnelle ; les meilleures approximations étant obtenues directement à partir de la suite des quotients partiels.

Malheureusement, il est souvent très difficile d'obtenir le développement en fraction continue d'une série de Laurent algébrique, que celle-ci soit donné par une équation algébrique ou par son développement en série de Laurent. La deuxième étude que nous présentons dans cette thèse fournit une information diophantienne *a priori* moins précise que la description du développement en fraction continue, mais qui à le mérite de concerner toutes les séries de Laurent algébriques (à coefficients dans un corps fini). L'idée principale est d'utiliser l'automaticité de la suite des coefficients de ces séries de Laurent afin d'obtenir une borne générale pour leur exposant d'irrationalité. Malgré la généralité de ce résultat, la borne obtenue n'est pas toujours satisfaisante. Dans certains cas, elle peut s'avèrer plus mauvaise que celle provenant de l'inégalité de Mahler. Cependant, dans de nombreuses situations, il est possible d'utiliser notre approche pour fournir, au mieux, la valeur exacte de l'exposant d'irrationalité, sinon des encadrements très précis de ce dernier.

Dans un dernier travail nous nous plaçons dans un cadre plus général que celui des séries de Laurent algébriques, à savoir celui des séries de Laurent dont la suite des coefficients a une « basse complexité [4] ». Nous montrons que cet ensemble englobe quelques fonctions remarquables, comme les séries algébriques et l'inverse de l'analogue du nombre π dans le module de Carlitz. Il possède, par ailleurs, des propriétés de stabilité intéressantes : entre autres, il s'agit d'un espace vectoriel sur le corps des fractions rationnelles à coefficients dans un corps fini (ce qui, d'un point de vue arithmétique, fournit un critère d'indépendance linéaire), il est de plus laissé invariant par diverses opérations classiques comme le produit de Hadamard.

[4] On rappelle que la complexité d'une suite infinie est le nombre de ses différents facteurs et qu'une suite automatique a une complexité d'ordre au plus linéaire.

Abstract

This thesis looks at the interplay of three important domains : combinatorics on words, theory of finite-state automata and number theory. More precisely, we show how tools coming from combinatorics on words and theory of finite-state automata intervene in the study of arithmetical problems concerning the Laurent series with coefficients in a finite field.

The starting point of this thesis is a famous theorem of Christol which characterizes algebraic Laurent series over the field $\mathbb{F}_q(T)$, q being a power of the prime number p, in terms of finite-state automata and whose statement is the following : " A Laurent series with coefficients in a finite field $\mathbb{F}_q$ is algebraic over $\mathbb{F}_q(T)$ if and only if the sequence of its coefficients is p-automatic." This result, which reveals, somehow, the simplicity of these Laurent series, has given rise to important works including numerous applications and generalizations. The theory of finite-state automata and the combinatorics on words naturally occur in number theory and, sometimes, prove themselves to be indispensable in establishing certain important results in this domain.

The main purpose of this thesis is, foremost, to exploit the simplicity of the algebraic Laurent series with coefficients in a finite field in order to obtain some diophantine results, then to try to extend this study to some interesting transcendental functions. First, we focus on a particular set of algebraic Laurent series that generalise the famous cubic introduced by Baum and Sweet. The main result we obtain concerning these Laurent series gives the explicit description of its continued fraction expansion, generalising therefore some articles of Mills and Robbins. Unfortunately, it is often very difficult to find the continued fraction representation of a Laurent series, whether it is given by an algebraic equation or by its Laurent series expansion. The second study that we present in this thesis provides a diophantine information which, although *a priori* less complete than the continued fraction expansion, has the advantage to characterize any algebraic Laurent series. The main idea is to use some the automaticity of the sequence of coefficients of these Laurent series in order to obtain a general bound for their irrationality exponent. In the last part of this thesis we focus on a more general class of Laurent series, namely the one of Laurent series of "low" complexity. We prove that this set includes some interesting functions, as for example the algebraic series or the inverse of the analogue of the real number π. We also show that this set satisfy some nice closure properties : in particular, it is a vector space over the field over $\mathbb{F}_q(T)$.

Structure de la thèse

Ce mémoire comprend deux parties, subdivisées en plusieurs chapitres.

Dans la *première partie*, nous introduisons les concepts combinatoires et arithmétiques qui seront étudiés et employés par la suite. Nous rappelons quelques notions classiques de la combinatoire des mots dans le **chapitre** 1. Le **chapitre** 2 est dédié au lien entre les séries de Laurent algébriques à coefficients dans un corps fini et la théorie des automates finis, notamment au théorème de Christol. Le **chapitre** 3 donne une brève introduction à l'approximation des séries de Laurent par des fractions rationnelles. Nous rappelons en particulier certaines analogies et différences avec le cas classique de l'approximation des nombres réels par des nombres rationnels. Pour terminer, nous présentons brièvement dans le **chapitre** 4 les principaux résultats obtenus, lesquels sont détaillés dans la deuxième partie de cette thèse.

La *deuxième partie*, composée de trois chapitres, regroupe les principaux travaux de cette thèse. Dans le **chapitre** 5, nous présentons une étude diophantienne de séries de Laurent algébriques généralisant la célèbre cubique introduite par Baum et Sweet. Ensuite, nous exposons dans le **chapitre** 6 une technique permettant d'obtenir une majoration générale de l'exposant d'irrationalité des séries de Laurent algébriques sur le corps $\mathbb{F}_q(T)$. Enfin, nous introduisons et nous étudions dans le **chapitre** 7 une notion de complexité pour les séries de Laurent à coefficients dans un corps fini.

Table des matières

I Introduction et aperçu des résultats **3**

1 Combinatoire des mots **5**
1.1 Notations et définitions : mots finis et infinis 5
1.2 Automates finis et suites automatiques 6
1.2.1 Suites automatiques et noyaux 8
1.2.2 Suites automatiques et morphismes de monoïdes 9
1.3 Fonction de complexité 11

2 Algébricité et automaticité : le théorème de Christol **15**
2.1 Quelques généralisations du théorème de Christol 18
2.1.1 Cas multidimensionnel-Théorème de Salon 18
2.1.2 Cas d'un corps quelconque de caractéristique non nulle . 18
2.1.3 Cas des séries généralisées-Théorème de Kedlaya 19
2.2 Quelques conséquences du théorème de Christol 21
2.2.1 Problème de changement de caractéristique 22
2.2.2 Les analogues de Carlitz 24
2.2.3 Application du théoréme de Christol à d'autres corps . . 25

3 Approximation diophantienne en caractéristique non nulle **27**
3.1 Le développement en fraction continue 29
3.2 Une classe spéciale de séries algébriques 32
3.3 Le théorème de Thue 33

4 Aperçu des résultats **35**
4.1 Généralisation de la cubique de Baum et Sweet et fractions continues 35
4.2 Approximation rationnelle 36
4.3 Complexité et séries formelles à coefficients dans un corps fini . 37

II Présentation des travaux 41

5 Sur une généralisation de la cubique de Baum et Sweet 43
5.1 Introduction . . . 43
5.2 Méthode employée et exemple de Mahler . . . 45
5.2.1 Premier exemple . . . 45
5.2.2 Le contexte général . . . 47
5.3 Généralisation de la cubique de Baum et Sweet . . . 49
5.3.1 Démonstration du théorème 5.3.1 . . . 50
A Sur l'exposant d'irrationalité des séries de Baum et Sweet généralisées . . . 57

6 Rational approximation for algebraic Laurent series 59
6.1 Introduction . . . 59
6.2 Proof of Theorem 6.1.2 . . . 63
6.2.1 Maximal repetitions in automatic sequences . . . 63
6.2.2 An approximation lemma . . . 63
6.2.3 Construction of rational approximations via Christol's theorem . . . 65
6.2.4 An equivalent condition for coprimality of P_n and Q_n . . . 68
6.3 Matrix associated with morphisms . . . 69
6.4 Examples . . . 74

7 Subword complexity and Laurent series with coefficients in a finite field 83
7.1 Introduction and motivations . . . 83
7.2 Terminology and basic notions . . . 86
7.2.1 Subword complexity and topological entropy . . . 87
7.2.2 Morphisms . . . 87
7.3 The analog of Π . . . 88
7.3.1 Proof of Part (a) of Theorem 7.1.2 . . . 89
7.3.2 Proof of Part (b) of Theorem 7.1.2 . . . 93
7.4 Closure properties of two classes of Laurent series . . . 96
7.4.1 Proof of Theorem 7.1.3 . . . 96
7.4.2 Other closure properties . . . 104
7.5 Cauchy product of Laurent series . . . 106
7.5.1 Products of automatic Laurent series . . . 106
7.5.2 A more difficult case . . . 109
7.6 Conclusion . . . 110
A Other examples of products of Laurent series . . . 111

Première partie

Introduction et aperçu des résultats

1

Combinatoire des mots

La combinatoire des mots remonte historiquement au début du XX-ième siècle avec les travaux de Thue [127, 129] sur les mots infinis sans carrés. Les domaines d'applications en sont multiples : la musique, la bio-informatique, ou le traitement du langage naturel comme il est décrit dans [89]. Il existe également de nombreuses interactions entre la combinatoire des mots et d'autres branches des mathématiques comme la théorie des nombres, l'algèbre, la géométrie discrète, la dynamique symbolique, la logique ... .

Le premier chapitre de cette thèse est consacré aux notions classiques de la combinatoire des mots comme la fonction de complexité, les mots automatiques, les morphismes de monoïdes libres. Ces notions apparaîtront tout au long de cette thèse. Pour plus de détails concernant la combinatoire des mots, nous renvoyons le lecteur à des références désormais classiques comme le livre d'Allouche et Shallit [20] ou la série de Lothaire [87, 88, 89].

1.1 Notations et définitions : mots finis et infinis

Un *mot* fini ou infini est une juxtaposition de symboles (ou lettres) appartenant à un ensemble non vide, fini ou infini, $\mathcal{A}$, appelé *alphabet*. Etant donné un alphabet $\mathcal{A}$, on note $\mathcal{A}^*$ l'ensemble des mots finis définis sur l'alphabet $\mathcal{A}$.

Soit $V := a_0 a_1 \cdots a_{m-1} \in \mathcal{A}^*$. La longueur d'un mot V est égal au nombre de ses lettres ; elle est notée $|V|$. Le mot de longueur 0 est appelé le mot vide ; il est noté ε. Etant donné un entier positif m, la notation $\mathcal{A}^m$ désigne l'ensemble

des mots finis de longueur m ; ainsi $\mathcal{A}^* := \cup_{k=0}^{\infty} \mathcal{A}^k$. On note aussi $\mathcal{A}^{\mathbb{N}}$ l'ensemble de tous les *mots infinis* définis sur l'alphabet $\mathcal{A}$.

Dans ce mémoire, nous utilisons typiquement des majuscules U, V, W pour désigner des éléments de $\mathcal{A}^*$ et des lettres minuscules en gras $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ pour désigner des mots infinis. Les éléments de $\mathcal{A}$ sont notés en général par des lettres minuscules ou des chiffres. Dans certains contextes, nous identifions les symboles avec les valeurs des entiers qu'ils représentent. Nous utilisons parfois la notation $\mathcal{A}_m$ pour désigner l'alphabet $\{0, 1, \ldots, m-1\}$.

On définit l'opération de concaténation de deux mots finis $U = u_1 u_2 \cdots u_m$ et $V = v_1 v_2 \cdots v_n$ comme le mot obtenu par juxtaposition :

$$UV = u_1 u_2 \cdots u_m v_1 v_2 \cdots v_n.$$

C'est une opération associative. L'ensemble $\mathcal{A}^*$ muni de l'opération de concaténation est un monoïde, l'élément neutre étant le mot vide ε.

Un mot fini $U = u_1 u_2 \cdots u_r$ est un *sous-mot* (ou *facteur*) d'un mot fini $V = v_1 v_2 \cdots v_m$ (respectivement infini $\mathbf{a} = a_1 a_2 \cdots$) s'il existe un entier i tel que $u_1 u_2 \cdots u_r = v_i v_{i+1} \cdots v_{i+r-1}$ (resp. $u_1 u_2 \cdots u_r = a_i a_{i+1} \cdots a_{i+r-1}$). On dit alors que l'entier i est *l'occurrence* ou *le rang d'apparition* de U dans V (resp. dans $\mathbf{a}$). Autrement dit, U est un facteur de V (resp. de $\mathbf{a}$) s'il existe deux mots, peut-être vides, A et B (respectivement A et $\mathbf{b}$) tels que $V = AUB$ (respectivement $\mathbf{a} = AU\mathbf{b}$). Dans le cas où A est le mot vide, alors on dit que U est un *préfixe* de V (resp. de $\mathbf{a}$).

Pour un entier $n \geq 1$, on note $U^n := UU \cdots U$ la concaténation n fois du mot fini U. Plus généralement, si ω est un nombre réel supérieur ou égal à 1, on note U^ω le mot $U^{\lfloor \omega \rfloor} U'$, où U' est le préfixe de U de longueur $\lceil (\omega - \lfloor \omega \rfloor)|U| \rceil$. Les notations $\lfloor \zeta \rfloor$ et $\lceil \zeta \rceil$ désignent, respectivement, la partie entière et la partie entière supérieure du nombre réel ζ. On note également $U^\infty := UU \cdots$ le mot infini obtenu en concaténant infiniment le mot U.

Un mot infini $\mathbf{a}$ est *ultimement périodique* s'il existe deux mots finis U et V (V non vide) tels que $\mathbf{a} = UV^\infty$. Lorsque $U = \varepsilon$, $\mathbf{a}$ est dit *purement périodique*, ou, plus simplement, *périodique*. Une suite qui n'est pas ultimement périodique est dite *apériodique*.

1.2 Automates finis et suites automatiques

La théorie des automates est apparue naturellement dans plusieurs domaines mathématiques. Les premiers à s'intéresser à ces objets ont été les logiciens. En particulier, Church et Turing ont introduit les notions de *calcul* ou de *machine*. Ensuite la théorie des systèmes dynamiques discrets (notamment avec les travaux de Morse), la théorie de l'information (avec les problèmes de codage étudiés par Schützenberger) ou la linguistique générative (développée par Chomsky en introduisant les concepts de mots, langages ou grammaires) ont eu une influence remarquable sur le développement de la théorie des automates. Les liens entres ces domaines et la théorie des automates font encore l'objet

de recherches très actives.

Les automates finis avec sortie, introduits vers les années 1950, sont, moins formellement, les « machines abstraites » les plus simples : on leur donne un mot sur un alphabet fixé et après lecture successive des lettres de ce mot, l'automate nous donne une réponse du type « vrai/faux » ou, plus généralement, il nous renvoie une information pouvant prendre un nombre fini de valeurs. Une suite infinie $\mathbf{a} = (a_n)_{n\geq 0}$ à valeurs dans un ensemble fini est dite k-automatique s'il existe un automate fini qui, lorsqu'on lui donne en entrée l'écriture en base k de n, nous renvoie le terme a_n.

Nous donnons à présent des définitions plus formelles de ces notions ; nous nous concentrons sur les suites automatiques et, par conséquent, nous n'allons définir que les k-automates.

Définition 1.2.1. *Soit $k \geq 2$ un entier. Un k-automate est la donnée d'un* 6*-uplet*

$$M = (Q, \mathcal{A}_k, \delta, q_0, \Delta, \varphi)$$

où Q est un ensemble fini d'états, $\mathcal{A}_k = \{0, 1, \ldots, k-1\}$, $\delta : Q \times \mathcal{A}_k \mapsto Q$ est la fonction de transition, q_0 est un élément de Q, appelé état initial, Δ est un alphabet fini appelé alphabet de sortie et $\varphi : Q \mapsto \Delta$ est la fonction de sortie.

Etant donnés un état $q \in Q$ et un mot fini $U = u_0u_1 \cdots u_r$ sur l'alphabet $\mathcal{A}_k$, on définit récursivement $\delta(q, U) := \delta(\delta(q, u_0u_1 \cdots u_{r-1}), u_r)$. Pour un mot fini $U = u_r u_{r-1} \cdots u_0 \in \mathcal{A}_k^{r+1}$, on note $[U]_k$ le nombre égal à $\sum_{i=0}^{r} u_i k^i$.

Définition 1.2.2. *Soit $\boldsymbol{a} = (a_n)_{n\geq 0}$ une suite infinie à valeurs dans un ensemble fini. On dit que $\boldsymbol{a}$ est k-automatique s'il existe un k-automate fini tel que $a_n = \varphi(\delta(q_0, U))$, pour tous les mots U tels que $[U]_k = n$.*

***Remarque* 1.2.1.** Soit $n \geq 0$. L'automate nous renvoie le terme a_n après avoir lu tous les chiffres de l'écriture en base k de n. La lecture se fait à partir du chiffre le plus significatif vers le chiffre le moins significatif. Nous pouvons dire que l'automate lit dans « le sens direct » ou bien que la suite est k-automatique dans « le sens direct ». Il est aussi possible de définir une suite automatique en lisant ces chiffres en ordre inverse, (c'est-à-dire du chiffre le moins significatif vers le chiffre le plus significatif) ; on dit alors que l'automate lit dans « le sens inverse ». En réalité, on peut démontrer que ces deux définitions sont équivalentes (voir, par exemple, [104], Proposition 1.3.4).

Exemple 1.2.1 (La suite de Thue-Morse)**.** L'exemple le plus célèbre de suite automatique est la suite de Thue–Morse[1], notée ici $\mathbf{t} = (t_n)_{n\geq 0}$. Une manière de la définir est la suivante : t_n est égal au nombre de « 1 », considéré modulo 2, dans la représentation binaire de n. Cette suite est 2-automatique car elle peut être engendrée par l'automate suivant

$$M = (\{q_0, q_1\}, \{0, 1\}, \delta, q_0, \{0, 1\}, \varphi),$$

où $\delta(q_0, 0) = \delta(q_1, 1) = q_0$, $\delta(q_0, 1) = \delta(q_1, 0) = q_1$ et $\varphi(q_0) = 0$, $\varphi(q_1) = 1$.

[1]Cette suite a été introduite indépendamment par Thue et par Morse au début du XX-ième siècle.

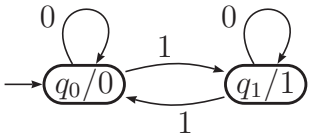


FIGURE 1.1 – L'automate engendrant la suite de Thue-Morse

En effet, on remarque que l'arrivée dans l'état q_0 représente la lecture d'une entrée avec un nombre pair de 1 (donc 0 modulo 2) et l'état q_1 représente la lecture d'une entrée avec un nombre impair de 1 (donc 1 modulo 2). Par exemple, si la donnée d'entrée est le mot $W = 1001100$, qui correspond à l'écriture en base 2 de 76, l'automate retourne le symbole 1, ce qui signifie que $\mathbf{t}_{76} = 1$.

La suite $\mathbf{t}$ possède des propriétés remarquables et nous renvoyons le lecteur à des références plus détaillées qui lui sont consacrées [18, 93]. Une propriété particulière concerne les répétitions dans les mots infinis. Il est facile de vérifier que sur un alphabet à deux lettres il n'existe pas de mots infinis « sans carré ». Un mot sans carré est un mot qui ne contient aucun motif de la forme XX. On peut naturellement se demander s'il contient des « chevauchements », c'est-à-dire des motifs de la forme WWx, où W est un mot (fini et non vide) et x la première lettre de W. Par exemple, le mot *ananas* contient le chevauchement *anana*. En 1912, Thue [129] a montré que la suite $\mathbf{t}$ ne contient aucun chevauchement. A fortiori, elle ne contient pas de « cube », c'est-à-dire, aucun motif de la forme XXX.

Exemple 1.2.2 (La suite de Baum-Sweet)**.** Une suite dont on parle souvent en approximation diophantienne est la suite introduite par Baum et Sweet dans [24]. Elle est définie de la façon suivante : $\mathbf{b} = (b_n)_{n\geq 0}$ où $b_n = 1$ si la représentation en base 2 ne contient aucun bloc de longueur impaire de 0 et $b_n = 0$ dans le cas contraire.

La suite $\mathbf{b}$ est 2-automatique puisqu'on peut montrer qu'elle est engendrée par le 2-automate suivant.

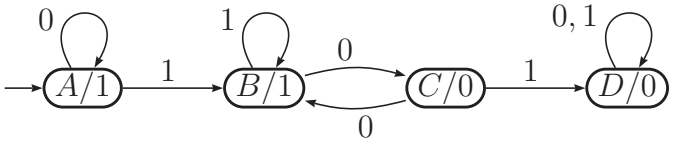


FIGURE 1.2 – L'automate engendrant la suite de Baum-Sweet

Cette suite sera abordée à nouveau dans les chapitres 3 et 5.

1.2.1 Suites automatiques et noyaux

Les suites automatiques s'avèrent, comme nous allons le voir plus tard, très utiles pour démontrer des résultats de théorie des nombres, notamment dans l'étude de l'algébricité des séries de Laurent à coefficients dans un corps fini.

Pour montrer qu'un mot infini est automatique, a priori, nous devons chercher l'automate qui l'engendre. Ceci peut être parfois très difficile et la définiton formelle des k-automates est alors un peu encombrante. C'est pourquoi nous introduisons ici une caractérisation des suites automatiques, qui utilise des sous-suites de la suite étudiée et qui à l'avantage d'être beaucoup plus maniable dans certains contextes.

Définition 1.2.3. *Soient $k \geq 2$ et $\boldsymbol{a} = (a_n)_{n\geq 0}$ une suite infinie. Le k-noyau de $\boldsymbol{a}$ est l'ensemble défini comme suit :*

$$\mathcal{N}_k(\boldsymbol{a}) = \{(a_{k^i n+j})_{n\geq 0} \,:\, i \geq 0 \text{ et } 0 \leq j \leq k^i\}.$$

Une caractérisation intéressante des suites automatiques est le résultat suivant, généralement attribué à Eilenberg [62].

Théorème 1.2.1 (Eilenberg)**.** *Une suite est k-automatique si et seulement si son k-noyau est un ensemble fini.*

L'idée sur laquelle repose ce résultat est le fait que l'écriture en base k de $k^i n + j$, pour $n, i \geq 0$ et $0 \leq j \leq k^i$ est $(n)_k \underbrace{0\cdots 0}_{i-[\log_k(j)]-1} (j)_k$, où, pour un entier $m \geq 0$, la notation $(m)_k$ désigne la représentation en base k de m.

Exemple 1.2.3. On peut vérifier que le 2-noyau de la suite $\mathbf{t}$ de Thue-Morse est formé de deux éléments : la suite $\mathbf{t}$ elle-même et la suite $1-\mathbf{t} := (1-t_n)_{n\geq 0}$, c'est-à-dire, la suite qui échange les lettres 0 et 1. Ceci est dû au fait que $t_{2n} = t_n$ et $t_{2n+1} = (t_n + 1) \mod 2$.

Exemple 1.2.4. Toute suite périodique, ou plus généralement ultimement périodique, est k-automatique, pour tout $k \geq 2$. Ceci peut être vu en construisant explicitement un automate avec un nombre d'états égal à la période de la suite, mais aussi en utilisant la notion de k-noyau et le théorème 1.2.1.

1.2.2 Suites automatiques et morphismes de monoïdes

Soit $\mathcal{A}$ (respectivement $\mathcal{B}$) un alphabet fini et soit $\mathcal{A}^*$ (resp. $\mathcal{B}^*$) le monoïde libre associé.

Définition 1.2.4. *Un morphisme σ est une application de $\mathcal{A}^*$ vers $\mathcal{B}^*$ telle que $\sigma(UV) = \sigma(U)\sigma(V)$ pour tous $U, V \in \mathcal{A}^*$.*

Puisque la concaténation est préservée, nous pouvons définir un morphisme sur $\mathcal{A}$, plutôt que sur $\mathcal{A}^*$. Si $\mathcal{A} = \mathcal{B}$, on peut itérer l'application σ. Ainsi $a \in \mathcal{A}$, $\sigma^0(a) = a$, $\sigma^i(a) = \sigma(\sigma^{i-1}(a))$, pour tout $i \geq 1$.

Définition 1.2.5. *Un morphisme σ est k-uniforme si $|\sigma(a)| = k$ pour tout $a \in \mathcal{A}$. Si $k = 1$, alors σ est appelé un codage.*

L'ensemble $\mathcal{A}^{\mathbb{N}}$ est muni de la topologie produit des topologies discrètes sur chaque copie de $\mathcal{A}$. Cette topologie est induite par la distance d définie de la manière suivante : si $\mathbf{a} = a_0 a_1 \cdots$ et $\mathbf{b} = b_0 b_1 \cdots$ sont deux mots infinis définis sur $\mathcal{A}$, alors

$$d(\mathbf{a}, \mathbf{b}) = \begin{cases} \dfrac{1}{2^{\inf\{i \in \mathbb{N},\, a_i \neq b_i\}}}, \text{ si } \mathbf{a} \neq \mathbf{b}; \\ 0 \text{ sinon.} \end{cases}$$

Moins formellement, on dit que deux mots sont proches s'ils ont un long préfixe commun. Nous pouvons ainsi étendre l'action d'un morphisme par continuité à l'ensemble $\mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$. On dit alors qu'un mot $\mathbf{a} \in \mathcal{A}^{\mathbb{N}}$ est *point fixe* d'un morphisme σ si $\sigma(\mathbf{a}) = \mathbf{a}$; on peut aussi dire que $\mathbf{a}$ est *engendré par le morphisme* σ. Un morphisme σ est *prolongeable* en $a \in \mathcal{A}$ si $\sigma(a) = aX$, où X est un mot non vide tel que $\sigma^k(X) \neq \varepsilon$ pour tout $k \geq 0$. Si σ est prolongeable alors le mot $(\sigma^i(a))_{i \geq 0}$ converge vers le mot infini

$$\sigma^\infty(a) = \lim_{i \to \infty} \sigma^i(a) = aX\sigma(X)\sigma^2(X)\sigma^3(X) \cdots .$$

Exemple 1.2.5 (Le mot de Thue-Morse)**.** Le morphisme τ défini sur l'alphabet $\{0, 1\}$ par $\tau(0) = 01$ et $\tau(1) = 10$ engendre le mot de Thue-Morse

$$\mathbf{t} = \tau^\infty(0) = 01101001100101101001010 \cdots .$$

Il est à noter que la suite de Thue-Morse, qui est 2-automatique, est engendrée par un morphisme 2-uniforme puisque $|\tau(0)| = |\tau(1)| = 2$. En fait, les suites automatiques et les morphismes uniformes sont en lien très fort, comme l'illustre le théorème suivant de Cobham [50].

Théorème 1.2.2 (Cobham)**.** *Un mot infini est k-automatique si et seulement si il est l'image par un codage d'un mot engendré par un morphisme k-uniforme.*

La preuve de ce théorème est constructive. Etant donné un k-automate qui engendre une suite infinie $\mathbf{a}$, alors il est possible de déterminer explicitement le codage et le morphisme qui engendre $\mathbf{a}$ et réciproquement.

Exemple 1.2.6 (Le mot de Baum-Sweet)**.** La suite de Baum-Sweet, définie précédemment par l'automate (1.2) peut être caractérisée comme suit :

$$\mathbf{b} = \tau(\sigma^\infty(A)) = 11011001010010011001 0 \cdots ,$$

où σ et $\tau : \{A, B, C, D\} \mapsto \{0, 1\}$ sont définis par :

$$\begin{array}{ll} \sigma(A) = AB & \tau(A) = 1 \\ \sigma(B) = CB & \tau(B) = 1 \\ \sigma(C) = BD & \tau(C) = 0 \\ \sigma(D) = DD & \tau(D) = 0. \end{array}$$

Mentionnons par ailleurs un théorème du à Eilenberg ([62], Proposition V.3.5).

Théorème 1.2.3. *Pour tout $m \geq 1$, une suite $\boldsymbol{a} = (a_n)_{n\geq 0}$ est k-automatique si et seulement si elle est k^m-automatique.*

Cependant, si une suite est k-automatique, non ultimement périodique, elle ne peut pas etre l-automatique, lorsque k et l sont multiplicativement indépendants (voir le théoréme 2.2.2).

1.3 Fonction de complexité

Etant donnée une suite infinie, on peut se demander à quel point celle-ci est « simple » ; par exemple, est-elle produite par une machine simple ou a-t-elle des propriétés de régularités évidentes ? Une autre façon naturelle de mesurer la « complexité » d'une suite infinie $\mathbf{a}$ est de calculer sa fonction de complexité[2], c'est-à-dire, le nombre de sous-mots (ou facteurs) qui apparaissent dans $\mathbf{a}$.

Définition 1.3.1. *Soit $\boldsymbol{a} = (a_n)_{n\geq 0}$ un mot infini défini sur l'alphabet $\mathcal{A}$. La fonction de complexité de $\boldsymbol{a}$ est la fonction qui associe à chaque $m \in \mathbb{N}$ le nombre $p(\boldsymbol{a}, m)$ défini par*

$$p(\boldsymbol{a}, m) = \mathrm{Card}\{(a_j, a_{j+1}, \ldots, a_{j+m-1}),\ j \in \mathbb{N}\}.$$

Pour tout mot infini $\mathbf{a}$, $p(\mathbf{a}, 0) = 1$, car l'unique mot de longueur 0 est le mot vide ε. Si l'on considère le mot infini $\mathbf{a} = aaa\cdots$, obtenu en concaténant infiniment la lettre a, alors, il est clair que $p(\mathbf{a}, m) = 1$, pour tout $m \geq 0$. Plus généralement, si $\mathbf{a}$ est ultimement périodique, alors, sa fonction de complexité est bornée. D'autre part, si on considère le mot infini de Champernowne

$$\mathbf{a} := 0123456789101112\cdots,$$

alors, pour tout $m \geq 0$, la fonction de complexité vérifie $p(\mathbf{a}, m) = 10^m$. Plus précisément, il est facile de montrer que la fonction de complexité d'un mot infini est croissante (au sens large) et que, pour tout $m \geq 0$, on a :

$$1 \leq p(\mathbf{a}, m) \leq (\mathrm{card}\,\mathcal{A})^m.$$

Une caractéristique importante des mots infinis apériodiques est donnée par le théorème de Morse et Hedlund [96].

Théorème 1.3.1 (Morse et Hedlund)**.** *Si $\boldsymbol{a}$ est un mot infini apériodique, alors sa fonction de complexité de $\boldsymbol{a}$ est strictement croissante. De plus,*

$$p(\boldsymbol{a}, m) \geq m + 1,$$

pour tout $m \geq 0$.

[2]Cette notion a été introduite en 1938 par Morse et Hedlund [96], comme un outil en dynamique symbolique mais le nom de complexité des sous-mots a été donné en 1975 par Ehrenfeucht, Lee et Rozenberg [61].

Des conséquences immédiates découlent de ce théorème. Tout d'abord, si la complexité d'une suite $\mathbf{a}$ satisfait $p(\mathbf{a}, m) \leq m$, pour un entier m, alors $\mathbf{a}$ est ultimement périodique. D'autre part, on déduit aussi de ce théorème que les fonctions équivalentes à $\sqrt{n}$, $\log n$ or $n/2$ ne peuvent pas être fonctions de complexité.

Cela conduit naturellement à sedemander quels types de fonctions peuvent être des fonctions de complexité. Dans cette problématique, plusieurs résultats dûs à Cassaigne [39, 40, 41], utilisant les facteurs spéciaux, donnent des conditions necessaires pour qu'une fonction soit une fonction de complexité. Il a aussi construit des familles d'exemples, permettant ainsi de montrer que de nombreuses fonctions sont des fonctions de complexité. Cependant, beaucoup de questions restent encore sans réponse sur ce sujet.

Mots sturmiens

Comme l'ont remarqué Morse et Hedlund, le théorème 1.3.1 est en fait optimal : il existe des mots dont la complexité est égale à $m + 1$ pour tout m. De tels mots sont appelés *mots sturmiens* ; ils constituent une classe très étudiée en théorie des nombres, géométrie discrète, combinatoire des mots, dynamique symbolique ... . L'exemple le plus célèbre de mot sturmien est le mot infini de Fibonacci défini comme l'unique point fixe qui commence par 0 du morphisme $\varphi : \varphi(0) = 01$ et $\varphi(1) = 0$:

$$\mathbf{f} = \varphi^{\infty}(0) = 0100101001001010010100100 0 \cdots .$$

Les mots sturmiens sont donc les suites apériodiques de complexité minimale définis sur un alphabet à deux lettres.

Mots automatiques

Le résultat suivant décrit la fonction de complexité des suites automatiques.

Théorème 1.3.2 (Cobham)**.** *Soit $\boldsymbol{a}$ une suite infinie apériodique k-automatique, $k \geq 2$. Alors la fonction de complexité vérifie*

$$p(\boldsymbol{a}, m) = \Theta(m),$$

où Θ désigne la notation usuelle de Landau : $f(n) = \Theta(g(n))$ s'il existe deux nombres réels strictement positifs tels que $C_1 g(n) \leq f(n) \leq C_2 g(n)$.

La démonstration de ce résultat repose sur le théorème de Cobham sur les morphismes uniformes (le lecteur pourrait consulter [20], Théorème 10.3.1, page 304).

On peut ainsi calculer explicitement la fonction de complexité de certaines suites k-automatiques. C'est l'objet de l'exemple suivant.

Exemple 1.3.1. La complexité du 2-automatique mot de Thue-Morse satisfait, pour tout $m \geq 0$:

$$p(\mathbf{t}, m) = \begin{cases} 1 & \text{si } n = 0, \\ 2 & \text{si } n = 1, \\ 4 & \text{si } n = 2, \\ 4m - 2 \cdot 2^M - 4 & \text{si } 2 \cdot 2^M < m \leq 3 \cdot 2^M, \\ 2m + 4 \cdot 2^M - 2 & \text{si } 3 \cdot 2^M < m \leq 4 \cdot 2^M. \end{cases}$$

Ce résultat a été démontré par Brlek [34], en utilisant des outils de combinatoire des mots tels que les facteurs spéciaux, mais aussi par Luca et Varrichio [90] et Avgustinovich [22], indépendamment. En fait, ces auteurs ont aussi demontré que la suite $(p(\mathbf{t}, m+1) - p(\mathbf{t}, m))_{m \geq 0}$ est 2-automatique. Plus généralement, Tapsoba [117] et Mossé [98] ont montré que si $\mathbf{a}$ est une suite k-automatique, vérifiant certaines propriétés supplémentaires, alors $(p(\mathbf{a}, m+1) - p(\mathbf{a}, m))_{m \geq 0}$ est aussi une suite k-automatique.

Mots morphiques

Des résultats décrivant la complexité des mots morphiques ont été donnés, plus généralement, par Ehrenfeucht, Lee et Rozenberg [61] : si $\mathbf{a}$ est le point fixe d'un morphisme alors $p(\mathbf{a}, m) = O(m^2)$. La notation O désigne la notation usuelle de Landau : $f(n) = O(g(n))$ s'il existe une constante strictement positive C telle que $f(n) \leq Cg(n)$. Les ordres de croissance possibles de la fonction de complexité ont ensuite été déterminés, de manière detaillée, par Pansiot [101].

Théorème 1.3.3 (Pansiot)**.** *Soit $\boldsymbol{a}$ un mot infini engendré par un morphisme prolongeable. Alors on a une des relations suivantes :*

(i) $p(\boldsymbol{a}, m) = \Theta(1)$,

(ii) $p(\boldsymbol{a}, m) = \Theta(m)$,

(iii) $p(\boldsymbol{a}, m) = \Theta(m \log\log m)$,

(iv) $p(\boldsymbol{a}, m) = \Theta(m \log m)$,

(v) $p(\boldsymbol{a}, m) = \Theta(m^2)$.

En particulier, on peut déduire du théorème de Pansiot que la complexité d'une suite automatique non ultimement périodique est $p(\mathbf{a}, m) = \Theta(m)$. Ce théorème met en évidence en fait 5 classes de morphismes sur un alphabet $\mathcal{A}$, en fonction des ordres de croissance des itérés $\sigma^n(a)$, pour $a \in \mathcal{A}$.

Entropie topologique

La notion de complexité d'une suite infinie est intimement liée à celle d'*entropie topologique* d'une suite laquelle peut être définie par

$$h(\mathbf{a}) = \lim_{m \to \infty} \frac{\log p(\mathbf{a}, m)}{m},$$

où la base de log est la taille de l'alphabet sur lequel le mot $\mathbf{a}$ prend ses valeurs.

Rappelons par ailleurs que l'entropie d'une suite correspond à l'entropie topologique du systeme dynamique topologique sous-jacent : le sous-shift associé à cette suite (voir [77] pour une démonstration).

Cette définition a bien un sens, puisque la limite existe ; ceci est une conséquence de la sous-additivité de la fonction de complexité : pour tous $m, n \geq 0$,

$$p(\mathbf{a}, m+n) \leq p(\mathbf{a}, m)p(\mathbf{a}, n).$$

On remarque que, pour toute suite infinie $\mathbf{a}$, on a

$$0 \leq h(\mathbf{a}) \leq 1.$$

Par exemple, les suites ultimement périodiques ont pour entropie 0, tandis que le mot de Champernowne est d'entropie maximale égale à 1.

Dans cette thèse, nous porterons un intérêt particulier aux suites d'entropie nulle.

2

Algébricité et automaticité : le théorème de Christol

Dans cette thèse, on s'interesse principalement à l'étude de séries formelles à coefficients dans un corps fini. L'objet de ce chapitre est de décrire le lien entre les séries algébriques à coefficients dans un corps fini et les suites automatiques.

Soient p un nombre premier, q une puissance de p et $\mathbb{F}_q$ le corps fini à q éléments. Commençons par rappeler le formalisme classique : $\mathbb{F}_q(T)$, $\mathbb{F}_q[[T^{-1}]]$ et $\mathbb{F}_q((T^{-1}))$ désignant, respectivement, le corps des fonctions rationnelles, l'anneau des séries formelles et le corps des séries de Laurent à coefficients dans $\mathbb{F}_q$. Nous rappelons aussi l'analogie entre l'anneau des entiers et l'anneau des polynômes à coefficients dans $\mathbb{F}_q$, le corps des rationnels et le corps des fractions rationnelles à coefficients dans $\mathbb{F}_q$, le corps des nombres réels (le complété de $\mathbb{Q}$ par la valeur absolue usuelle) et le corps des séries de Laurent (le complété de $\mathbb{F}_q(T)$ par la valeur absolue ultramétrique usuelle), et enfin le corps des nombres complexes (clôture algébrique de $\mathbb{R}$) et le corps $\mathcal{C}_\infty$ (défini comme la complétion d'une clôture algébrique de $\mathbb{F}_q((T^{-1}))$).

Remarquons que les coefficients dans $\mathbb{F}_q$ jouent le rôle de « chiffres » dans la base donnée par les puissances de la variable T. Il y a toutefois une différence importante : dans le cas des nombres réels, il est difficile de contrôler les restes lorsque l'on additionne ou que l'on multiplie, tandis que dans le cas des séries formelles à coefficients dans un corps fini, cette difficulté disparaît.

Les nombres rationnels correspondent aux nombres dont le développement

dans une base entière $b \geq 2$ est ultimement périodique ; les séries de Laurent rationnelles à coefficients dans un corps fini correspondent, elles aussi, aux séries dont la suite des coefficients est ultimement périodique. Pour aller plus loin, on peut se demander comment on peut décrire le développement des nombres réels algébriques ou, parallèlement, les séries de Laurent algébriques sur le corps $\mathbb{F}_q(T)$. On rappelle d'abord la definition suivante.

Définition 2.0.2. *Une série f est algébrique sur $\mathbb{F}_q(T)$ s'il existe un entier $d \geq 1$ et des polynômes $A_0(T), \ldots, A_d(T) \in \mathbb{F}_q[T]$, non tous nuls, tels que :*

$$A_0 + A_1 f + A_2 f^2 + \cdots + A_d f^d = 0.$$

Si le polynôme $P(X) = A_0 + A_1 X + A_2 X^2 + \cdots + A_d X^d \in \mathbb{F}_q(T)[X]$ est irreductible, alors, d est appelé le degré d'algébricité de f.

Cette question sur les développements des nombres réels algébriques a été posée pour la première fois par Borel [33] qui a conjecturé que tout nombre irrationnel algébrique est normal. Rappelons qu'un nombre réel est *normal* s'il est normal en toute base entière $b \geq 2$; c'est-à-dire, si, pour tout entier n, chacun des b^n mots de longueur n sur l'alphabet $\{0, 1, \ldots, b-1\}$ apparaît dans son développement b-adique avec la même fréquence $1/b^n$. Bien que cette conjecture soit considérée comme hors d'atteinte, des progrès ont été fait concernant la fonction de complexité du développement b-adique d'un nombre irrationnel algébrique [4, 64]. D'autre part, il a été conjecturé en 1968 par Cobham [48] que la suite des chiffres du développement b-adique d'un nombre algébrique irrationnel est trop complexe pour pouvoir être engendrée par un automate fini. En 2006, Adamczewski et Bugeaud ont démontré cette conjecture [4] en utilisant une méthode fondée sur le théorème du sous-espace de Schmidt.

Théorème 2.0.4 (Adamczewski et Bugeaud)**.** *Soit $b \geq 2$ un entier et ξ un réel. Si le développement en base b de ξ est engendré par un automate fini, alors ξ est soit rationnel, soit transcendant.*

Contrairement à ce qui se passe pour les nombres irrationnels algébriques, où des informations sur le développement b-adique sont difficiles à obtenir, la situation des séries irrationnelles algébriques s'avère beaucoup mieux comprise. Le premier progrès majeur dans cette direction a été fait en 1967 par Furstenberg [66] avec le résultat suivant.

Théorème 2.0.5 (Furstenberg)**.** *Soit k le corps $\mathbb{Q}$ ou le corps $\mathbb{F}_q$. L'ensemble des séries algébriques sur $k(T)$ est égal à l'ensemble des diagonales des séries rationnelles à deux variables. De plus, la diagonale d'une série rationnelle à plusieurs variables à coefficients dans $\mathbb{F}_q$ est algébrique.*

On rappelle que la diagonale d'une série à k variables

$$R = \sum r_{n_1,n_2,\ldots,n_k} T_1^{n_1} T_2^{n_2} \cdots T_k^{n_k}$$

est définie par $D(T) = \sum r_{n,n,\ldots,n} T^n$. Ce résultat, dans le cas de la caractéristique non nulle, a été généralisé par Deligne [54] qui a montré que toute

diagonale d'une série formelle algébrique à plusieurs variables est aussi algébrique.

Mais le résultat le plus remarquable est sans doute le théorème de Christol [46, 47] qui décrit en terme d'automates finis le développement en série de Laurent des éléments de $\mathbb{F}_q((T^{-1}))$ algébriques sur le corps de fractions rationnelles $\mathbb{F}_q(T)$.

Théorème 2.0.6 (Christol)**.** *Soit $f(T) = \sum_{n \geq -n_0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$. Alors, $f(T)$ est algébrique sur le corps $\mathbb{F}_q(T)$ si et seulement si la suite $(a_n)_{n \geq 0}$ est p-automatique.*

Considérons par exemple la série $f_{\mathbf{t}}(T) := \sum t_n T^{-n} \in \mathbb{F}_2((T^{-1}))$, où $\mathbf{t} = (t_n)_{n \geq 0}$ est la suite de Thue-Morse définie dans le chapitre précédent. On peut aisément montrer que $t_{2n} = t_n \mod 2$ et $t_{2n+1} = (t_n + 1) \mod 2$. Ceci nous permet d'écrire

$$\begin{aligned} f_{\mathbf{t}}(T) &= \sum t_{2n} T^{-2n} + \sum t_{2n+1} T^{-2n-1} \\ &= \sum t_n T^{-2n} + T^{-1} \sum (t_n + 1) T^{-2n} \\ &= f_{\mathbf{t}}(T^2) + \frac{1}{T} f_{\mathbf{t}}(T^2) + \frac{T}{T^2 - 1}. \end{aligned}$$

Puisque la caractéristique du corps est 2, nous obtenons, par le morphisme de Frobenius, que $f_{\mathbf{t}}(T^2) = f_{\mathbf{t}}(T)^2$. En conséquence, la série $f_{\mathbf{t}}$ satisfait l'équation algébrique suivante :

$$(T+1)^3 f_{\mathbf{t}}^2(T) + T(T+1)^2 f_{\mathbf{t}}(T) + T^2 = 0.$$

Comme il est possible de le remarquer dans l'exemple précédent, l'ingrédient principal de la preuve du théorème de Christol est la finitude du p-noyau. En effet, puisqu'une suite est p-automatique si et seulement si son p-noyau est fini, il est possible de reformuler le théorème de Christol de la manière suivante.

Théorème 2.0.7. *Soit $f(T) = \sum_{n \geq -n_0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$. Alors, $f(T)$ est algébrique sur le corps $\mathbb{F}_q(T)$ si et seulement si il existe un nombre fini de sous-suites de la forme $(a_{p^i n + j})_{n \geq 0}$, avec $0 \leq j < p^i$.*

***Remarque* 2.0.1.** Le théorème de Christol est, en général, cité pour les séries de Laurent de la forme $f(T) = \sum a_n T^n \in \mathbb{F}_q((T))$ en utilisant la même définition pour le p-noyau. En effet, la série $\sum a_n T^n$ est transcendante sur $\mathbb{F}_q(T)$ si et seulement si la série $\sum a_n T^{-n}$ est transcendante sur $\mathbb{F}_q(T)$.

***Remarque* 2.0.2.** Nous rappelons aussi qu'une série de Laurent est algébrique sur $\mathbb{F}_q(T)$ si et seulement si elle est algébrique sur $\mathbb{F}_p(T)$, où q est une puissance du nombre premier p.

2.1 Quelques généralisations du théorème de Christol

Dans ce paragraphe, nous rappelons brièvement quelques généralisations du théorème de Christol, en passant d'un corps fini à un corps de caractéristique non nulle, des séries formelles en une variable aux séries formelles à plusieurs variables, des séries formelles classiques aux séries de Hahn généralisées (pour plus de détails, voir aussi [**?**, 13]).

2.1.1 Cas multidimensionnel-Théorème de Salon

Une généralisation remarquable du théorème de Christol a été obtenue par Salon [107, 108]. Elle concerne les séries de Laurent à plusieurs variables à coefficients dans un corps fini.

Définition 2.1.1. *Une série* $f := \sum a(n_1, n_2 \ldots, n_d) T_1^{n_1} T_2^{n_2} \cdots T_d^{n_d}$ *est algébrique sur le corps* $\mathbb{F}_q(T_1, T_2, \ldots, T_d)$ *s'il existe un polynôme non trivial* P*, à coefficients dans* $\mathbb{F}_q(T_1, T_2, \ldots, T_d)$ *tel que* $P(f) = 0$*.*

La généralisation naturelle du p-noyau d'une suite multidimensionnelle $\mathbf{a} := (a(n_1, n_2 \ldots, n_d))$ est donnée par :

$$N_q(\mathbf{a}) = \{a(p^i n_1 + j_1, \ldots, p^i n_d + j_d) \mid i \geq 0, 0 \leq j_m \leq p^i - 1, 1 \leq m \leq d\}.$$

Ainsi, le théorème de Salon s'énonce de la manière suivante.

Théorème 2.1.1 (Salon)**.** *La série* $\sum a_{n_1,n_2,\ldots,n_d} T_1^{n_1} T_2^{n_2} \cdots T_d^{n_d}$ *est algébrique sur* $\mathbb{F}_q(T_1, T_2, \ldots, T_d)$ *si et seulement si le q-noyau de la suite* $\boldsymbol{a} = (a_{n_1,n_2,\ldots,n_d})$ *est fini.*

Notons que le théorème de Salon permet d'obtenir, de façon élémentaire, le résultat de Deligne sur la diagonale d'une série formelle à plusieurs variables, lorsque le corps considéré est fini.

2.1.2 Cas d'un corps quelconque de caractéristique non nulle

Le théorème de Christol est valable dans le contexte d'un corps de base fini. Lorsque le corps de base est un corps quelconque de caractéristique non nulle, une généralisation a été obtenue indépendamment par Sharif et Woodcock dans [112] et Harase dans [72] (on pourra aussi consulter l'article de Denef et Lipshitz [55] et celui d'Allouche [11]).

Théorème 2.1.2 (Sharif et Woodcock, Harase)**.** *Soit* K *un corps de caractéristique non nulle* p*,* $\overline{K}$ *le corps parfait contenant* K *et* $q = p^s$*, où* $s \geq 1$ *un*

entier. La série $\sum a_n T^{-n} \in K((T^{-1}))$ *est algébrique sur* $K(T)$ *si et seulement si l'espace vectoriel engendré sur* $\overline{K}$ *par le q-noyau « modifié »*

$$N_q(\boldsymbol{a}) := \{(a^{1/q^i}_{q^i n+j})_{n\geq 0} \mid i \geq 0,\ 0 \leq j \leq q^i - 1\}$$

est de dimension finie.

On remarque que l'ensemble $N_q(\mathbf{a})$, qui généralise la notion de q-noyau de la suite infinie $\mathbf{a} = (a_n)$, est exactement le q-noyau de $\mathbf{a}$ dans le cas où le corps K est fini. Ainsi, dans ce cas, ce théorème est équivalent au théorème de Christol. Mentionnons que le théorème 2.1.2 a été en fait démontré pour des séries à plusieurs variables ; ce dernier permet ainsi de prouver le théorème de Deligne de façon élémentaire.

Une toute récente description, en termes d'automates finis, de séries formelles à plusieurs variables, à coefficients dans un corps arbitraire de caractéristique non nulle, algébriques, a été fournie par Adamczewski et Bell dans [6].

Théorème 2.1.3 (Adamczewski et Bell)**.** *Soient* K *un corps arbitraire de caractéristique non nulle et*

$$f := \sum a(n_1, n_2 \ldots, n_d) T_1^{n_1} T_2^{n_2} \cdots T_d^{n_d} \in \mathbb{K}[[T_1, T_2, \ldots, T_d]]$$

algébrique sur le corps $\mathbb{F}_q(T_1, T_2, \ldots, T_d)$*. Alors l'ensemble*

$$Z(f) = \{(n_1, n_2, \ldots, n_d) \in \mathbb{N}^d \text{ tel que } a(n_1, n_2, \ldots, n_d) = 0\}$$

est p-automatique.

Notons que ce résultat donne une généralisation du théorème de Christol, mais également de celui de Derksen [56] sur l'analogue du théorème de Skolem-Mahler-Lech.

2.1.3 Cas des séries généralisées-Théorème de Kedlaya

Le théorème de Christol donne une description concrète des éléments de $\mathbb{F}_q((T^{-1}))$ qui sont algébriques sur $\mathbb{F}_q(T)$; il montre, comme nous l'avons vu, qu'une série est algébrique sur $\mathbb{F}_q(T)$ si et seulement si la suite de ses coefficients est q-automatique. Puisque le corps $\mathbb{F}_q((T^{-1}))$ est loin d'être algébriquement clos, le théorème de Christol n'est pas entièrement satisfaisant, n'offrant qu'une description incomplète des éléments algébriques sur $\mathbb{F}_q(t)$. En effet, il existe des polynômes à coefficients dans $\mathbb{F}_q(T)$ qui n'ont aucune racine dans le corps des séries formelles $\mathbb{F}_q((T^{-1}))$. Par exemple, le polynôme d'Artin–Schreier

$$P(X) = X^p - X - T$$

n'a pas de racine dans le corps $\mathbb{F}_q((T^{-1}))$. Ses racines sont les séries de la forme

$$x = c + T^{1/p} + T^{1/p^2} + \cdots \text{ pour } c = 0, 1, 2, \ldots, p-1.$$

Elles font partie d'un corps bien plus gros, le corps des séries formelles généralisées à coefficients dans $\mathbb{F}_q$ (introduites par Hahn [71] en 1907), que l'on note $\mathbb{F}_q((t^{\mathbb{Q}}))$. Rappelons ci-dessous la définition des séries formelles généralisées.

Un groupe abélien G est dit *totalement ordonné* s'il existe une relation binaire « $>$ » telle que, pour tous $a, b, c \in G$:

$$a \not> a;$$

$$a \not> b, b \not> a \Rightarrow a = b;$$

$$a > b, b > c \Rightarrow a > c;$$

$$a > b \Leftrightarrow a + c > b + c.$$

Un sous-ensemble S de G est *bien ordonné* si tout sous-ensemble de S a un plus petit élément. Ceci est équivalent à dire qu'il n'existe pas de suite infinie décroissante dans S. Soit R un anneau commutatif et G un groupe abélien totalement ordonné. On note $R((T^G))$ l'ensemble de tous les éléments de la forme $f = \sum_{\alpha \in G} r_\alpha T^\alpha$ qui vérifient :

- $r_\alpha \in G$, pour tout $\alpha \in G$,
- le support de f, c'est-à-dire, l'ensemble $\{\alpha \,/ r_\alpha \neq 0\}$, est bien ordonné.

On définit les lois « $+$ » et « $\times$ » de la manière suivante :

$$\sum_{\alpha \in G} r_\alpha T^\alpha + \sum_{\alpha \in G} s_\alpha T^\alpha = \sum_{\alpha \in G} (r_\alpha + s_\alpha) T^\alpha$$

$$\sum_{\alpha \in G} r_\alpha T^\alpha \times \sum_{\alpha \in G} s_\alpha T^\alpha = \sum_{\alpha \in G} \sum_{\beta \in G} (r_\beta s_{\alpha - \beta}) T^\alpha.$$

Dans ce cas, l'ensemble $R((T^G))$, muni des deux lois définies précédemment, constitue un anneau qu'on appelle *l'anneau des séries formelles généralisées de R à exposants dans G* ou *l'anneau des séries de Hahn–Mal'cev–Neumann à coefficients dans R et à exposants dans G*.

Un élément non nul est une unité de l'anneau si et seulement si son premier coefficient est non nul. En particulier, si R est un corps, alors $R((T^G))$ est aussi un corps. De plus, si K est un corps algébriquement clos et G est un groupe divisible, alors $K((T^G))$ est algébriquement clos [74].

Dans le cas où K est le corps fini à q éléments et G est le groupe divisible des rationnels $\mathbb{Q}$, on obtient la suite d'inclusions suivante :

$$\mathbb{F}_q(T) \subset \mathbb{F}_q((1/T)) \subset \mathbb{F}_q((T^{\mathbb{Q}})).$$

Le corps $\mathbb{F}_q((T^{\mathbb{Q}}))$ n'est pas algébriquement clos ; en revanche, le corps

$$(\bigcup_{n \geq 1} \mathbb{F}_{p^n})((T^{\mathbb{Q}}))$$

l'est, puisque $\bigcup_{n\geq 1}\mathbb{F}_{p^n}$ fournit une clôture algébrique de $\mathbb{F}_p$.
Notons que $\mathbb{F}_q((T^{\mathbb{Q}}))$ est bien une généralisation du corps des séries de Laurent à coefficients dans $\mathbb{F}_q$, ce dernier étant le corps des séries de Hahn à support dans $\mathbb{Z}$.

Kedlaya généralise dans [75] le théorème de Christol, en se plaçant dans le corps des séries formelles de Hahn pour déterminer la clôture algébrique de $\mathbb{F}_p(T)$. Pour cela, il définit la notion de quasi-p-automaticité. Une série formelle généralisée $\sum_{i\in I} x_i T^i$ est *quasi-p-automatique* si elle est telle que, si on effectue une transformation affine rationnelle sur les exposants de T, leur dénominateur devient des puissances de p et les coefficients peuvent alors être calculés par un p-automate fini. Le résultat de Kedlaya s'énonce alors comme suit.

Théorème 2.1.4 (Kedlaya)**.** *Soit $f(T)=\sum_{i\in I} x_i T^i \in \mathbb{F}_q((T^{\mathbb{Q}}))$. Alors $f(T)$ est une série algébrique sur le corps $\mathbb{F}_q(T)$ si et seulement si f est quasi-p-automatique.*

Il est à noter que le sens « automatique $\Rightarrow$ algébrique » est le sens le plus facile de la preuve du théorème de Kedlaya ; il utilise uniquement des propriétés élémentaires des automates. Quant au sens « algébrique $\Rightarrow$ automatique », il est beaucoup plus complexe et fait appel à des outils puissants concernant les corps valués (en particulier la théorie des polygones de Newton constitue un ingrédient principal de cette preuve).

2.2 Quelques conséquences du théorème de Christol

Dans ce paragraphe, nous discuterons l'utilisation du théorème de Christol pour démontrer des résultats de transcendance en caractéristique non nulle, mais aussi en caractéristique nulle ; autrement dit, en utilisant quelques propriétés des suites automatiques, il est possible de démontrer la transcendance de certaines séries de Laurent, comme par exemple les analogues de Carlitz des périodes classiques (en particulier π) ou des valeurs des fonctions ζ de Riemann.

Il est à noter, par ailleurs, que des résultats d'automaticité peuvent aussi être obtenus de façon élémentaire, en utilisant des résultats de théorie des nombres. En particulier, nous pouvons regarder l'exemple suivant.

Exemple 2.2.1. Considérons deux séries

$$f(T)=\sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_p((T^{-1}))$$

et

$$g(T)=\sum_{n\geq 0} b_n T^{-n} \in \mathbb{F}_p((T^{-1}))$$

algébriques sur $\mathbb{F}_p(T)$.

Alors, si l'on considère le produit $f(T)g(T) = \sum_{n\geq 0} c_n T^{-n}$, où la suite $(c_n)_{n\geq 0}$ est définie, pour $n \geq 0$, par

$$c_n = \sum_{k=0}^{k=n} a_k b_{n-k} \mod p,$$

il est bien connu que le produit $f(T)g(T)$ est aussi une série algébrique sur $\mathbb{F}_p(T)$. Par conséquent, la suite $(c_n)_{n\geq 0}$ est une suite p-automatique ce qui n' était pas évident à partir de la définition.

Contrairement au produit de Cauchy de deux séries formelles, pour lequel on peut obtenir directement l'algébricité, le cas du produit d'Hadamard est plus difficile. Rappelons d'abord la définition du produit de Hadamard.

Définition 2.2.1. *Le produit d'Hadamard de deux séries formelles* $f(T) = \sum_{n\geq 0} a_n T^{-n}$ *et* $g(T) = \sum_{n\geq 0} b_n T^{-n}$ *est défini par* $f(T)\odot g(T) = \sum_{n\geq 0} a_n b_n T^{-n}$.

Le résultat suivant peut être obtenu rapidement, en passant par le théorème de Christol.

Théorème 2.2.1. *Le produit d'Hadamard de deux séries de Laurent à coefficients dans un corps fini qui sont algébriques est aussi algébrique.*

Le produit d'Hadamard a été introduit dans [70] et est intimement lié aux diagonales des séries formelles à plusieurs variables. En réalité, la première preuve de ce théorème fut donnée par Furstenberg [66] mais il peut être aussi vu comme un corollaire immédiat du théorème de Christol.

Ce théorème est aussi vrai sur un corps quelconque de caractéristique non nulle, en le voyant comme une conséquence du théorème de Sharif et Woodcock. On remarque, par ailleurs, qu'il est faux en caractéristique 0. Ceci peut être argumenté en considérant l'exemple suivant :

$$f(T) = \sum_{n\geq 0} \binom{2n}{n} T^{-n} = (1 - 4T^{-1})^{-1/2}.$$

Il est clair que f est algébrique sur tout corps K et il est possible de montrer que le produit d'Hadamard $f \odot f$ est transcendant sur $\mathbb{Q}(T)$ (voir à ce sujet [17, 63, 113, 116].

2.2.1 Problème de changement de caractéristique

Nous nous concentrons maintenant sur une autre conséquence du théorème de Christol, qui, de plus, utilise un résultat important de Cobham [49].

Théorème 2.2.2 (Cobham)**.** *Soit* $\boldsymbol{a} = (a_n)_{n\geq 0}$ *une suite à valeurs dans un alphabet fini. Soient* k *et* l *deux entiers multiplicativement indépendants (c'est-à-dire* $\log k/\log l$ *est irrationnel). Alors* $\boldsymbol{a}$ *est* k *et* l*-automatique si et seulement si* $\boldsymbol{a}$ *est ultimement périodique.*

Remarquons que ce théorème peut également fournir des séries transcendantes sur $\mathbb{F}_p(T)$. Par exemple, nous considérons la série formelle dont la suite des coefficients est la suite de Thue-Morse :

$$f(T) = \sum_{n\geq 0} t_n T^{-n}.$$

Comme il a été vu précédemment, cette série, considérée à coefficients dans $\mathbb{F}_2$, est algébrique sur $\mathbb{F}_2(T)$. On peut se demander alors si, considérée à coefficients dans $\mathbb{F}_3$, elle est toujours algébrique sur $\mathbb{F}_3(T)$. La réponse est non, grâce aux théorèmes de Christol et Cobham. En effet, si la série était algébrique sur $\mathbb{F}_3(T)$ alors la suite $\mathbf{t} = (t_n)_{n\geq 0}$ serait 3-automatique. Mais, il est bien connu que $\mathbf{t}$ est 2-automatique. Puisque 2 et 3 sont multiplicativement indépendants, alors, d'après le théorème de Cobham, la suite $\mathbf{t}$ serait ultimement périodique, ce qui n'est pas le cas.

Plus généralement, on peut en déduire le résultat suivant de Christol, Kamae, Mendès France et Rauzy [47].

Théorème 2.2.3 (CKRM)**.** *Soit $\boldsymbol{a} = (a_n)_{n\geq 0}$ une suite à valeurs dans $\{0,1\}$ et soient p et q deux nombres premiers différents. Alors les deux séries de Laurent*

$$f(T) = \sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_p((T^{-1})) \text{ et } g(T) = \sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$$

sont algébriques (respectivement sur le corps $\mathbb{F}_p(T)$ et $\mathbb{F}_q(T)$) si et seulement si ce sont des fractions rationnelles.

Autrement dit, si une série irrationnelle est algébrique sur $\mathbb{F}_{p_1}(T)$, alors elle est transcendante sur $\mathbb{F}_{p_2}(T)$, si p_1 et p_2 sont deux nombres premiers distincts. Dans le cas des séries de Hahn, Adamczewski et Bell [2] ont généralisé ce résultat, afin d'avoir une caractérisation globale de tous les éléments algébriques sur $\mathbb{F}_p(T)$.

Théorème 2.2.4 (Adamczewski et Bell)**.** *Soit $h : \mathbb{Q} \mapsto \{0,1\}$ une fonction dont le support est bien ordonné. Soient p et q deux nombres premiers différents. Alors, les deux séries de Hahn*

$$f(T) = \sum_{\alpha\in\mathbb{Q}} h(\alpha) T^{\alpha} \in \mathbb{F}_p((T^{\mathbb{Q}})) \text{ et } g(T) = \sum_{\alpha\in\mathbb{Q}} h(\alpha) T^{\alpha} \in \mathbb{F}_q((T^{\mathbb{Q}}))$$

sont algébriques (respectivement sur le corps $\mathbb{F}_p(T)$ et $\mathbb{F}_q(T)$) si, et seulement s'il existe un entier $n \geq 1$ tel que $f(T^n)$ et $g(T^n)$ soient deux fractions rationnelles.

Le théorème 2.2.3 est intimement lié à une question très difficile, concernant le développement de nombres réels dans deux bases multiplicativement indépendantes : *étant donnée une suite $\boldsymbol{a} = (a_n)_{n\geq 0} \in \{0,1\}^{\infty}$, apériodique, peut-on montrer qu'au moins un de deux nombres $\sum_{n\geq 0} a_n/2^n$ et $\sum_{n\geq 0} a_n/3^n$ est transcendant* ? Malheureusement, le développement binaire d'un nombre réel n'offre aucune information concernant son développement en base 3. Ainsi, cette question, qui est en fait attribuée à Mahler par Mendès France, semble très loin d'être résolue (voir [3]).

2.2.2 Les analogues de Carlitz

Dans cette partie, nous rappelons quelques fonctions définies par analogie avec certaines constantes réeles comme le réel π ou certaines valeurs de la fonction ζ de Riemann. Nous rappelons également quelques résultats de transcendance les concernant : en particulier, le théoréme de Christol permet de montrer la transcendance de certaines de ces fonctions.

En 1935, Carlitz [38] a construit en caractéristique non nulle une fonction e_C par analogie avec la fonction exponentielle définie pour les nombres réels :

$$e_C(0) = 0,\ d/dz(e_C(z)) = 1 \text{ et } e_C(Tz) = Te_C(z) + e_C(z)^q.$$

Celle-ci est appelée l'exponentielle de Carlitz et l'action $u \to Tu + u^q$ conduit à la définition du module de Carlitz, qui est en fait un cas particulier d'un module de Drinfeld. La fonction exponentielle de Carlitz $e_C(z)$ peut être définie par le produit infini suivant :

$$e_C(z) = z \prod_{a \in \mathbb{F}_q[T],\, a \neq 0} (1 - \frac{z}{a\widetilde{\Pi}_q})$$

où

$$\widetilde{\Pi}_q = (-T)^{\frac{q}{q-1}} \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j-1}}\right)^{-1}.$$

Puisque $e^z = 1$ si et seulement si $z \in 2i\pi\mathbb{Z}$ et puisque $e_C(z)$ a été construit par analogie tel que $e_C(z) = 0$ si et seulement si $z \in \widetilde{\Pi}_q \mathbb{F}_q[T]$ alors, si l'on choisit

$$\Pi_q = \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j-1}}\right)^{-1},$$

Π_q est considéré comme l'analogue du nombre réel π.

On rappelle aussi que la fonction ζ de Riemann est définie pour tous les nombres complexes s, $\Re(s) > 1$, de la manière suivante :

$$\zeta(s) = \sum_{n \geq} \frac{1}{n^s}.$$

Par analogie, la fonction ζ_q de Carlitz est définie pour $n \geq 1$ par :

$$\zeta_q : \mathbb{N}^* \to \mathbb{F}_q\left[\left[\frac{1}{T}\right]\right]; \qquad \zeta_q(n) = \sum_{\substack{\mathrm{P} \in \mathbb{F}_q[T] \\ \mathrm{P\ unitaire}}} \frac{1}{P^n}. \tag{2.1}$$

La fonction ζ_q de Carlitz a de propriétés analogues à la fonction usuelle ζ. En particulier, Carlitz a démontré que $\zeta_q(n)/\Pi_q^n \in \mathbb{F}_q(T)$, pour tout n multiple de $q-1$. Cette propriété est l'analogue dans le cas des séries formelles du résultat d'Euler pour les entiers pairs. On note, par ailleurs, que le groupe des unités de $\mathbb{F}_q(T)$ est $\mathbb{F}_q^*$ et que son cardinal est $q-1$, alors que le groupe multiplicatif

de $\mathbb{Z}$ est $\{+1, -1\}$ avec pour cardinal 2. Ceci est la raison pour laquelle les entiers divisibles par $q-1$ sont considérés comme les entiers « pairs ».

Puisque le nombre réel π est transcendant, les valeurs aux entiers pairs de ζ sont aussi transcendantes ; de même, les nombres $\zeta(2n)/\pi^{2n}$ sont rationnels. Il est donc naturel de se demander si la série formelle Π_q ou les valeurs de la fonction ζ_q sont, elles aussi, transcendantes. A ce propos, plusieurs méthodes sont connues pour montrer la transcendance de ces séries formelles : la première, utilisée par Wade dans les année quarante, ressemble à une méthode classique de transcendance de nombres réels sur le corps des rationnels [135, 136, 137]. Ensuite, l'utilisation des modules de Drinfeld [133] donne des résultats plus complets (les valeurs $\zeta_q(n)$ sont transcendantes, pour tout $n \in \mathbb{N}^*$ et $\zeta_q(n)/\Pi_q^n$ est transcendant pour tout n non multiple de $q-1$). Enfin, la méthode d'approximation diophantienne, exposée par De Mathan et Chérif [44], qui donne des mesures d'irrationalité pour différentes valeurs de ζ_q ; et la dernière méthode, considérée comme la plus « élémentaire », la preuve appelée « automatique ». Elle utilise principalement le théorème de Christol et des propriétés des automates finis ou des suites automatiques. Le premier résultat obtenu par cette méthode a été la transcendance de la série formelle Π_q, en utilisant le développement en série formelle de $1/\Pi_q$, donné explicitement par Allouche [12], et en particulier l'infinitude du q-noyau de cette suite.

D'autres résultats utilisant la méthode automatique ont été obtenus par Berthé [30] et concernent la transcendance de toute combinaison linéaire, à coefficients dans $\mathbb{F}_q(T)$, de $\zeta_q(n)/\Pi_q^n$, pour $1 \leq n \leq q-2$ et $q \neq 2$. Notons que ce résultat a été démontré précédemment par Yu [133] pour tout n non multiple de $q-1$. En particulier, cela implique la transcendance de $\zeta_q(n)/\Pi_q^n$ pour $1 \leq n \leq q-2$ et $q \neq 2$.

Dans le même esprit, des résultats de transcendance ont été obtenus par Berthé concernant les valeurs du logarithme de Carlitz [29] ou bien par Thakur [121] pour des valeurs de la fonction Gamma de Carlitz.

Mentionnons également que des résultats importants ont été obtenus par Papanikolas pour l'analogue de Carlitz de la fonction logarithme. Plus précisément, l'auteur a démontré dans [102] l'analogue de la conjecture qui stipule que les nombres $\pi, \log 2, \log 3, \ldots$ sont algébriquement indépendants ; ce résultat est une conséquence d'une variante de la conjecture de Grothendieck dans le corps de fonctions en caractéristique non nulle, démontrée par le même auteur. Une autre application de cette conjecture est donnée par Chang et Yu [43] qui ont décrit les relations algébriques entre les valeurs aux entiers strictement positifs de la fonction ζ_q de Carlitz. Pour une référence plus détaillée, le lecteur peut consulter [103].

2.2.3 Application du théoréme de Christol à d'autres corps

Nous venons de rappeler des résultats de transcendance dans le cas de séries formelles en caractéristique non nulle. Cependant, le théorème de Christol

a aussi des conséquences intéressantes dans le cadre des séries formelles en caractéristique 0. En effet, si une série formelle $f \in \mathbb{Q}((T^{-1}))$ est algébrique sur $\mathbb{Q}(T)$, alors sa réduction modulo tout nombre premier p est aussi algébrique sur $\mathbb{F}_p(T)$. A l'aide de cette observation, on peut, par exemple, donner une preuve élémentaire de la transcendance sur $\mathbb{Q}(T)$ de la série classique

$$\Theta_3(T) = \sum_{-\infty<n<\infty} T^{-n^2};$$

il suffit de réduire cette série modulo 3 et de démontrer que la suite caractéristique des carrés n'est pas 3-automatique. Cela peut se faire en étudiant la complexité des facteurs de cette suite en prouvant qu'elle est quadratique. D'après le théorème 1.3.2, la suite n'est pas 3-automatique.

Concluons cette partie en accentuant à nouveau la différence entre les représentations des nombres réels algébriques et celles des séries formelles algébriques à coefficients dans un corps fini. En combinant le théorème de Christol avec le théorème 2.0.4, le résultat suivant peut être déduit.

Théorème 2.2.5 (Adamczewski et Bugeaud)**.** *Soit* $\boldsymbol{a} := (a_n)_{n\geq 0}$ *une suite apériodique à valeurs dans* $\mathbb{F}_p$*. Si la série formelle* $f_{\boldsymbol{a}}(T) = \sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_p((T^{-1}))$ *est algébrique sur* $\mathbb{F}_p(T)$*, alors le nombre réel* $\xi_{\mathbf{a}} = \sum_{n\geq 0} a_n p^{-n}$ *est transcendant. Réciproquement, si* $\xi_{\mathbf{a}}$ *est un nombre réel algébrique, alors la série* $f_{\boldsymbol{a}}(T)$ *est transcendante.*

3

Approximation diophantienne en caractéristique non nulle

L'approximation diophantienne est un thème classique de la théorie des nombres qui, sous sa forme primitive, a pour but de répondre à la question suivante : dans quelle mesure peut-on approcher les nombres irrationnels par des nombres rationnels ? En particulier, étant donné un nombre réel ξ, comment peut-on construire des nombres rationnels p/q approchant ξ et tels que l'écart $|\xi - p/q|$ soit majoré par une puissance de q ?

A cet effet, *l'exposant (ou la mesure) d'irrationalité* a été défini comme étant le supremum des nombres réels μ pour lesquels l'inégalité

$$\left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^{\mu}}$$

possède une infinité de solutions p/q.

Une première réponse dans cette direction a été apportée par Dirichlet [58] en 1842. En utilisant le principe des tiroirs, il a montré que, pour tout nombre réel irrationnel ξ, il existe une infinité de nombres rationnels p/q avec $|\xi - p/q| < q^{-2}$.

Deux années plus tard, Liouville [86] s'intéresse à un ensemble particulier de nombres réels, à savoir celui des nombres algébriques et démontre le résultat suivant.

Théorème 3.0.6 (Liouville)**.** *Soit ξ un nombre réel algébrique de degré $q \geq 2$.*

Alors, il existe une constante réelle $c(\xi)$, telle que

$$\left|\xi - \frac{p}{q}\right| \geq \frac{c}{q^d},$$

pour tout nombre rationnel p/q avec $q \geq 1$.

Ainsi, nous remarquons que la mesure d'irrationalité d'un nombre algébrique irrationnel ξ vérifie $2 \leq \mu(\xi) \leq \deg(\xi)$. Une première amélioration de ce résultat a été faite par Thue [128] en 1909, qui a montré que $2 \leq \mu(\xi) \leq \deg(\xi)/2$. Une application relativement directe de ce résultat concerne l'étude des équations diophantiennes : si $f(x, y)$ est un polynôme homogène, irréductible, à coefficients entiers, de degré supérieur ou égal à 3, alors l'équation

$$f(x, y) = m$$

a un nombre fini de solutions entières x, y (voir [128, 138]). Par la suite, d'autres améliorations ont été apportées par Siegel [114] en 1921 et Dyson [60] en 1947 ; elles culminent avec le célèbre résultat de Roth [106].

Théorème 3.0.7 (Roth)**.** *Soit ξ un nombre algébrique irrationnel. Pour tout $\varepsilon > 0$, il existe une constante réelle positive $c(\xi, \varepsilon)$ telle que l'inégalité*

$$\left|\xi - \frac{p}{q}\right| > \frac{c(\xi, \varepsilon)}{q^{2+\varepsilon}}$$

soit vraie pour tout nombre rationnel p/q, avec $q > 0$.

La caractéristique commune à tous les résultats obtenus par Thue, Siegel, Dyson et Roth est qu'ils sont ineffectifs, c'est-à-dire que la constante $c(\xi, \varepsilon)$ ne peut pas être calculée explicitement. Ceci implique, en particulier, qu'on ne peut pas majorer la taille des solutions des équations diophantiennes, l'un des enjeux majeurs de la théorie de l'approximation diophantienne.

Plusieurs extensions et généralisations du théorème de Roth ont été ensuite obtenues. En particulier, Schmidt a établi dans [109] un résultat remarquable, connu sous le nom du théorème du sous-espace et s'exprimant en terme d'approximation simultanée des formes linéaires à coefficients algébriques. Il s'agit d'un résultat très profond qui a donné lieu à de nombreuses applications intéressantes. A ce sujet, on peut consulter l'article de survol de Bilu [31] correspondant à son l'exposé au séminaire Bourbaki. Nous mentionnons en particulier les résultats de Corvaja et Zannier [51, 53] sur les équations diophantiennes du type $P(u(n), y) = 0$, $u(n)$ étant une somme de la forme $b_1 a_1^n + \cdots + b_m a_m^n$ ou bien leur travaux autour du théorème de Siegel sur les points entiers sur les courbes [52]. D'autres applications intéressantes, concernant notamment le développement décimal des nombres algébriques, sont dues à Adamczewski et Bugeaud [4].

Des questions similaires d'approximation diophantienne peuvent être étudiées en remplaçant les entiers par les polynômes, les nombres rationnels par

les fractions rationnelles, ou les nombres algébriques par les séries de Laurent algébriques. Comme il a été observé par Mahler [91], les théorèmes de Dirichlet et Liouville ont des analogues dans le corps des séries de Laurent : pour toute série de Laurent irrationnelle f à coefficients dans un corps K, algébrique sur $K(T)$ on a : $2 \leq \mu(f) \leq \deg(f)$, où la mesure d'irrationalité est définie comme dans le cas réel, mais les rationnels p et q sont remplacés par des polynômes à coefficients dans K.

De plus, Uchiyama [131] a démontré que le théorème de Roth est aussi vrai dans le cas où K est un corps de caractéristique 0 ; cependant, ceci n'est plus vrai dans le corps des fonctions en caractéristique non nulle. En effet, Mahler a été le premier a avoir observé que le théorème de Liouville est optimal. Il a donné l'exemple, probablement, le plus simple d'une série algébrique de degré q dont l'exposant d'irrationalité est aussi égal à q : il s'agit de $f(T) = \sum_{n\geq 0} T^{-q^n} \in \mathbb{F}_q((T^{-1}))$ vérifiant l'équation $f^q - f + T^{-1} = 0$.

Au début, Mahler a suggéré que les séries dont les mesures d'irrationalité atteignent la borne de Liouville doivent satisfaire une certaine condition sur leur degré, comme par exemple la divisibilité du degré par la caractéristique. Des années plus tard, Osgood [99] a donné des exemples de séries formelles de degré arbitraire, dont les exposants d'irrationalité atteignent la borne de Liouville. Ceci nous amène à nous demander, par exemple, comment les mesures d'irrationalité des séries algébriques sont distribuées ; quelles sont les séries qui atteignent la borne de Liouville ou quelles sont les séries qui vérifient ou qui ne vérifient pas le théorème de Roth ? Les questions que nous venons d'évoquer, ainsi que de nombreuses autres liées à ce sujet, ne sont toujours pas résolues (voir quelques commentaires à ce propos dans [126]).

3.1 Le développement en fraction continue

Nous allons nous intéresser, à présent, aux questions d'approximation diophantienne dans le corps des séries de Laurent définies sur un corps de caractéristique p. Ce domaine s'est beaucoup développé au fil des ans, notamment avec les travaux de Baum et Sweet, Mills, Robbins, Schmidt, Osgood, De Mathan, Lasjaunias, Thakur... . Evidemment, la théorie des fractions continues intervient naturellement, puisque les meilleures approximations rationnelles sont obtenues en tronquant le développement en fraction continue. Malheureusement, trouver le développement en fraction continue d'une série formelle est souvent une question difficile ; c'est à cet effet que plusieurs algorithmes et méthodes ont été développés.

Soit K un corps de caractéristique p. Nous allons utiliser les notations de Cassels [42] pour rappeler quelques notions de base des fractions continues dans ce cadre. Pour une série de Laurent

$$f(T) = a_{i_0} T^{-i_0} + a_{i_0+1} T^{-i_0-1} + \cdots ,$$

où $i_0 \in \mathbb{Z}, a_i \in K, a_{i_0} \neq 0$, la valeur absolue ultramétrique est définie par $|f| = |T|^{-i_0}$, où $|T|$ est un nombre réel positif strictement supérieur à 1 (dans la littérature, il existe plusieurs notations : il est parfois considéré égal à e, p ou 2).

La partie entière $[f]$ de f est définie par

$$[f] = \sum_{i=i_0}^{i=0} a_i T^{-i},$$

et on fixe $||f|| = |f - [f]|$. On définit une suite des meilleures approximations P_n/Q_n de f de la façon suivante :

$$Q_0 = 1,$$

$$|Q_n| < |Q_{n+1}|, \text{ c'est-à-dire, } \deg Q_n < \deg Q_{n+1},$$
$$|Q_{n+1}f - P_{n+1}| < |Q_n f - P_n| < 1,$$
$$|Q_n f - P_n| \leq |Qf - P|, \text{ pour } |Q_n| \leq |Q| < |Q_{n+1}|.$$

La théorie des fractions continues fournit un algorithme pour calculer P_n/Q_n. De plus, il est bien connu que :

$$P_n Q_{n+1} - P_{n+1} Q_n = (-1)^{n+1} \text{ et}$$
$$|Q_n f - P_n| = |Q_{n+1}|^{-1}.$$

Par conséquent, il résulte de la première égalité que $\text{pgcd}(P_{n+1}, Q_{n+1}) = 1$ et que

$$(P_{n+2} - P_n)Q_{n+1} = (Q_{n+2} - Q_n)P_{n+1}.$$

Il existe alors des polynômes $a_{n+2} \in K(T)$ tels que, pour tout $n \geq 0$, on a :

$$P_{n+2} - P_n = a_{n+2} P_{n+1},$$
$$Q_{n+2} - Q_n = a_{n+2} Q_{n+1}.$$

Pour que ce résultat reste toujours vrai pour $n = -1, -2$, on fixe

$$Q_{-1} = P_{-2} = 0, Q_{-2} = P_{-1} = 1, a_0 = P_0, a_1 = Q_1.$$

Si on note

$$\delta_n = Q_n f - P_n$$
$$\theta_n = -\delta_{n-2}/\delta_{n-1},$$

alors $a_n = \theta_n - \theta_{n+1}^{-1}$. En fait, $|\theta_n| > 1$ et alors, pour tout $n \geq 0$, $a_n = [\theta_n]$. Ainsi, la série de Laurent f peut s'écrire comme :

$$f = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}},$$

et

$$P_n = a_n P_{n-1} + P_{n-2} \text{ et } Q_n = a_n Q_{n-1} + Q_{n-2}.$$

De façon plus courte, nous notons $f = [a_0, a_1, a_2, \ldots]$. Ceci est connu comme le développement en fraction continue de f (ces notions sont analogues avec le cas réel et le cas de séries formelles sur des corps arbitraires). Les fractions rationnelles P_n/Q_n, pour $n \geq 0$, désignent *les convergents* de f et les polynômes a_n *les quotients partiels* ; la relation suivante relie les deux notions :

$$P_n/Q_n = [a_0; a_1, \ldots, a_n].$$

Grâce à la valeur absolue ultramétrique, on a :

$$|f - P_n/Q_n| = |P_{n+1}/Q_{n+1} - P_n/Q_n| = |Q_n Q_{n+1}|^{-1} = |a_{n+1}|^{-1}|Q_n|^{-2}. \quad (3.1)$$

Par ailleurs, il est intéressant de noter que si $P, Q \in K[T]$ et $Q \neq 0$ alors P/Q est un convergent de f si et seulement si

$$|f - P/Q| < |Q|^{-2}.$$

Notons que f est rationnelle si et seulement si la suite de ses quotients partiels est finie. En revanche, pour les développements en fraction continue ultimement périodiques on dispose du résultat suivant, dont l'analogue dans le cas réel est le théorème clasique d'Euler–Lagrange.

Théorème 3.1.1. *Soit K un corps fini et $f \in K((T^{-1}))$ irrationnelle. Alors la suite des quotients partiels du développement en fraction continue de f est ultimement périodique si et seulement si f est quadratique sur $K(T)$.*

Outre les quadratiques réels, on ne connaît aucun développement en fraction continue d'un nombre réel algébrique de degré supérieur à 2 ; de plus, nous ne savons pas non plus si la suite des quotients partiels est bornée ou non. Il est très difficile d'obtenir des informations sur le développement en fraction continue des nombres algébriques, puisque les actions des opérations usuelles comme l'addition ou la multiplication ne sont pas du tout transparentes sur le développement en fraction continue. En général, on conjecture que la suite des quotients partiels de nombres algébriques de degré superieur ou égal à 3 est non bornée ; cette question, suggerée par Khintchine, constitue un problème majeur en approximation diophantienne.

En revanche, en caractéristique p, nos connaissances sont sensiblement plus satisfaisantes (bien qu'on ne dispose pas d'analogue au théorème de Roth) : on connaît des séries algébriques de degré au moins 3 dont le développement en fraction continue est connu explicitement ; ceci est dû en partie au fait que le morphisme de Frobenius agit de manière très transparente sur le développement en fraction continue : en effet, si $f = [a_0; a_1, a_2, \ldots]$ alors $f^p = [a_0^p; a_1^p, a_2^p, \ldots]$. Avoir des informations sur le développement en fraction

continues permet clairement d'obtenir des résultats très fins concernant l'approximation rationnelle [35, 79, 78]. En effet, d'après l'égalité (3.1), l'exposant d'irrationalité peut aussi être calculé grâce à la formule suivante :

$$\mu(f) = 2 + \limsup_k \frac{\deg a_{k+1}}{\deg Q_k}.$$

Par ailleurs, par définition, on a :

$$\deg Q_k = \sum_{1 \le i \le k} \deg a_i,$$

ce qui implique que l'exposant d'irrationalité est directement lié à la croissance de la suite des degrés des quotients partiels. En particulier, si la suite $(\deg a_i)_{i \ge 1}$ est bornée, alors $\mu(f) = 2$; cependant, la réciproque est fausse car il est possible de fournir des exemples concrets de séries dont l'exposant est égal à 2 et dont la suite des degrés des quotients partiels est non bornée. Toujours d'après l'égalité (3.1) et d'après l'analogue du théorème de Liouville pour les séries algébriques, on peut en déduire que, si $\deg(f) \ge 2$, alors on a

$$\mu(f) \in [2, \deg(f)].$$

En s'inspirant des exemples de Mahler et d'Osgood, Voloch a donné plusieurs exemples de séries formelles de degré d (n'utilisant pas les fractions continues) en montrant que l'exposant d'irrationalité peut être égal à tout nombre rationnel compris entre $1 + \sqrt{q}$ et d, où d prend des valeurs quelconques. Ensuite, en utilisant des fractions continues, Thakur [123] a construit des exemples de séries formelles de différentes degrés pour lesquelles les exposants d'irrationalité sont des nombres rationnels. De plus, il a montré que pour tout rationnel $\nu \ge 2$, il existe une série algébrique f telle que $\mu(f) = \nu$. Ce théorème a été indépendamment démontré par Schmidt [110]. Notons qu'avec ce résultat, il semble alors raisonnable de penser que l'ensemble des exposants d'irrationalité des séries algébriques de degré fixé d est en fait exactement l'ensemble des nombres rationnels compris entre 2 et d.

3.2 Une classe spéciale de séries algébriques

En 1976, Baum et Sweet ont donné dans [24] le premier exemple d'une série formelle algébrique de degré 3 dont la suite des quotients partiels ne prend qu'un nombre fini de valeurs. C'est d'ailleurs l'exemple le plus célèbre, avec bien sûr les séries de Laurent quadratiques, qui vérifie le théorème de Roth (voir aussi [25]). Quelques années plus tard, Mills et Robbins ont étudié le développement en fraction continue de cette série ; de plus, ils ont donné d'autres exemples de séries algébriques en explicitant leur développement en fraction continue. Pour cela, ils ont mis en lumière un algorithme qui permet d'expliciter les quotients partiels d'une classe spéciale de séries formelles : la classe des

hyperquadratiques, notée $\mathcal{H}(K) := \bigcup_{s\geq 1} \mathcal{H}_s$, où $\mathcal{H}_s$ est défini par :

$$\mathcal{H}_s(K) = \left\{ f \in K((T^{-1})),\ f = \frac{Af^{p^s} + B}{Cf^{p^s} + D} \right\},$$

et $A, B, C, D \in K(T)$ tels que $AD - BC \neq 0$ et K est un corps (arbitraire) de caractéristique p . Notons que la cubique de Baum et Sweet appartient à cette classe, comme d'ailleurs toutes les séries algébriques sur $K(T)$ de degré inférieur ou égal à 3 (puisque si q est une puissance de p, alors $1, \alpha, \alpha^q, \alpha^{q+1}$ sont linéairement dépendants sur $K(T)$). Toujours dans [95], les auteurs ont exhibé, pour chaque $p \geq 3$, un exemple de série algébrique non-quadratique à coefficients dans $\mathbb{F}_p$ dont les quotients partiels sont tous de degré égal à 1.

Les exemples trouvés par Baum et Sweet et ensuite ceux trouvés par Mills et Robbins sont à l'origine de nombreuses questions. Parallèlement au théorème de Christol, Mendès France a posé la question d'automaticité du développement en fraction continue d'une série algébrique, dont les quotients partiels ne prennent qu'un nombre fini des valeurs. Une réponse positive a été donnée dans [10, 16] pour chaque exemple de [95] en caractéristique strictement supérieure à 2 ; cependant, la méthode employée par ces auteurs ne permettait pas d'étudier l'exemple de Baum et Sweet. Mkaouar [94] (voir aussi [132]) a donné une réponse négative à la question de Mendès France en montrant que la suite des quotients partiels de la série de Baum et Sweet n'est p-automatique pour aucun $p \geq 2$. En revanche, il a demontré que cette suite des quotients partiels est une suite substitutive ; ainsi, il pose la question suivante : *Si f est algébrique sur $\mathbb{F}_p(T)$ et son développement en fraction continue est à quotients partiels bornés, alors ce développement est-il toujours engendré par une substitution sur un alphabet fini* ? On ne connaît pour l'instant pas la réponse à cette question, mais, il est clair que les séries formelles hyperquadratiques ont un intérêt particulier parmi l'ensemble des séries formelles algébriques. Tous les exemples pour lesquels on sait décrire, de façon plus ou moins explicite, le développement en fraction continue sont des séries de Laurent hyperquadratiques et il semble que leurs quotients partiels soient ou bien de degré borné, ou bien de degré très grand (par rapport à leur indice) pour une infinité d'entre eux (voir à ce sujet les travaux de Lasjaunias [32, 81, 82, 78, 80, 83]).

3.3 Le théorème de Thue

Malgré l'existence de l'algorithme de Mills et Robbins, la description complète de tous les éléments hyperquadratiques semble hors d'atteinte. Plusieurs résultats en approximation diophantienne sont connus pour cet ensemble de séries ou pour son complémentaire dans l'ensemble des séries algébriques.

Voloch [134] a demontré que, si f est une série hyperquadratique, à coefficients dans un corps arbitraire de caractéristique non nulle, d'exposant d'irra-

tionalité ν, alors il existe une constante C telle que :

$$\left| f - \frac{P}{Q} \right| \geq \frac{C}{|Q|^{\nu}},$$

pour tous les polynômes $P, Q \in K[T]$, $Q \neq 0$. En d'autres termes les séries hyperquadratiques d'exposant d'irrationalité égal à 2 ont toutes des quotients partiels bornés.

Quelques années plus tard, De Mathan et Lasjaunias ont démontré l'analogue du théorème de Thue pour les séries formelles non-hyperquadratiques [84].

Théorème 3.3.1 (De Mathan et Lasjaunias)**.** *Soit K un corps de caractéristique non nulle et $f \in K((T^{-1}))$ algébrique sur $K(T)$ de degré $n > 1$. Si $f \notin \mathcal{H}(K)$, alors pour tout $\varepsilon > 0$, il existe une constante C telle que :*

$$\left| f - \frac{P}{Q} \right| \geq \frac{C}{|Q|^{[n/2]+1+\varepsilon}},$$

pour tous les polynômes $P, Q \in K[T]$, $Q \neq 0$.

Dans le cas où K est un corps fini, ce résultat peut être légèrement amélioré, en éliminant le réel ε dans l'inégalité précédente. Ceci est l'objet de l'article [85].

Notons qu'un analogue de Thue a été démontré d'abord par Osgood [99, 100], en utilisant la méthode de différentiation (un autre outil puissant en caractéristique non nulle) pour les séries formelles qui ne satisfont pas une équation de Ricatti. Il est connu qu'une série algébrique hyperquadratique satisfait aussi une équation de Ricatti ; ainsi le résultat démontré par De Mathan et Lasjaunias est plus général : il est par ailleurs moins précis puisque, dans la preuve du théorème d'Osgood, la constante C est effective.

Nous renvoyons aussi aux travaux de Thakur [118, 120, 122] où l'auteur construit, en utilisant différentes combinaisons linéaires de type Mahler, des séries de Laurent ne satisfaisant pas d'équation de Ricatti, pour lesquelles les développements en fraction continue sont explicites. La plupart de ces exemples ont des quotients partiels non-bornés ; il y a aussi des constructions qui ont des quotients partiels bornés, et l'auteur fait alors un lien avec la théorie des automates (pour plus de détails à ce sujet voir aussi [124]). Notons que les techniques employées dans [124] ont été précédemment utilisées dans ses travaux sur les développements en fraction continue des analogues de e ou des nombres de Hurwitz $(ae^{2/n} + b)(ce^{2/n} + d)$ dans le module de Carlitz.

4

Aperçu des résultats

4.1 Généralisation de la cubique de Baum et Sweet et fractions continues

Dans le Chapitre 5, nous présentons une étude diophantienne de séries de Laurent algébriques qui généralisent la célèbre cubique introduite par Baum et Sweet dans [24]. Plus précisément, nous obtenons une description explicite du développement en fraction continue des éléments de cette classe de séries de Laurent.

Dans un premier temps, nous observons que l'équation

$$TX^{r+1} + X - T = 0, \tag{4.1}$$

où r est une puissance de p, a une unique solution dans le corps $\mathbb{F}_p((T^{-1}))$. Pour $r = p = 2$, cette solution est exactement la cubique de Baum et Sweet.

Afin de décrire le développement en fraction continue d'une telle solution, nous utilisons une technique qui a déjà été employée par Lasjaunias dans [82] ; celle-ci ressemble, partiellement, à l'algorithme de Mills et Robbins introduit pour décrire le développement en fraction continue de la cubique de Baum et Sweet, puis pour décrire le développement en fraction continue de plusieurs exemples de séries algébriques hyperquadratiques.

Le chapitre 5 est divisé en trois parties. Dans un premier temps, nous énonçons le résultat principal sur lequel la preuve repose (voir le lemme 5.2.2) ;

une application assez directe de ce résultat nous permet de décrire au passage le développement en fraction continue de la série de Mahler

$$\Theta_r(T) := \sum_{r\geq 0} \frac{1}{T^{r^k}} \in \mathbb{F}_p((T^{-1})).$$

Dans la deuxième partie de ce chapitre, nous décrivons de façon plus détaillée l'algorithme permettant, en général, d'obtenir le développement en fraction continue d'une série de Laurent z vérifiant une relation du type

$$Pz_m^r = Qz_n + R,$$

où $z = [a_1, a_2, \ldots, z_m]$, $P, Q, R \in \mathbb{F}_p[T]^3$ et $m < n$ sont deux entiers strictement positifs. Les séries de Laurent qui satisfont à une telle équation sont dites de type (P, Q, R, m, n). Elles ont un intérêt particulier pour la théorie des fractions continues : en effet, dans le cas où P, Q, R sont « bien choisis », il est généralement possible d'expliciter le développement en fraction continue d'une telle série de Laurent. Pour quelques résultats dans cette direction, nous renvoyons le lecteur aux articles [81, 82, 83].

Dans la dernière partie du chapitre 5, nous utilisons cet algorithme afin de décrire le développement en fraction continue des séries de Laurent qui sont solutions de l'équation (4.1). Cette description, donnée dans le théorème 5.3.1, semble à première vue assez complexe, mais elle révèle la présence de motifs intéressants apparaissant dans la suite des quotients partiels de ces séries de Laurent. De tels motifs traduisent une structure combinatoire riche, dotée en particulier d'une symétrie assez surprenante. Cette dernière est traduite par l'occurrence de blocs de quotients partiels « pseudo-palindromiques ».

4.2 Approximation rationnelle des séries de Laurent algébriques à coefficients dans un corps fini

Le Chapitre 6 est consacré à l'étude de l'exposant d'irrationalité des séries de Laurent qui sont algébriques sur le corps des fractions rationnelles $\mathbb{F}_q(T)$. Le point de départ de ce travail est l'article [8] portant sur l'exposant d'irrationalité des nombres réels automatiques.

D'après le théorème de Christol, nous savons que la suite des coefficients d'une série algébrique sur $\mathbb{F}_p(T)$ est p-automatique. En utilisant une approche analogue à celle de [8], nous donnons dans un premier temps une majoration générale de l'exposant d'irrationalité des séries de Laurent algébriques sur $\mathbb{F}_p(T)$. Cette borne dépend de deux paramètres : le cardinal du p-noyau de la suite des coefficients et le nombre d'états de l'automate minimal qui engendre la suite des coefficients (dans le sens direct)[1].

[1]Rappelons que le cardinal du p-noyau d'une suite p-automatique est aussi égal au nombre d'etats de l'automate qui l'engendre, mais dans le sens inverse

Afin de majorer l'exposant d'irrationalité d'une série de Laurent algébrique f, nous cherchons classiquement à construire une suite « dense » de « bonnes » approximations rationnelles. Pour ce faire, l'idée principale consiste à tronquer le développement en série de Laurent de f à certains endroits « bien choisis » à l'aide du théorème de Cobham sur les morphismes uniformes, puis à compléter le développement fini obtenu par périodicité. Ainsi, nous construisons une suite de fractions rationnelles (P_n/Q_n) approchant f et vérifiant certaines conditions décrites dans le lemme 6.2.2, lequel est une variante d'un résultat connu dans ce contexte sous le nom de lemme de Voloch. Notons que le lemme 6.2.2 présente un intérêt indépendant.

Evidemment, la majoration de l'exposant d'irrationalité que nous obtenons par cette approche dépend fortement de la construction de la suite d'approximation (P_n/Q_n). En particulier, la borne générale du théorème 6.1.2 est la plupart du temps bien loin d'être optimale. Dans l'esprit de [9], nous remarquons qu'il est néanmoins souvent possible d'optimiser, dans la pratique, la construction de la suite (P_n/Q_n) afin d'améliorer significativement la majoration donnée par le théorème 6.1.2. Par exemple, le fait de savoir que les polynômes P_n et Q_n sont premiers entre eux, conduit à une amélioration substantielle. Ce problème de coprimalité est bien connu pour être difficile dans le cas de constructions analogues faisant intervenir des nombres réels (voir [9]).

Dans la suite du chapitre 6, nous proposons une nouvelle approche afin de surmonter cette difficulté. Plus exactement, nous décrivons un algorithme qui permet de vérifier si les polynômes P_n et Q_n sont premiers entre eux. Pour cela, nous remarquons que les dénominateurs des approximations construites ont une forme particulière permettant de calculer leurs racines (dans une clôture algébrique de $\mathbb{F}_p$) grâce aux simplifications de calculs induites par l'endomorphisme de Frobenius. En outre, il suffit de vérifier que ces racines n'annulent pas les numérateurs P_n pour garantir la coprimalité des polynômes P_n et Q_n. Dans ce but, nous développons ensuite un calcul des numérateurs. Nous associons pour cela à chaque p-morphisme une matrice de polynômes $M(T)$ et nous détaillons dans la partie 6.3 quelques propriétés de ces matrices. Notons que ces matrices généralisent la matrice d'incidence du p-morphisme. Plus précisément, si $T = 1$, alors $M(1)$ est exactement la matrice d'incidence du morphisme sous-jacent.

Dans la dernière partie du Chapitre 6, nous illustrons notre approche à l'aide de quelques exemples. Nous donnons en particulier plusieurs séries de Laurent algébriques pour lesquelles nous calculons, à l'aide de l'algorithme précédent, la valeur exacte de l'exposant d'irrationalité.

4.3 Complexité et séries formelles à coefficients dans un corps fini

La suite des chiffres décimaux ou binaires de nombres réels algébriques irrationnels, comme $\sqrt{2}$, est source de nombreux problèmes difficiles. On s'attend

généralement à ce que ces nombres soient des nombres normaux, mais cette conjecture semble malheureusement hors d'atteinte. En effet, nous sommes par exemple toujours incapables de savoir si le nombre 2 apparaît une infinité de fois dans le développement décimal de $\sqrt{2}$. Une approche pertinente pour étudier ce type de questions consiste à introduire une notion de complexité pour les nombres réels, fondée sur la notion de complexité des facteurs des mots infinis (voir [5, 4, 64]). La normalité d'un nombre réel implique alors que sa complexité doit être maximale en toute base entière, en particulier celle-ci doit être d'ordre exponentiel. On peut alors chercher à minorer la complexité des nombres algébriques irrationnels. Voir à ce sujet, les avancées récentes [4, 36]. Au-delà des nombres algébriques irrationnels, un intérêt particulier est porté à l'étude des constantes classiques comme π, $\log 2$, $\zeta(3)$ ou e. Kontsevitch et Zagier [76] ont défini les notions de « période » et de « période exponentielle » afin d'offrir un cadre commun pour l'étude de ces nombres. Ainsi, les nombres algébriques, π ou $\log 2$ sont des périodes ; quant au nombre e, c'est l'exemple typique d'une période exponentielle. Malgré un résultat récent [1] donnant une minoration de la complexité du nombre e et de certaines périodes exponentielles transcendantes, nos connaissances sur la complexité de périodes transcendantes sont encore plus limitées que celles concernant les nombres algébriques.

Motivés par l'analogie, mais aussi par les différences, entre le corps des nombres réels et les corps des séries de Laurent à coefficients dans un corps fini, nous étudions dans le chapitre 7 des questions liées à la complexité des éléments du corps $\mathbb{F}_q((T^{-1}))$, où q est une puissance d'un nombre premier p. Dans ce travail, nous introduisons tout d'abord une notion de complexité pour de telles séries de Laurent à l'aide de la complexité des facteurs des mots infinis. Comme dans le cas de nombres réels, nous observons qu'une série de Laurent est une fraction rationnelle si et seulement si sa fonction de complexité est bornée. De plus, comme conséquence du théorème de Christol et d'un théorème de Cobham, on obtient que la complexité d'une série de Laurent algébrique est majorée par une fonction affine (on dira alors qu'elle est de complexité sous-linéaire ou au plus linéaire). Il est alors naturel de se demander si certaines fonctions transcendantes classiques ont également une faible complexité.

Nous nous intéressons, dans un premier temps, à l'analogue du nombre π, la série formelle Π_q de Carlitz. Malheureusement, il semble difficile d'obtenir des renseignements sur le développement en série de Laurent de Π_q. Toutefois, nous parvenons à obtenir des résultats précis sur la complexité de l'inverse de Π_q. Plus exactement, nous obtenons dans le théorème 7.1.2 que la complexité de $1/\Pi_q$ est d'ordre linéaire lorsque $q \geq 3$, tandis qu'elle est d'ordre quadratique si $q = 2$. Ce résultat se démontre en étudiant certains motifs répétitifs qui apparaissent dans le développement en série de Laurent de $1/\Pi_q$, comme cela est décrit dans [12]. Notons que ce travail fournit, au passage, une nouvelle preuve de la transcendance de Π_q lorsque $q = 2$.

La notion de complexité que nous avons introduite induit naturellement

une hiérarchie des séries de Laurent à coefficients dans un corps fini. Dans la seconde partie de ce chapitre, nous considérons certaines classes de cette hiérarchie, à savoir la classe des séries de Laurent de complexité (au plus) polynomiale et celle des série de Laurent d'entropie nulle, et nous étudions leurs propriétés de stabilité. Ces deux classes sont d'autant plus naturelles qu'elles contiennent les séries algébriques ainsi que des séries de Laurent transcendantes arithmétiquement intéressantes comme $1/\Pi_q$. On peut d'ailleurs se demander s'il existe des analogues de périodes qui n'appartiendraient pas à l'une de ces deux classes. Nous montrons que chacune de ces deux classes est un espace vectoriel (de dimension infinie) sur le corps des fractions rationnelles $\mathbb{F}_q(T)$ (voir le théorème 7.1.3). Ceci conduit, bien évidement, à un résultat d'indépendance linéaire sur le corps $\mathbb{F}_q(T)$ (voir la proposition 7.4.4). De plus, nous prouvons que ces classes sont laissées stables par l'application d'un certain nombre d'opérations classiques comme le produit d'Hadamard, l'opérateur de dérivation ou les opérateurs de Cartier.

Une question plus difficile consisterait à déterminer si l'ensemble des séries de complexité polynomiale ou, plus largement, celles d'entropie nulle, forment un anneau ou un corps. Une des motivations serait l'obtention, *in fine*, de résultats d'indépendance algébrique. Comme préambule à une telle étude, nous considérons dans la dernière partie du chapitre 7, quelques produits (de Cauchy) de séries formelles de faible complexité. De façon un peu surprenante, nous remarquons qu'il semble déjà délicat de majorer la complexité d'un produit aussi simple que celui de deux séries de Laurent lacunaires. Ainsi, pour démontrer le théorème 7.5.1, nous devons utiliser un résultat classique de la théorie des équations en S-unités, puis le combiner à des idées analogues à celles développées dans la preuve du théorème 7.1.2 à propos de la complexité de l'inverse de Π_q.

Deuxième partie

Présentation des travaux

5

Sur une généralisation de la cubique de Baum et Sweet

Dans ce chapitre nous considérons les équations de la forme $TX^{r+1}+X-T=0$, où r est une puissance du nombre premier p. Nous montrons qu'une équation de ce type admet une unique solution dans le corps $\mathbb{F}_p((T^{-1}))$ et nous nous interessons à son développement en fraction continue. Notons que dans le cas où $r=p=2$, cette équation a comme solution la fameuse cubique proposée par Baum et Sweet en 1976. Inspirés par les travaux de Mills et Robbins et Lasjaunias, nous présentons une méthode qui nous permet de décrire la suite des quotients partiels de cette famille de séries de Laurent algébriques. Ce chapitre a fait l'objet d'un article accepté pour publication au *Journal de Théorie des Nombres de Bordeaux*.

5.1 Introduction

Il y a une trentaine d'années, les travaux de Baum et Sweet [24] ont ouvert un nouveau domaine de recherche sur l'approximation diophantienne dans les corps de séries formelles à coefficients dans un corps fini, par le biais du développement en fraction continue. Ces auteurs ont notamment donné l'exemple d'une série formelle à coefficients dans le corps fini $\mathbb{F}_2$, algébrique de degré 3 sur $\mathbb{F}_2(T)$, ayant un développement en fraction continue avec des quotients partiels qui sont tous des polynômes en T de degré 1 ou 2. Dix ans plus tard, Mills et Robbins [95] ont décrit un algorithme qui leur a permis de donner le développement explicite en fraction continue pour la série formelle cubique de Baum

et Sweet. Ces travaux ont mis en lumière un sous ensemble de séries formelles algébriques, obtenues comme points fixes de la composée d'une homographie à coefficients entiers (polynômes) avec le morphisme de Frobenius ; ces séries sont alors appelées hyperquadratiques. Le développement en fraction continue a pu être donné explicitement (voir [79, 95, 110] pour plus de références) pour de nombreux exemples de ces séries formelles.

Nous avons observé que pour tout nombre premier p et $r = p^t$, l'équation

$$TX^{r+1} + X - T = 0$$

a une unique solution dans le corps $\mathbb{F}_p((T^{-1}))$. Pour $r = p = 2$ cette solution est la cubique de Baum et Sweet. Dans cet article, nous donnons le développement en fraction continue de cette solution pour $r > 2$ (voir la partie 5.3). On peut par ailleurs remarquer que si l'on remplace r par 2 dans les formules obtenues, alors le développement obtenu est impropre (une sous-suite de quotients partiels tend vers 0 dans $\mathbb{F}_p((T^{-1}))$). Cependant, cette expression, lorsqu'elle est tronquée et rendue propre, donne le développement qui a été obtenu dans [95].

Pour obtenir ce développement nous utilisons une méthode déjà utilisée par Lasjaunias dans [82], qui, bien que proche de l'algorithme de Mills et Robbins, en diffère un peu. Pour illustrer cette méthode, nous l'appliquons dans un premier temps à un autre exemple de série formelle hyperquadratique, à coefficients dans $\mathbb{F}_p$. Cet exemple, très célèbre, a été introduit par Mahler [91] dans un article fondateur sur l'approximation diophantienne dans les corps de fonctions.

Nous rappelons brièvement les notations utilisées. Dans ce texte, p est un nombre premier, $\mathbb{F}_p$ désigne le corps fini à p éléments, $\mathbb{F}_p[T]$, $\mathbb{F}_p(T)$ et $\mathbb{F}_p((T^{-1}))$ sont, respectivement, l'anneau des polynômes, le corps des fonctions rationnelles et le corps des séries formelles (en $1/T$) de la variable T sur $\mathbb{F}_p$. Ainsi

$$\mathbb{F}_p((T^{-1})) = \{0\} \cup \left\{ \sum_{k \leq k_0} u_k T^k, k_0 \in \mathbb{Z}, u_k \in \mathbb{F}_p, u_{k_0} \neq 0 \right\}.$$

Ce corps de séries formelles est muni d'une valeur absolue ultramétrique définie par $|\alpha| = |T|^{k_0}$ et $|0| = 0$, où $|T|$ est un réel fixé strictement supérieur à 1. De plus, il est connu que $\mathbb{F}_p((T^{-1}))$ est le completé de $\mathbb{F}_p(T)$ pour cette valeur absolue.

Dans la suite, r est une puissance de p, $r = p^t$, avec $t \geq 1$ entier. Le morphisme de Frobenius défini dans $\mathbb{F}_p((T^{-1}))$ est noté $\alpha \mapsto \alpha^r$. Une série formelle, $\alpha \in \mathbb{F}_p((T^{-1}))$, est dite hyperquadratique si l'on a $\alpha = f(\alpha^r)$ où f est une homographie à coefficients dans $\mathbb{F}_p[T]$.

Tout élément $\alpha \in \mathbb{F}_p((T^{-1}))$ a un développement en fraction continue (infini si α n'est pas une fraction rationnelle) que l'on notera

$$\alpha = [a_0, a_1, a_2, \ldots, a_n, \alpha_{n+1}] \tag{5.1}$$

où les $a_i \in \mathbb{F}_p[T]$ (avec $\deg(a_i) > 0$ pour $i > 0$) sont appelés les quotients partiels et les $\alpha_i \in \mathbb{F}_p((T^{-1}))$ sont les quotients complets.

5.2 Méthode employée et exemple de Mahler

Dans cette partie nous présentons le raisonnement sur lequel les preuves reposent. Nous commençons par un lemme élémentaire concernant les fractions continues. Une courte démonstration en est donnée dans l'article [82].

Nous rappelons déjà la notation suivante. Soit $P/Q \in \mathbb{F}_p(T)$ tel que $P/Q := [a_1, a_2, \ldots, a_n]$. Pour tout $x \in \mathbb{F}_p(T)$, nous noterons

$$\Big[[a_1, a_2, \ldots, a_n], x\Big] := \frac{P}{Q} + \frac{1}{x}.$$

Lemme 5.2.1. *Soient $a_1, \ldots, a_n, x \in \mathbb{F}_p(T)$. On a la relation suivante :*

$$\Big[[a_1, a_2, \ldots, a_n], x\Big] = [a_1, a_2, \ldots, a_n, x'],$$

avec

$$x' = f_n x + g_n, \tag{5.2}$$

où les f_n, g_n sont des éléments de $\mathbb{F}_p(a_1, a_2, \ldots, a_n)$ (voir [82], page 330).

A l'exception du paragraphe 5.2.2, nous utilisons ce lemme uniquement dans les cas $n = 2$ et $n = 3$, qui s'énoncent comme suit.

Lemme 5.2.2. *Soient $a_1, a_2, a_3, x \in \mathbb{F}_p(T)$. On a les relations suivantes :*

$$\Big[[a_1, a_2], x\Big] = [a_1, a_2, y], \ \text{où } y = -a_2^{-2}x - a_2^{-1},$$

$$\Big[[a_1, a_2, a_3], x\Big] = [a_1, a_2, a_3, y], \ \text{où } y = (a_2a_3 + 1)^{-2}x - a_2(a_2a_3 + 1)^{-1}.$$

5.2.1 Premier exemple

Nous allons à présent décrire la suite de quotients partiels d'un ensemble de séries formelles vérifiant un certain type d'équation. Comme corollaire, nous obtenons le développement en fraction continue de la série de Mahler :

$$\Theta_r = 1/T + 1/T^r + 1/T^{r^2} + \cdots + 1/T^{r^k} + \cdots \in \mathbb{F}_p((T^{-1})). \tag{5.3}$$

On peut remarquer que Θ_r est une série algébrique de degré r vérifiant l'équation

$$Tz^r - Tz + 1 = 0. \tag{5.4}$$

Il convient de noter que le développement en fraction continue de Θ_r, donné dans le corollaire 5.2.1 est déjà connu, même s'il n'a jamais été présenté sous cette forme. En effet, il est exposé dans [79] (p. 215) et peut aussi être déduit de travaux plus anciens de Shallit sur les fractions continues de certains nombres réels [111].

Théorème 5.2.1. *Soit p un nombre premier et $r = p^t$, $t \geq 1$, avec $r > 2$. Soit $\ell \in \mathbb{N}$, $\ell \geq 1$ et soit $(a_1, a_2, \ldots, a_l)$ un ℓ-uplet de polynômes dans $F_p[T]$, avec $a_i(T) \in T\mathbb{F}_p[T]$, pour tout i impair, $1 \leq i \leq l$. Si z est la fraction continue $z = [a_1, a_2, \ldots, a_\ell, z_{\ell+1}]$ vérifiant l'équation :*

$$z^r = -T^2 z_{\ell+1} - T, \tag{5.5}$$

alors la suite de quotients partiels de z, $(a_n)_{n \geq l+1} \in (\mathbb{F}_p[T])^{\mathbb{N}}$ est définie pour $k \geq 0$ par :

$$\begin{array}{rclcrcl} a_{\ell+4k+1} & = & -\dfrac{a_{2k+1}^r}{T^2}, & & a_{\ell+4k+2} & = & -T, \\ a_{\ell+4k+3} & = & a_{2k+2}^r, & & a_{\ell+4k+4} & = & T. \end{array}$$

***Remarque* 5.2.1.** L'existence de la fraction continue vérifiant (5.5) découle du théorème 1 de l'article [82].

Revenons maintenant à la série Θ_r. On pose $y := 1/\Theta_r$ et $y := [a_1, a_2, \ldots, a_n, \ldots]$. D'après (5.3), pour $r > 2$, on a

$$\left| T - \frac{1}{\Theta_r} \right| = \frac{1}{|T^{r-1}\Theta_r|} = \frac{1}{|T|^{r-2}} < 1,$$

et par conséquent $y = T + 1/y_2 = [T, y_2]$. D'après (5.4) on a

$$\frac{T}{y^r} = \frac{T}{y} - 1 = -\frac{1}{Ty_2 + 1},$$

et donc y_2 vérifie la relation :

$$y^r = -T^2 y_2 - T. \tag{5.6}$$

En remarquant que l'équation (5.6) est un cas particulier de l'équation (5.5), où $\ell = 1$ at $a_1 = T$, nous en déduisons directement le corollaire suivant.

Corollaire 5.2.1. *On a $\Theta_r = [0, a_1, a_2, \ldots, a_n, \ldots]$ où la suite $(a_i)_{i \geq 1}$ est définie par récurrence, pour $k \geq 0$, par :*

$$\begin{array}{rclcrcl} a_{4k+1} & = & T, & & a_{4k+2} & = & -a_{2k+1}^r/T^2, \\ a_{4k+3} & = & -T, & & a_{4k+4} & = & a_{2k+2}^r. \end{array}$$

Démonstration du théorème 5.2.1. Nous partons de la relation :

$$z^r = -T^2 z_{\ell+1} - T,$$

qui peut être écrite aussi sous la forme $[a_1^r, z_2^r] = -T^2 z_{\ell+1} - T$, ou encore :

$$\left[[-\frac{a_1^r}{T^2}, -T], -T^2 z_2^r \right] = z_{\ell+1}.$$

En utilisant le lemme 5.2.2 et en tenant en compte du fait que a_1 est divisible par T, nous en déduisons les relations suivantes :

$$z_{\ell+1} = [-\frac{a_1^r}{T^2}, -T, z'], \text{ avec } z' = z_2^r + T^{-1}.$$

Puisque $|z'| > 1$, on en déduit que $z' = z_{l+3}$. On a donc

$$a_{\ell+1} = -\frac{a_1^r}{T^2},\ a_{\ell+2} = -T \text{ et } z_{l+3} = z_2^r + \frac{1}{T}.$$

Nous appliquons de nouveau le même raisonnement et nous obtenons :

$$z_{l+3} = z_2^r + \frac{1}{T} = [a_2^r + \frac{1}{T}, z_3^r] = [a_2^r, T, z''],$$

avec $z'' = -T^{-2}z_3^r - T^{-1}$. Puisque $r > 2$, on a $|z''| > 1$ et alors, par identification, $z'' = z_{l+5}$. On a donc

$$a_{\ell+3} = a_2^r,\ a_{\ell+4} = T \text{ et } z_{l+5} = -T^{-2}z_3^r - T^{-1}.$$

En résumé,

$$z_{\ell+1} = [-\frac{a_1^r}{T^2}, -T, a_2^r, T, z_{\ell+5}].$$

Plus généralement, par une simple récurrence sur k, on obtient que :

$$z_{2k+1}^r = -T^2 z_{\ell+4k+1} - T.$$

Puisque $a_{2k+1} \in T\mathbb{F}_p[T]$, nous obtenons comme précédemment :

$$z_{\ell+4k+1} = [-\frac{a_{2k+1}^r}{T^2}, -T, a_{2k+2}^r, T, z_{\ell+4k+5}],$$

ce qui termine la démonstration. □

5.2.2 Le contexte général

Dans cette partie nous énonçons le raisonnement général et les notations utilisées dans les preuves qui suivent. Ceux-ci restent proches de ceux utilisés par Mills et Robbins dans [95].

Soit $(P, Q, R) \in \mathbb{F}_p[T]^3$ et soient m, n deux entiers strictement positifs, $m < n$. On dit que z satisfait une relation du type (P, Q, R, m, n) si on a :

$$Pz_m^r = Qz_n + R, \tag{5.7}$$

où la notation z_k, pour $k \in \mathbb{N}^*$, designe un quotient complet de z, comme défini en (5.1).

Dans la suite, nous considérons une série $z = [a_1, a_2, \ldots] \in \mathbb{F}_p((T^{-1}))$ satisfaisant (5.7), avec un triplet (P, Q, R) bien choisi. En fait, il est probable que cette relation soit vraie pour presque toutes les séries formelles hyperquadratiques, mais à notre connaissance, aucun résultat général ne le confirme (le lecteur peut consulter [82], page 333, pour quelques commentaires à ce sujet).

On suppose connus les $n-1$ premiers coefficients partiels : $a_1, a_2, \ldots, a_{n-1}$ (qui peuvent être vus comme une « donnée de départ »). La relation (5.7) implique :

$$Pa_m^r + \frac{P}{z_{m+1}^r} = Qz_n + R,$$

car, par définition, $z_m = [a_m, z_{m+1}]$. Ainsi, nous obtenons :

$$\frac{Pa_m^r - R}{Q} + \frac{P}{Qz_{m+1}^r} = z_n. \tag{5.8}$$

Puisque

$$\frac{Pa_m^r - R}{Q} \in \mathbb{F}_p(T),$$

cette expression a un développement en fraction continue qui est fini. Il existe ainsi des polynômes $\lambda_1, \lambda_2, \ldots, \lambda_\ell \in \mathbb{F}_p[T]$, que l'on peut calculer, tels que :

$$\frac{Pa_m^r - R}{Q} = [\lambda_1, \lambda_2, \ldots, \lambda_\ell].$$

La relation (5.8) devient :

$$\left[[\lambda_1, \lambda_2, \ldots, \lambda_\ell], \frac{Qz_{m+1}^r}{P}\right] = z_n,$$

et le lemme 5.2.1 implique alors que :

$$z_n = [\lambda_1, \lambda_2, \ldots, \lambda_\ell, z'].$$

A ce moment, si $|z'| > 1$, nous pouvons déjà identifier $a_n = \lambda_1, a_{n+1} = \lambda_2, \ldots, a_{n+l-1} = \lambda_\ell$ et $z_{n+\ell} = z'$.
Par le lemme 5.2.1 on déduit donc :

$$z_{n+\ell} = f_\ell \frac{Qz_{m+1}^r}{P} + g_\ell,$$

où f_ℓ et g_ℓ appartiennent à $\mathbb{F}_p(T)$ (voir la formule (5.2)). Il existe donc trois polynômes $P_1, Q_1, R_1 \in \mathbb{F}_p[T]$, qu'on peut déterminer explicitement, tels que :

$$P_1 z_{m+1}^r = Q_1 z_{n+\ell} + R_1.$$

En résumé, la connaissance de P, Q, R et a_m nous permet de déterminer les ℓ nouveaux quotients partiels : $a_n, a_{n+1}, \ldots, a_{n+\ell-1}$ et une nouvelle équation :

$$P_1 z_{m+1}^r = Q_1 z_{n+\ell} + R_1.$$

Dans la suite, nous noterons ce raisonnement par :

$$(P, Q, R, m, n : a_m) \to (P_1, Q_1, R_1, m+1, n+\ell : a_n, a_{n+1}, \ldots, a_{n+\ell-1}).$$

Plus généralement, si $X := (P, Q, R, m, n)$ et $Y := (P_\ell, Q_\ell, R_\ell, m+\ell, n+k)$, nous noterons :

$$(X : a_m, a_{m+1}, \ldots, a_{m+\ell-1}) \to (Y : a_n, a_{n+1}, \ldots, a_{n+k-1}) \tag{5.9}$$

pour désigner la suite des raisonnements suivants :

$$\begin{cases} (X : a_m) & \to (X_1 : a_n, a_{n+1}, \ldots, a_{n+k_1-1}) \\ (X_1 : a_{m+1}) & \to (X_2 : a_{n+k_1}, a_{n+k_1+1}, \ldots a_{n+k_2-1}) \\ & \vdots \\ (X_{\ell-1} : a_{m+\ell-1}) & \to (Y : a_{n+k_{\ell-1}}, a_{n+k_{\ell-1}+1}, \ldots, a_{n+k-1}). \end{cases}$$

où $X_i := (P_i, Q_i, R_i, m+i, n+k_i)$, pour $1 \leq i \leq \ell - 1$. La nouvelle équation va donc relier $z_{m+\ell}$ et z_{n+k} de la manière suivante :

$$P_\ell z^r_{m+\ell} = Q_\ell z_{n+k} + R_\ell.$$

Nous remarquons que, lorsque $\ell = n - m$, la relation (5.9) s'écrit :

$$(P, Q, R, m, n : a_m, a_{m+1}, \ldots, a_{n-1}) \to (P_{n-m}, Q_{n-m}, R_{n-m} : a_n, a_{n+1} \ldots, a_{n+k-1}),$$

ce qui signifie que la connaissance de $a_m, a_{m+1}, \ldots, a_{n-1}$ nous permet de calculer les k quotients partiels suivants, à savoir $a_n, a_{n+1}, \ldots, a_{n+k-1}$. Nous pouvons alors appliquer de nouveau le même raisonnement, à une équation du type $(P_{n-m}, Q_{n-m}, R_{n-m}, n, n+k)$ ayant comme « donnée de départ » les quotients partiels $a_n, a_{n+1}, \ldots, a_{n+k-1}$. L'itération de ce procédé permet ainsi d'obtenir tous les quotients partiels de la série z.

Soient W, W' deux suites finies de polynômes à coefficients dans $\mathbb{F}_p$. On note $|W|$ la longueur de W, c'est-à-dire le nombre de ses termes. Alors la relation (5.9) peut être écrite aussi :

$$(P, Q, R, m, n : W) \to (P_{|W|}, Q_{|W|}, R_{|W|}, m + |W|, n + |W'| : W').$$

Pour alléger les notations, nous nous permettons parfois d'écrire :

$$(P, Q, R, m, n : W) \to (P_{|W|}, Q_{|W|}, R_{|W|} : W').$$

5.3 Généralisation de la cubique de Baum et Sweet

Avant d'énoncer notre résultat principal, nous allons tout d'abord introduire quelques définitions.

Définition 5.3.1. *On définit la suite de liste de polynômes à coefficients dans $\mathbb{F}_p$, $(\Gamma_k)_{k\geq 1}$, de la façon suivante. On pose $\Gamma_1 := -T, T^r, T$. Pour $k > 1$, Γ_k est défini récursivement par :*

$$\Gamma_k := a_1, a_2, \ldots, a_{2^{k+1}-1} \text{ et } \Gamma_{k+1} := b_1, b_2, \ldots, b_{2^{k+2}-1},$$

où

$$b_{2i-1} = (-1)^{i+k} T \quad \text{pour } 1 \leq i \leq 2^{k+1}$$

et

$$b_{4i} = a^r_{2i}/T^2 \text{ pour } 1 \leq i \leq 2^k - 1, \quad b_{4i-2} = -a^r_{2i-1} \text{ pour } 1 \leq i \leq 2^k.$$

Définition 5.3.2. *On définit la suite de liste de polynômes à coefficients dans* $\mathbb{F}_p$, $(\Lambda_k)_{k\geq 1}$, *de la façon suivante. On pose* $\Lambda_1 := T+1, T-1$, $\Lambda_2 := T, -T^r+1, -T$. *Pour* $k \geq 3$, Λ_k *est défini récursivement par :*

$$\Lambda_k := \Lambda_{k-2}, -T^{\lambda_{k-1}}, \Gamma_{k-2},$$

où $(\lambda_k)_{k\geq 1}$ *est défini comme suit :*

$$\lambda_1 = r, \quad \lambda_{k+1} = r\lambda_k - 2.$$

Etant donnée une suite $W = a_1, \ldots, a_m$ à valeurs dans $\mathbb{F}_p[T]$, on note $-W := -a_1, \ldots, -a_m$ et $\overline{W} := a_m, a_{m-1}, \ldots, a_1$.

Définition 5.3.3. *On définit la suite de liste de polynômes à coefficients dans* $\mathbb{F}_p$, $(\Omega_k)_{k\geq 1}$, *de la façon suivante. On pose* $\Omega_1 := -T^{\omega_1}$. *Pour* $k \geq 2$, Ω_k *est défini récursivement par :*

$$\Omega_k := \Omega_{k-1}, \Lambda_{k-1}, -T^{\omega_k}, -\overline{\Lambda}_{k-1}, -\overline{\Omega}_{k-1}, \tag{5.10}$$

où $(\omega_k)_{k\geq 1}$ *est défini comme suit :*

$$\omega_1 = r-2, \quad \omega_{k+1} = r\omega_k - 2.$$

Nous notons Ω_∞ *la suite infinie commençant par* Ω_k, *pour tout* $k \geq 1$.

Avec les notations ci-dessus, nous présentons maintenant notre résultat principal.

Théorème 5.3.1. *Soit* p *un nombre premier et* $r = p^t$, *où* $t \geq 1$, *avec* $r > 2$. *L'équation*

$$TX^{r+1} + X - T = 0$$

a une unique racine $f_r(T)$ *dans le corps* $\mathbb{F}_p((T^{-1}))$ *dont le développement en fraction continue est :*

$$[1, -T-1, \Omega_\infty].$$

Nous remarquons par ailleurs que, si $r > 2$, la suite des quotients partiels de $f_r(T)$ n'est pas bornée, tandis que la cubique de Baum et Sweet, qui correspond à $f_2(T)$ lorsque $p = 2$, a une suite des quotients partiels bornée.

5.3.1 Démonstration du théorème 5.3.1

Tout d'abord, nous montrons que l'équation :

$$TX^{r+1} + X - T = 0 \tag{5.11}$$

a une unique solution dans le corps $\mathbb{F}_p((T^{-1}))$.
En effet, de (5.11), on déduit que :

$$X = \frac{T}{TX^r + 1}. \tag{5.12}$$

Nous considérons maintenant l'application $f(X) = T/(TX^r + 1)$ définie sur $\mathbb{F}_p((T^{-1}))$ à valeurs dans $\mathbb{F}_p((T^{-1}))$. Il n'est pas difficile de voir que f est une application strictement contractante (puisque $r > 2$) et nous savons que $\mathbb{F}_p((T^{-1}))$ est un espace complet pour la distance ultramétrique usuelle. Ainsi, par utilisation du théorème du point fixe, l'équation $f(X) = X$ a une unique solution dans $\mathbb{F}_p((T^{-1}))$. Celle-ci est donc l'unique solution de l'équation (5.11), que nous noterons dans la suite BS_r.
Soit z la série définie par :

$$BS_r = 1 + \frac{1}{(-T-1) + 1/z}. \tag{5.13}$$

A partir de (5.12) et (5.13), un calcul simple nous conduit à :

$$z = \frac{(-T^r + T - 1)z^r + 1}{T^2 z^r}, \tag{5.14}$$

ce qui entraîne

$$\left| z - \frac{-T^r + T - 1}{T^2} \right| = \frac{1}{|T^2 z^r|} < \frac{1}{|T^4|},$$

l'inegalité $|z^r| > |T^2|$ etant due au fait que $|z| > |T|$ et $r > 2$.

Autrement dit, $(-T^r+T-1)/T^2$ est une réduite de z, donc les trois premiers quotients partiels de z sont $-T^{r-2}, T+1$ et $T-1$. Ainsi $z = [-T^{r-2}, T+1, T-1, z_4]$ et d'après (5.14) on obtient la relation suivante :

$$z^r = T^2 z_4 + (T+1).$$

Le théorème 5.3.1 est alors une conséquence directe de proposition ci-dessous.

Proposition 5.3.1. *Soit z la fraction continue infinie $z := [-T^{r-2}, T+1, T-1, z_4]$ satisfaisant :*

$$z^r = T^2 z_4 + (T+1). \tag{5.15}$$

Alors la suite de quotients partiels de z est Ω_∞.

Notation 5.3.1. On dit qu'une équation est du type A_1 si $P = 1$, $Q = T^2$ et $R = T + 1$ et on note $A_1 := (1, T^2, T + 1)$. De la même manière, nous allons définir les équations des types suivants :

$$\begin{aligned} A_2 &:= (1, T^2, -T + 1), \\ A_3 &:= (T, -T, -1), \\ A_4 &:= (1, T^2, -T), \\ A_5 &:= (T, -T, 1), \\ A_6 &:= (1, T^2, T). \end{aligned}$$

Dans ce qui suit, nous utiliserons les notations décrites dans la partie 5.2.2.

Lemme 5.3.1. *Soient $m, n \in \mathbb{N}$ tels que $m < n$ et soit $a \in \mathbb{F}_p[T]$. On a les relations suivantes :*

$$(A_1, m, n : a) \to \left(A_2, m+1, n+3 : \frac{a^r}{T^2}, -T+1, -T-1\right) \text{ si } a \equiv 0[T],$$
$$(A_1, m, n : a) \to \left(A_5, m+1, n+2 : \frac{(a-1)^r}{T^2}, -T\right) \text{ si } a \equiv 1[T],$$
$$(A_2, m, n : a) \to \left(A_1, m+1, n+3 : \frac{a^r}{T^2}, T+1, T-1\right) \text{ si } a \equiv 0[T],$$
$$(A_2, m, n : a) \to \left(A_3, m+1, n+2 : \frac{(a-1)^r}{T^2}, T\right) \text{ si } a \equiv 1[T],$$
$$(A_3, m, n : a) \to (A_4, m+1, n+2 : -a^r, -T) \text{ pour tout } a \in \mathbb{F}_p[T],$$
$$(A_4, m, n : a) \to \left(A_3, m+1, n+2 : \frac{a^r}{T^2}, T\right) \text{ si } a \equiv 0[T],$$
$$(A_4, m, n : a) \to \left(A_2, m+1, n+3 : \frac{(a+1)^r}{T^2}, T+1, T-1\right) \text{ si } a \equiv -1[T],$$
$$(A_5, m, n : a) \to (A_6, m+1, n+2 : -a^r, T) \text{ pour tout } a \in \mathbb{F}_p[T],$$
$$(A_6, m, n : a) \to \left(A_5, m+1, n+2 : \frac{a^r}{T^2}, -T\right) \text{ si } a \equiv 0[T],$$
$$(A_6, m, n : a) \to \left(A_2, m+1, n+3 : \frac{(a+1)^r}{T^2}, -T+1, -T-1 :\right) \text{ si } a \equiv -1[T].$$

Démonstration. Nous allons prouver le premier cas, c'est-à-dire le cas où z satisfait une relation du type A_1, avec $a \equiv 0[T]$:

$$z_m^r = T^2 z_n + (T+1).$$

Cette relation s'écrit aussi sous la forme : $[a^r, z_{m+1}^r] = T^2 z_n + (T+1)$ ou bien

$$\frac{a^r}{T^2} - \frac{T+1}{T^2} + \frac{1}{T^2 z_{m+1}^r} = z_n.$$

Puisque a est divisible par T, nous obtenons :

$$\left[[\frac{a^r}{T^2}, -T+1, -T-1], T^2 z_{m+1}^r\right] = z_n.$$

En appliquant le lemme 5.2.2, on en déduit que

$$z_n = [\frac{a^r}{T^2}, -T+1, -T-1, z_{n+3}]$$

et

$$z_{m+1}^r = T^2 z_{n+3} + (-T+1).$$

Nous remarquons que z_{n+3} est bien un quotient complet puisque $|z_{n+3}| > 1$ (en utilisant la relation précèdente, le fait que $|z_{m+1}| > 1$ et le fait que $r > 2$).

Ainsi nous obtenons les nouveaux quotients partiels $a_n = \frac{a^r}{T^2}, a_{n+1} = -T + 1, a_{n+2} = -T - 1$ et une équation du type $A_2 := (1, T^2, -T + 1)$.
Les autres cas se déduisent de manière analogue, en appliquant le raisonnement précédent et, en particulier, le lemme 5.2.2. □

Notation 5.3.2. Soient $i, j \in \mathbb{N}^*$, $i, j \leq m$ et $W = a_1, a_2, \ldots, a_m$. On notera

$$^{(i)}W := a_{i+1}, a_{i+2}, \ldots, a_m \text{ et } W^{(j)} := a_1, a_2, \ldots, a_{m-j}.$$

Lorsque $i + j < m - 1$, on notera

$$^{(i)}W^{(j)} := a_{i+1}, \ldots, a_{m-j}.$$

La proposition 5.3.1 est une conséquence immédiate de la proposition 5.3.2.

Proposition 5.3.2. *Pour tout $k \in \mathbb{N}^*$, on a la relation suivante :*

$$(A_1, 1, 4 : \Omega_k) \to \left(A_2, 1 + |\Omega_k|, |\Omega_{k+1}| : {}^{(3)}\Omega^{(1)}_{k+1}\right).$$

***Remarque* 5.3.1.** Soit $k \in \mathbb{N}$. Nous notons ℓ_k la longueur du mot fini Ω_k. La proposition précédente peut être traduite de la manière suivante. On suppose connus les ℓ_k premiers quotients de z. Alors, si on applique ℓ_k fois le procédé énoncé dans le paragraphe 5.2.2 à l'équation (5.15) nous obtenons :

$$z_4 = [{}^{(3)}\Omega^{(1)}_{k+1}, z_{\ell_{k+1}}]$$

et la nouvelle relation sera :

$$z_{\ell_k+1} = T^2 z_{\ell_{k+1}} + (-T + 1).$$

Pour démontrer la proposition 5.3.2, nous allons utiliser les lemmes suivants.

Lemme 5.3.2. *Soient k, m, n des entiers strictement positifs, $m < n$. On a les relations suivantes :*

- *si k est pair, alors :*

$$(A_5, m, n : \Gamma_k) \to (A_6 : {}^{(1)}\Gamma_{k+1}),$$
$$(A_5, m, n : -\overline{\Gamma}_k) \to (A_6 : {}^{(1)}(-\overline{\Gamma}_{k+1}));$$

- *si k est impair, alors :*

$$(A_3, m, n : \Gamma_k) \to (A_4 : {}^{(1)}\Gamma_{k+1}),$$
$$(A_3, m, n : -\overline{\Gamma}_k) \to (A_4 : {}^{(1)}(-\overline{\Gamma}_{k+1})).$$

Démonstration. Tout d'abord, on remarque que, pour tout $k \in \mathbb{N}$, les termes de Γ_k sont des polynômes divisibles par T.
Soient $a, b \in T\mathbb{F}_p(T)$. Alors, à l'aide du lemme 5.3.1 on peut déduire :

$$(A_5, m, n : a, b) \to (A_5, m+2, n+4 : -a^r, T, b^r/T^2, -T).$$

Plus généralement, si $a_1, a_2, \ldots, a_{2i} \in T\mathbb{F}_p[T]$ alors :

$$(A_5, m, n : a_1, a_2 \ldots, a_{2i}) \to (A_5 : -a_1^r, T, a_2^r/T^2, -T, \ldots, -a_{2i-1}^r, T, a_{2i}^r/T^2, -T).$$

Soit $k \in \mathbb{N}$ et $\Gamma_k := a_1, a_2, \ldots, a_{2^{k+1}-1}$. La suite Γ_k a un nombre impair d'éléments ; ainsi, nous devons appliquer le lemme 5.3.1 à son dernier terme aussi. On obtient :

$$(A_5, m, n : \Gamma_k) \to (A_6 : -a_1^r, T, a_2^r/T^2, -T, \ldots, -a_{2i-1}^r, T, a_{2i}^r/T^2, -T, -a_{2i+1}^r, T).$$

Dans le cas où k est pair, le premier terme de Γ_{k+1} est $b_1 := -T$. Donc, par définition,

$$^{(1)}\Gamma_{k+1} = b_2, \ldots, b_{2^{k+2}-1} = -a_1^r, T, \ldots, -a_{2i-1}^r, T, a_{2i}^r/T^2, -T, -a_{2i+1}^r, T,$$

ce qui coïncide avec notre résultat.
Les autres cas se démontrent de manière analogue. □

Lemme 5.3.3. *Soient k, m, n des entiers strictement positifs, $m < n$. On a les relations suivantes :*

- *si k est pair, alors :*

$$(A_2, m, n : \Lambda_k) \to (A_6 : T^{r-2}, \Lambda_{k+1}), \tag{5.16}$$
$$(A_5, m, n : -\overline{\Lambda}_k) \to (A_1 : {}^{(1)}(-\overline{\Lambda}_{k+1}), T^{r-2}, T+1, T-1). \tag{5.17}$$

- *si k est impair, alors :*

$$(A_2, m, n : \Lambda_k) \to (A_4 : T^{r-2}, \Lambda_{k+1}), \tag{5.18}$$
$$(A_3, m, n : -\overline{\Lambda}_k) \to (A_1 : {}^{(1)}(-\overline{\Lambda}_{k+1}), T^{r-2}, T+1, T-1). \tag{5.19}$$

Démonstration. Nous allons démontrer ici la relation (5.16). Les relations (5.17), (5.18) et (5.19) peuvent être démontrées de manière analogue. On raisonne par récurrence sur k.
Nous commençons par le cas où $k = 2$. Par définition :

$$\Lambda_2 = T, -T^r + 1, -T$$

et

$$\Lambda_3 = T+1, T-1, -T^{r^2-2}, -T, T^r, T.$$

En appliquant les formules du lemme 5.3.1, on a :

$$\begin{aligned}(A_2, m, n : T) &\to (A_1, m+1, n+3 : T^{r-2}, T+1, T-1)\\ (A_1, m+1, n+3 : -T^r+1) &\to (A_5, m+2, n+5 : -T^{r^2-2}, -T)\\ (A_5, m+2, n+5 : -T) &\to (A_6 : -T^r, T)\end{aligned}$$

Par conséquent, (5.16) est prouvé pour $k = 2$.
Nous supposons à présent que k est un entier pair, $k > 2$, et que la relation (5.16) est vraie pour $k-2$. Nous allons la montrer maintenant pour k.
Par définition,

$$\Lambda_k = \Lambda_{k-2}, -T^{\lambda_{k-1}}, \Gamma_{k-2}.$$

Par l'hypothèse de récurrence,

$$(A_2, m, n : \Lambda_{k-2}) \to (A_6, m + |\Lambda_{k-2}|, n + |\Lambda_{k-1}| + 1 : T^{r-2}, \Lambda_{k-1}).$$

En appliquant les lemmes 5.3.1 et 5.3.2 nous avons aussi :

$$\begin{aligned}(A_6, m + |\Lambda_{k-2}|, n + |\Lambda_{k-1}| + 1 : -T^{\lambda_{k-1}}) &\to (A_5 : -T^{\lambda_k}, -T)\\ (A_5, m + |\Lambda_{k-2}| + 1, n + |\Lambda_{k-1}| + 3 : \Gamma_{k-2}) &\to (A_6 : {}^{(1)}\Gamma_{k-1}).\end{aligned}$$

En réunissant tous ces relations et en tenant en compte que, pour tous les k pairs, Γ_{k-1} commence par $-T$, on obtient le résultat. □

Démonstration de la proposition 5.3.2. Nous allons prouver par récurrence sur k que, pour tous m, n tels que $m < n$, on a les relations suivantes :

$$(A_1, m, n : \Omega_k) \to \left(A_2 : {}^{(3)}\Omega^{(1)}_{k+1}\right), \tag{5.20}$$

$$(A_1, m, n : -\overline{\Omega}_k) \to \left(A_2 : {}^{(3)}(-\overline{\Omega}_{k+1})^{(1)}\right). \tag{5.21}$$

Lorsque $m = 1, n = 4$, la relation (5.20) implique la proposition 5.3.2. On commence par le cas où $k = 1$. On a :

$$\begin{aligned}\Omega_1 &= -T^{r-2}, -\overline{\Omega_1} = T^{r-2},\\ \Omega_2 &= -T^{r-2}, T+1, T-1, -T^{\omega_2}, -T+1, -T-1, T^{r-2},\\ -\overline{\Omega}_2 &= -T^{r-2}, T+1, T-1, T^{\omega_2}, -T+1, -T-1, T^{r-2}.\end{aligned}$$

D'autre part, le lemme 5.3.1 implique que pour tous m, n, $m < n$:

$$(A_1, m, n : \pm T^{r-2}) \to (A_2 : \pm T^{r(r-2)-2}, -T+1, -T-1),$$

et puisque $\omega_2 = r(r-2) - 2$ les relations (5.20) et (5.21) sont établies pour $k = 1$.
Soit $k > 1$. Nous supposons maintenant que (5.20) et (5.21) sont vraies pour $k-1$ et tous m, n, $m < n$. A cause du lemme 5.3.3, nous devons distinguer deux cas : le cas où k est pair et le cas où k est impair. Nous allons supposer

maintenant que k est pair.
Fixons $m, n \in \mathbb{N}$, $m < n$.

D'après (5.10) on peut écrire les deux relations suivantes :

$$\begin{aligned}\Omega_k &= \Omega_{k-1}, \Lambda_{k-1}, -T^{\omega_k}, -\overline{\Lambda}_{k-1}, -\overline{\Omega}_{k-1},\\ -\overline{\Omega}_k &= \Omega_{k-1}, \Lambda_{k-1}, T^{\omega_k}, -\overline{\Lambda}_{k-1}, -\overline{\Omega}_{k-1},\end{aligned}$$

afin d'appliquer notre algorithme à chaque sous-suite qui apparaît dans les expressions de Ω_k et $-\overline{\Omega}_k$.
Soient $1 \leq i, i' \leq 6$ et W, W' des suites finies de polynômes sur $\mathbb{F}_p$. Pour alléger l'écriture, nous nous permettons d'utiliser la notation :

$$(A_i : W) \to (A_{i'} : W'),$$

les indices m et n étant sous-entendus.
Par l'hypothèse de récurrence, on a :

$$\begin{aligned}&(A_1 : \Omega_{k-1}) \to (A_2 : {}^{(3)}\Omega_k^{(1)}),\\ &(A_1 : -\overline{\Omega}_{k-1}) \to \left(A_2 : {}^{(3)}(-\overline{\Omega}_k)^{(1)}\right).\end{aligned}$$

Les lemmes 5.3.3 et 5.3.1 impliquent :

$$\begin{aligned}&(A_2 : \Lambda_{k-1}) \to (A_4 : T^{r-2}, \Lambda_k),\\ &(A_4 : \pm T^{\omega_k}) \to (A_3 : \pm T^{\omega_{k+1}}, T),\\ &(A_3 : -\overline{\Lambda}_{k-1}) \to (A_1 : {}^{(1)}(-\overline{\Lambda}_k), T^{r-2}, T+1, T-1).\end{aligned}$$

En résumé, on a :

$$\begin{aligned}&(A_1 : \Omega_k) \to (A_2 : {}^{(3)}\Omega_k^{(1)}, T^{r-2},\\ &\quad \Lambda_k, -T^{\omega_{k+1}}, T, {}^{(1)}(-\overline{\Lambda}_k), T^{r-2}, T+1, T-1, {}^{(3)}(-\overline{\Omega}_k)^{(1)}),\\ &(A_1 : -\overline{\Omega}_k) \to (A_2 : {}^{(3)}\Omega_k^{(1)}, T^{r-2},\\ &\quad \Lambda_k, T^{\omega_{k+1}}, T, {}^{(1)}(-\overline{\Lambda}_k), T^{r-2}, T+1, T-1, {}^{(3)}(-\overline{\Omega}_k)^{(1)}).\end{aligned}$$

Nous remarquons que pour tous les $k > 1$, Ω_k commence par $-T^{r-2}, T+1, T-1$ et il se termine par $-T+1, -T-1, T^{r-2}$. De même, $-\overline{\Lambda}_k$ commence par T (puisque Γ_{k-2} se termine par $-T$). En combinant les relations précédentes, nous en déduisons (5.20) et (5.21).

Le cas où k est impair se traite de manière analogue. □

Nous avons observé que tout le raisonnement ci-dessus pour obtenir le développement en fraction continue de BS_r est basé sur le fait que le premier

des 3 quotients partiels de départ $(-T^{r-2}, T+1, T-1)$ est divisible par T. Ainsi, nous pouvons obtenir un résultat plus général en remplaçant $-T^{r-2}$ par un polynôme P arbitraire, divisible par T. En utilisant les mêmes notations qu'au début de ce paragraphe, nous avons le théorème ci-dessous dont la preuve s'obtient comme précédemment.

Théorème 5.3.2. *Soit $P \in T\mathbb{F}_p[T]$. On définit la suite des polynômes à coefficients dans $\mathbb{F}_p$, $(\Omega_k(P))_{k\geq 0}$, de la façon suivante.*

On pose $\Omega_1(P) := P$. Pour $k \geq 2$, $\Omega_k(P)$ est défini récursivement par :

$$\Omega_k(P) := \Omega_{k-1}(P), \Lambda_{k-1}, T^{\omega_{k-1}}(P/T)^{r^{k-1}}, -\overline{\Lambda}_{k-1}, -\overline{\Omega}_{k-1}.$$

Si z est la fraction continue infinie $z := [P, T+1, T-1, z_4]$ satisfaisant :

$$z^r = T^2 z_4 + (T+1), \tag{5.22}$$

alors la suite de quotients partiels de z est $\Omega_\infty(P)$, où $\Omega_\infty(P)$ est la suite infinie commençant par $\Omega_k(P)$, pour tout $k \geq 1$.

***Remarque* 5.3.2.** L'existence de la fraction continue z vérifiant (5.22) découle aussi du théorème 1 de l'article [82]. De plus, il est facile de voir que z satisfait l'équation algèbrique :

$$T^2 z^{r+1} = (PT^2 + T - 1)z^r + 1.$$

A Sur l'exposant d'irrationalité des séries de Baum et Sweet généralisées

Dans cette appendice, nous étudions l'approximation rationnelle des séries de Baum et Sweet généralisées. La connaissance des quotients partiels de ces séries de Laurent nous permet de déduire aisément le résultat suivant.

Théorème A.1. *L'exposant d'irrationalité de la série de Laurent vérifiant Théorème 5.3.1 satisfait*

$$r - 1 \leq \mu(f_r) \leq r + 1.$$

Démonstration. Nous allons utiliser la suite de quotients partiels $\Omega := (\Omega_n)_{n\geq 0}$, décrite dans le Chapitre précédent et, puisque nous n'avons pas une formule exacte pour le n-ième terme de cette suite (c'est-à-dire pour chaque quotient partiel), nous nous servirons d'une sous-suite de Ω.

La majoration de $\mu(f_r)$ est immédiate puisque, d'après le théorème de Liouville, l'exposant d'irrationalité de f_r est inférieur à son degré, donc à $r+1$. Pour obtenir la minoration souhaitée, il suffit d'exhiber une suite infinie de fractions rationnelles $(P_n/Q_n)_{n\geq 0}$ qui vérifie :

$$\left| f_r - \frac{P_n}{Q_n} \right| \leq \frac{1}{|Q_n|^{r-1}}.$$

Nous rappelons que $f_r(T) = [1, -T-1, \Omega_\infty]$, où la suite Ω est donnée dans la définition (5.3.3) par la formule récursive :

$$\Omega_{n+1} := \Omega_n, \Lambda_n, -T^{\omega_{n+1}}, -\overline{\Lambda}_n, -\overline{\Omega}_n. \tag{5.23}$$

Soit $n \geq 1$. Notons P_n/Q_n la suite des convergents de f_r. Nous notons également β_n la longueur du bloc de quotients partiels $1, -T-1, \Omega_n, \Lambda_n$ et nous affirmons que la sous-suite de quotients partiels $P_{\beta_n}/Q_{\beta_n} = [1, -T-1, \Omega_n, \Lambda_n]$ vérifie l'inégalité précédente. En effet, pour tout $n \geq 1$, on a :

$$\left| f_r - \frac{P_{\beta_n}}{Q_{\beta_n}} \right| = \frac{1}{|a_{\beta_n+1}||Q_{\beta_n}|^2},$$

où a_{β_n+1} est le β_n+1-ème quotient partiel du développement en fraction continue de f_r ; d'après la formule (5.23), $a_{\beta_n+1} = -T^{\omega_{n+1}}$. Puisque

$$\deg Q_{\beta_n} = \sum_{k=1}^{\beta_n} \deg a_i,$$

il est facile d'obtenir, d'après les formules récursives qui définissent les suites (Ω_n), (Λ_n) et ω_n, que

$$\deg Q_{\beta_n} = \frac{r^{n+1}-1}{r-1},$$

$$\omega_n = r^n - 2\frac{r^n-1}{r-1}.$$

Par conséquent,

$$|a_{\beta_n+1}||Q_{\beta_n}|^2 = r^{n+1}$$

et donc

$$\left| f_r - \frac{P_{\beta_n}}{Q_{\beta_n}} \right| = \frac{1}{|T|^{r^{n+1}}} \leq \frac{1}{|T|^{r^{n+1}-1}} = \frac{1}{|Q_{\beta_n}|^{r-1}}.$$

On en déduit ainsi que $\mu(f_r) \geq r-1$, ce qui termine la démonstration. □

***Remarque* A.1.** Une étude plus fine de la suite Ω permettrait sans doute d'obtenir à partir du théorème 5.3.1 la valeur exacte de l'exposant d'irrationalité de f_r.

6

Rational approximation for algebraic Laurent series

Ce chapitre est consacré à l'étude de l'exposant d'irrationalité des séries de Laurent algébriques sur le corps des fractions rationnelles $\mathbb{F}_p(T)$. En utilisant une approche analogue à celle de [8], nous donnons, dans un premier temps, une majoration générale de l'exposant d'irrationalité de ces séries de Laurent. Pour ce faire, nous cherchons classiquement à construire une suite de bonnes approximations rationnelles, en utilisant le théorème de Christol, puis le théorème de Cobham sur les morphismes uniformes. Dans la seconde partie de ce chapitre, nous introduisons une nouvelle approche permettant d'améliorer, dans de nombreux cas, la borne précédemment obtenue. Plus exactement, nous décrivons un algorithme qui permet de vérifier si la suite d'approximations rationnelles construite satisfait à une certaine propriété de coprimalité. Enfin, nous illustrons dans la dernière partie, notre approche à l'aide de quelques exemples. Nous utilisons en particulier cet algorithme pour calculer la valeur exacte de l'exposant d'irrationalité de plusieurs séries de Laurent algébriques. Le contenu de ce chapitre fait l'objet d'un article en cours de rédaction.

6.1 Introduction

One of the basic question in Diophantine approximation is how well real numbers can be approximated by rationals. The theory of rational approximation

of real numbers has then been transposed to function fields thanks to the pioneering works of Maillet [92] in 1906 and Gill [68] in 1930. In the present work, we are interested in Diophantine approximation for Laurent series with coefficients in a finite field. More precisely, our aim is to study the irrationality exponent of Laurent series that are algebraic over the field of rational fractions.

All along this chapter, p is a prime number and q is a power of p. We will let $\mathbb{F}_q(T)$, $\mathbb{F}_q[[T^{-1}]]$ and $\mathbb{F}_q((T^{-1}))$ denote, respectively, the field of rational functions, the ring of formal series and the field of Laurent series over the finite field $\mathbb{F}_q$. We also consider the absolute value defined on $\mathbb{F}_q(T)$ by

$$|P/Q| = e^{\deg P - \deg Q},$$

for $P, Q \in \mathbb{F}_q[T]$. The field of Laurent series in $1/T$, usually denoted by $\mathbb{F}_q((T^{-1}))$ should be seen as a completion of the field $\mathbb{F}_q(T)$ for this absolute value. Thus, if f is a nonzero element of $\mathbb{F}_q((T^{-1}))$ defined by

$$f(T) = a_{i_0}T^{-i_0} + a_{i_0+1}T^{-i_0-1} + \cdots,$$

where $i_0 \in \mathbb{Z}, a_i \in \mathbb{F}_q, a_{i_0} \neq 0$, we have $|f| = e^{-i_0}$.

We recall that the irrationality exponent (or measure) of a given Laurent series f, denoted by $\mu(f)$, is the supremum of the real numbers τ for which the inequality

$$\left| f - \frac{P}{Q} \right| < \frac{1}{|Q|^\tau}$$

has infinitely many solutions $(P, Q) \in \mathbb{F}_q[T]^2$, $Q \neq 0$. Thus, $\mu(f)$ measures the quality of the best rational approximations to f.

In general, it is a difficult problem to find irrationality measures for a given Laurent series. Concerning rational approximations to algebraic real numbers, a fundamental result is the so–called Liouville's inequality. In 1949, Mahler [91] observed that a similar result holds for Laurent series over a field of positive characteristic.

Theorem 6.1.1. *(Mahler, 1949) Let $\mathbb{K}$ be a field of positive characteristic and $f \in \mathbb{K}((T^{-1}))$ be an algebraic series over $\mathbb{K}(T)$ of degree $n > 1$. Then, there exists a positive real number C such that*

$$\left| f - \frac{P}{Q} \right| \geq \frac{C}{|Q|^n},$$

for all $P, Q \in \mathbb{K}(T)$, with $Q \neq 0$.

In other words, Mahler's theorem tells us that the irrationality exponent of an algebraic irrational Laurent series is at most equal to its degree. In the case of real numbers, Liouville's theorem was improved by the works of Thue, Siegel, Dyson and others, leading to the famous Roth's theorem. In 1960, Uchiyama obtained an analogue of Roth's theorem [131], for the case of

Laurent series with coefficients in a field of characteristic 0. More precisely, if $\mathbb{K}$ is a field of characteristic 0 and f is an algebraic element of $\mathbb{K}((T^{-1}))\setminus\mathbb{K}(T)$, then, for every $\varepsilon > 0$, we have

$$\left| f - \frac{P}{Q} \right| \geq \frac{1}{|Q|^{-(2+\varepsilon)}},$$

for all $P, Q \in \mathbb{K}(T)$, with $|Q|$ sufficiently large.

On the other hand, it is well-known that, when the base field has positive characteristic, there is no analogue of Roth's theorem. In fact, the Mahler theorem is optimal. In order to see this, it is sufficient to consider the element $f \in \mathbb{F}_q((T^{-1}))$ defined by $f(T) = \sum_{i\geq 0} T^{-q^i}$. It is not difficult to see that f is an algebraic series of degree q (since it verifies the equation $f^q - f + T^{-1} = 0$) and that the irrationality exponent of f is equal to q. Note that this example is known as the Mahler's algebraic Laurent series. In the same direction, Osgood [99] and Baum and Sweet [24] gave examples of algebraic series of various degrees for which Liouville's bound is the best possible.

For a special class of algebraic series, the bound given by Liouville for the irrational exponent was improved by Osgood [99, 100]. In 1976, this author proved an analog of Thue's theorem for algebraic Laurent series which are not solutions of a rational Ricatti differential equation. In 1996, de Mathan and Lasjaunias [84] proved that Thue's theorem actually holds for every algebraic Laurent series in $\mathbb{K}((T^{-1}))$, $\mathbb{K}$ being an arbitrary field of characteristic p, which satisfies no equation of the form $f = (Af^{p^s} + B)/(Cf^{p^s} + D)$, where $A, B, C, D \in \mathbb{K}[T]$, not all zero. Laurent series satisfying such an equation are called hyperquadratic and they were studied by many authors [79, 110, 123, 134]. Note that every hyperquadratic Laurent series does satisfy a Ricatti differential equation.

Continued fractions are naturally used in this setting since the best rational approximations are obtained by truncating the continued fraction expansion. Unfortunately, it is generally difficult to obtain the continued fraction expansion of a given algebraic Laurent series. Indeed, the effect of basic operations such as addition or multiplication is not clear at all on the continued fraction expansion. The aim of this chapter is to introduce a different approach which is based on the use of the Laurent series expansion.

As a starting point, we use Christol's theorem [46] which describes in terms of automata the Laurent series expansion of formal power series that are algebraic over $\mathbb{F}_q(T)$. More precisely, we recall that $f(T) = \sum_{i\geq 0} a_i T^{-i} \in \mathbb{F}_q[[T^{-1}]]$ is algebraic over $\mathbb{F}_q(T)$ if, and only if, the sequence $(a_i)_{i\geq 0}$ is generated by a p-automaton. Furthermore, we recall that by a classical result of Eilenberg [62] the so-called p-kernel of a p-automatic sequence is always finite. More details about these notions can be found in the first part of this thesis. Our main result is the following explicit general upper bound for the irrationality exponent of algebraic Laurent series in $\mathbb{F}_q((1/T))$.

Theorem 6.1.2. *Let $f(T) = \sum_{i \geq -k} a_i T^{-i}$ be a Laurent series with coefficients in a finite field of characteristic p, algebraic over $\mathbb{F}_q(T)$. Let s be the cardinality of the p-kernel of $\boldsymbol{a} = (a_i)_{i \geq 0}$ and d be the number of states of the minimal automaton generating $\boldsymbol{a}$ (in direct reading). Then the irrationality exponent $\mu(f)$ satisfies*

$$\mu(f) \leq p^{s+1} d. \tag{6.1}$$

The approach we use to prove Theorem 6.1.2 already appears in a different framework in [8]. It is essentially based on repetitive patterns occurring in automatic sequences. More precisely, each algebraic formal series $f(T) = \sum_{i \geq 0} a_i T^{-i}$ is identified with a p-automatic sequence $\mathbf{a} := (a_i)_{i \geq 0}$ over $\mathbb{F}_q$. Then we use a theorem of Cobham which characterizes p-automatic sequences in terms of p-morphisms (see Section 6.2.3). As a consequence of this result and of the pigeonhole principle, we are able to find infinitely many pairs of finite words (U_n, V_n) and a real number $\omega > 1$ such that $U_n V_n^\omega$ is prefix of $\mathbf{a}$ for every positive integer n. Furthermore, the length of U_n and V_n are respectively of the form kp^n and ℓp^n. Hence, there exists an infinite sequence of pairs of polynomials (P_n, Q_n) such that the Laurent series expansion of the rational function P_n/Q_n is the (ultimately periodic) sequence $\mathbf{c}_n := U_n V_n^\infty$. The sequence of rational functions P_n/Q_n provides good rational approximations to f since the words $\mathbf{a}$ and $\mathbf{c}_n$ have the common prefix $U_n V_n^\omega$. Using such approximations we are able to prove the following result (see Theorem 6.2.1):

$$\frac{k + \omega\ell}{k + \ell} \leq \mu(f) \leq \frac{p^{s+1}(k + \ell)}{(\omega - 1)\ell}. \tag{6.2}$$

In practice, it may happen that we can choose U_n and V_n such that $\mathbf{a}$ and $\mathbf{c}_n$ have the same first $(k + \omega\ell)p^n$ digits, while the $(k + \omega\ell)p^n + 1$th are different. In this case, the upper bound we obtain for the irrationality exponent may be a much better one and in particular does not depend anymore on the cardinality of the p-kernel. Furthermore, in the case where we can prove that $(P_n, Q_n) = 1$ for all n large enough, we obtain a significative improvement on the upper bound, as it will be explained in Remark 6.2.1. Note that, when working with similar constructions involving real numbers, it is well-known that this coprimality assumption is usually very difficult to check.

In the second part of this chapter, we introduce a new approach in order overcome this difficulty in our setting. We provide an algorithm that allows us to check, in a finite amount of time, wether the sequences of polynomials $(P_n)_{n \geq 0}$ and $(Q_n)_{n \geq 0}$, associated with an algebraic series f, are relatively prime for all n large enough. In order to do this, we observed that the rational approximations we have obtained have a very specific form: the roots of Q_n can only be 0 and the ℓth roots of unity (see Section 6.2.4). Then we have to develop a calculus allowing to compute the polynomials $P_n(T)$. In order to do this, we introduce some matrices associated with p-morphisms. These matrices generalize the so–called incidence matrix of the underlying morphism

(see Section 6.3). This could be of independent interest.

In the last part of this chapter, we illustrate the relevance of our approach with few examples. We give in particular several algebraic Laurent series for which we are able to compute the exact value of the irrationality exponent.

6.2 Proof of Theorem 6.1.2

Theorem 6.1.2 is an easy consequence of the more precise result established in Theorem 6.2.1. In order to prove Theorem 6.2.1, we first establish our approximation lemma which is the analog of a classical result in Diophantine approximation (Lemma 6.2.2). Then we show how to construct, starting with an arbitrary algebraic Laurent series f with coefficients in a finite field, an infinite sequence of rational approximations of f satisfying the assumptions of our approximation lemma. We thus deduce the expected upper bound for the irrationality exponent of f.

All along this section, we provide comments and remarks allowing one to improve in most cases this general upper bound (see in particular Remark 6.2.1 and Section 6.2.4).

6.2.1 Maximal repetitions in automatic sequences

Before stating our approximation lemma, we first recall a useful result, which will allow us later to control repetitive patterns occurring as prefixes of automatic sequences. The proof of the following lemma can be found in [8] (Lemma 5.1, page 1356). Before stating it, we recall that the kernel K_k of a k-automatic sequence $\mathbf{a} = (a_i)_{i\geq 0}$ is defined as the set of subsequences of the form $(a_{k^n i+l})_{i\geq 0}$, where $n \geq 0$ and $0 \leq l < k^n$. Furthermore, we recall that by a result of Eilenberg a sequence $\mathbf{a}$ is k-automatic if, and only if, $K_k(\mathbf{a})$ is finite.

Lemma 6.2.1. *Let $\boldsymbol{a}$ be a non-ultimately periodic k-automatic sequence defined on a alphabet $\mathcal{A}$. Let $U \in \mathcal{A}^*, V \in \mathcal{A}^* \setminus \{\varepsilon\}$ and $\omega \in \mathbb{Q}$ be such that UV^ω is a prefix of the sequence $\boldsymbol{a}$. Let s be the cardinality of the k-kernel of $\boldsymbol{a}$. Then we have the following inequality:*

$$\frac{|UV^\omega|}{|UV|} < k^s.$$

6.2.2 An approximation lemma

We start with the following result which is, in fact, an analog version of Lemma 4.1 in [9] for Laurent series with coefficients in a finite field. We also recall the proof, since it is not very long and it may be of independent interest.

Lemma 6.2.2. *Let $f(T)$ be a Laurent series with coefficients in $\mathbb{F}_q$. Let δ, ρ and θ be real numbers such that $0 < \delta \leq \rho$ and $\theta \geq 1$. Let us assume that there*

exists a sequence $(P_n/Q_n)_{n\geq 1}$ *of rational fractions with coefficients in* $\mathbb{F}_q$ *and some positive constants* c_0, c_1 *and* c_2 *such that:*

(i)
$$|Q_n| < |Q_{n+1}| \leq c_0|Q_n|^\theta$$

(ii)
$$\frac{c_1}{|Q_n|^{1+\rho}} \leq |f - \frac{P_n}{Q_n}| \leq \frac{c_2}{|Q_n|^{1+\delta}}.$$

Then, the irrationality measure $\mu(f)$ *satisfies*

$$1 + \delta \leq \mu(f) \leq \frac{\theta(1+\rho)}{\delta}. \tag{6.3}$$

Furthermore, if we assume that there is $N \in \mathbb{N}^*$ *such that for any* $n \geq N$, $(P_n, Q_n) = 1$, *then, we have*

$$1 + \delta \leq \mu(f) \leq \max(1 + \rho, 1 + \frac{\theta}{\delta}).$$

Proof. The left-hand side inequality is clear. We thus turn our attention to the second inequality. Let $P/Q \in \mathbb{F}_q(T)$ such that $|Q|$ is large enough. Then there exists a unique integer $n = n(Q) \geq 2$ such that

$$|Q_{n-1}| < (2c_2|Q|)^{\frac{1}{\delta}} \leq |Q_n|. \tag{6.4}$$

If $P/Q \neq P_n/Q_n$ then

$$\left|\frac{P}{Q} - \frac{P_n}{Q_n}\right| \geq \frac{1}{|QQ_n|},$$

and using Eq. (6.4) and (ii) we get that

$$\left|f - \frac{P_n}{Q_n}\right| \leq \frac{c_2}{|Q_n|^{1+\delta}} = \frac{c_2}{|Q_n||Q_n|^\delta} \leq \frac{1}{2|Q||Q_n|}.$$

By the triangle inequality, we have that

$$\left|f - \frac{P}{Q}\right| \geq \left|\frac{P}{Q} - \frac{P_n}{Q_n}\right| - \left|f - \frac{P_n}{Q_n}\right|.$$

Eq. (i) together with Eq. (6.4) imply that $|Q_n| \leq c_0|Q_{n-1}|^\theta < c_0(2c_2|Q|)^{\theta/\delta}$. Thus,

$$\left|f - \frac{P}{Q}\right| \geq \frac{1}{2|Q||Q_n|} \geq \frac{1}{2|Q|c_0(2c_2|Q|)^{\theta/\delta}} \geq \frac{c_3}{|Q|^{\frac{\theta(1+\rho)}{\delta}}}$$

since $1 + \theta/\rho \leq \theta + \theta/\rho$ (because $\theta \geq 1$) and $c_3 := 1/2c_0(2c_2)^{\theta/\delta}$.

On the other hand, if $P/Q = P_n/Q_n$, then

$$\left|f - \frac{P}{Q}\right| = \left|f - \frac{P_n}{Q_n}\right| \geq \frac{c_1}{|Q_n|^{1+\rho}} \geq \frac{c_1}{(c_0(2c_2|Q|)^{\theta/\rho})^{1+\rho}} = \frac{c_4}{|Q|^{\frac{\theta(1+\rho)}{\delta}}},$$

where $c_4 = c_1/c_0^{1+\rho}(2c_2)^{\frac{\theta(1+\rho)}{\delta}}$. This ends the proof.

The case where $(P_n, Q_n) = 1$ is treated in a similar way and we refer the reader to Lemma 4.1 in [9] (page 10). The proof consists, as previously, of two cases, but, when $P_n/Q_n = P/Q$ and P/Q is reduced, then $Q_n = Q$; this permits to obtain an improved upper bound. □

Note that the second part of Lemma 6.2.2 is also known as a Voloch's Lemma and for more details, we refer the reader to [134].

6.2.3 Construction of rational approximations via Christol's theorem

Let

$$f(T) := \sum_{i\geq 0} a_i T^{-i} \in \mathbb{F}_q[[T^{-1}]]$$

be an irrational algebraic Laurent series over $\mathbb{F}_q(T)$.

We recall that, by Christol's theorem, the sequence $\mathbf{a} := (a_i)_{i\geq 0}$ is p-automatic. According to Cobham's theorem, there exist $m \geq 1$,

$$\sigma : \mathcal{A}_m \mapsto \mathcal{A}_m^*$$

a p-morphism and

$$\varphi : \mathcal{A}_m \mapsto \mathbb{F}_q$$

a coding such that $\mathbf{a} = \varphi(\sigma^\infty(a))$ where $a \in \mathcal{A}_m$.

In all that follows, we let $f_{\mathbf{a}}(T) = \sum_{i\geq 0} a_i T^{-i}$ denote the formal series associated with the infinite word $\mathbf{a} = (a_i)_{i\geq 0}$.

We also give the following definition for a polynomial associated with a finite word.

Definition 6.2.1. *Let $U = a_0 a_1 \cdots a_{k-1}$ be a finite word over a finite alphabet. We associate with U the polynomial $P_U(T) := \sum_{j=0}^{k-1} a_{k-1-j} T^j$. If $U = \varepsilon$, we set $P_U(T) = 0$.*

For example, if we consider the word $U = 1020031 \in \mathcal{A}_5$ then

$$P_U(T) = T^6 + 2T^4 + 3T + 1$$

is a polynomial with coefficients in $\mathbb{F}_5$.

Using this notation, we have the following two lemmas.

Lemma 6.2.3. *Let U, V be two finite words such that $|U| = k \in \mathbb{N}$ and $|V| = \ell \in \mathbb{N}$ and let $\boldsymbol{a} := UV^\infty$. Then,*

$$f_{\boldsymbol{a}}(T) = \frac{P_U(T)(T^\ell - 1) + P_V(T)}{T^{k-1}(T^\ell - 1)}.$$

If $k = 0$, we have

$$f_{\boldsymbol{a}}(T) = \frac{T P_V(T)}{T^\ell - 1}.$$

Proof. Let $U := a_0a_1\cdots a_{k-1}$ and $V := b_0b_1\cdots b_{\ell-1}$. Writing the associated series with $\mathbf{a} := UV^\infty$, we have

$$f_{\mathbf{a}}(T) = (a_0 + a_1T^{-1} + \cdots + a_{k-1}T^{-(k-1)}) + (b_0T^{-k} + \cdots + b_{\ell-1}T^{-(k+l-1)}) + \cdots$$

and then, factorizing $T^{-(k-1)}, T^{-k}, T^{-(k+\ell)}, T^{-(k+2\ell)}, \ldots$ and using the definition of $P_U(T)$ and $P_V(T)$ (see Def. (6.2.1)), we obtain

$$\begin{aligned} f_{\mathbf{a}}(T) &= T^{-(k-1)}P_U(T) + T^{-k}(b_0 + b_1T^{-1} + \cdots + b_{\ell-1}T^{-(\ell-1)}) + \\ &\quad + T^{-(k+\ell)}(b_0 + \cdots + b_{\ell-1}T^{-(\ell-1)}) + \cdots \\ &= T^{-(k-1)}P_U(T) + T^{-(k+\ell-1)}P_V(T)(1 + T^{-\ell} + T^{-2\ell} + T^{-3\ell} + \cdots) \\ &= \frac{P_U(T)(T^\ell - 1) + P_V(T)}{T^{k-1}(T^\ell - 1)}. \end{aligned}$$

When $k = 0$, that is, when $U = \varepsilon$, we just have to replace $k = 0$ in the identity above. □

Lemma 6.2.4. *Let $\boldsymbol{a} = (a_i)_{i\geq 0}$ and $\boldsymbol{b} = (b_i)_{i\geq 0}$ two infinite sequences over a finite alphabet, satisfying $a_i = b_i$ for $0 \leq i \leq L-1$, where $L \in \mathbb{N}^*$. Then, we have*

$$|f_{\boldsymbol{a}} - f_{\boldsymbol{b}}| \leq \frac{1}{e^L}.$$

the equality being obtained when $a_L \neq b_L$.

Proof. This lemma immediately follows from the definition of an ultrametric norm. □

We now construct a sequence of rational fractions $(P_n/Q_n)_{n\geq 0}$ that satisfies the assumptions of Lemma 6.2.2. The approach we use appears in [8] and is essentially based on the repetitive patterns occurring in automatic sequences.

The sequence $\mathbf{a}$ being p-automatic, the p-kernel is finite. We let d denote the number of states of the minimal automaton generating $\mathbf{a}$ (in *direct reading*) and s the cardinality of the p-kernel. Let consider a prefix P of $\sigma^\infty(a)$ of length $d+1$. Observe that d is greater than or equal to the cardinality of the internal alphabet of $\mathbf{a}$, that is, $\mathcal{A}_m$. It follows, from the pigeonhole principle, that there exists a letter $b \in \mathcal{A}_m$ occurring at least twice in P. This means that there exists two (possibly empty) words U' and V' and a letter b (all defined over $\mathcal{A}_m$) such that $P := U'bV'b$. Now, if we set $U := U'$, $V := bV'$, $|U| := k$, $|V| := \ell$, $\omega := 1 + 1/\ell$, then UV^ω is a prefix of $\sigma^\infty(a)$.

Let $n \in \mathbb{N}$, $U_n := \varphi(\sigma^n(U))$ and $V_n := \varphi(\sigma^n(V))$. Since

$$\mathbf{a} = \varphi(\sigma^\infty(a))$$

then, for any $n \in \mathbb{N}$, $U_nV_n^\omega$ is a prefix of $\mathbf{a}$. Notice also that $|U_n| = |U|p^n$ and $|V_n| = |V|p^n$ and the sequence $(|V_n|)_{n\geq 1}$ is increasing.

For any $n \geq 1$, we set

$$Q_n(T) = T^{kp^n-1}(T^{\ell p^n} - 1). \tag{6.5}$$

Let $\mathbf{c}_n$ denote the infinite word $U_n V_n^\infty$. There exists $P_n(T) \in \mathbb{F}_q[T]$ such that

$$f_{\mathbf{c}_n}(T) = \frac{P_n(T)}{Q_n(T)}.$$

More precisely, by Lemma 6.2.3, the polynomial $P_n(T)$ may be defined by the following formula

$$P_n(T) = P_{U_n}(T)(T^{\ell p^n} - 1) + P_{V_n}(T). \tag{6.6}$$

Since $\mathbf{a}$ and $\mathbf{c}_n$ have the common prefix $U_n V_n^\omega$ it follows by Lemma 6.2.4 that

$$\left| f - \frac{P_n}{Q_n} \right| \leq \frac{c_2}{|Q_n|^{\frac{k+\omega\ell}{k+\ell}}}, \tag{6.7}$$

where $c_2 = e^{\frac{k+l}{k+\omega l}}$.

Furthermore, if $\mathbf{a}$ and $\mathbf{c}_n$ have the common prefix $U_n V_n^\omega$ and the $(k+\omega\ell)p^n+1$-th letters are different, then, by Lemma 6.2.4, Inequality (6.7) becomes an equality. On the other hand, we also have the following result, which is an easy consequence of Lemma 6.2.1.

Lemma 6.2.5. *Let s be the cardinality of the p-kernel of the sequence $\boldsymbol{a} := (a_i)_{i\geq 0}$. We have*

$$\left| f - \frac{P_n}{Q_n} \right| \geq \frac{1}{|Q_n|^{p^s}}.$$

Proof. Using Lemma 6.2.1 we obtain

$$|U_n V_n^\omega| < p^s |U_n V_n|.$$

This implies that $\mathbf{a}$ and $\mathbf{c}_n$ cannot have the same first $p^s|U_nV_n|$ digits. Hence

$$\begin{aligned} \left| f - \frac{P_n}{Q_n} \right| &\geq \frac{1}{e^{(|U_n|+|V_n|)p^s}} \\ &= \frac{1}{(e|Q_n|)^{p^s}} \geq \frac{c_1}{|Q_n|^{p^s}}. \end{aligned}$$

where $c_1 := 1/e^{p^s}$. □

This shows that $(P_n/Q_n)_{n\geq 1}$ satisfies the assumptions of Lemma 6.2.2 with $\theta = p$, $\rho = p^s - 1$ and $\delta = \frac{(\omega-1)\ell}{k+\ell}$. With this notation, we obtain the following theorem.

Theorem 6.2.1. *Let $f(T) = \sum_{i\geq 0} a_i T^{-i} \in \mathbb{F}_q[[T^{-1}]]$ be an algebraic Laurent series over $\mathbb{F}_q(T)$. Let k, l, ω, s be the parameters of f defined above. Then, the irrationality measure of f, $\mu(f)$, satisfies the following inequality:*

$$\frac{k+\omega\ell}{k+\ell} \leq \mu(f) \leq \frac{p^{s+1}(k+\ell)}{(\omega-1)\ell}. \tag{6.8}$$

***Remark* 6.2.1.** If, for every n, $\mathbf{a}$ and $\mathbf{c}_n$ have the same first $(k+\omega\ell)p^n$ digits, while the $(k+\omega\ell)p^n+1$-th digits are different, then, we have

$$\left| f - \frac{P_n}{Q_n} \right| = \frac{c_2}{|Q_n|^{1+\delta}}.$$

Thus, the inequality (6.8) does not depend on s (the cardinality of p-kernel) anymore. More precisely, in this case, P_n and Q_n satisfy Lemma 6.2.2 with $\theta = p$, $\rho = \delta = \frac{(\omega-1)\ell}{k+\ell}$ and we have

$$\frac{k+\omega\ell}{k+\ell} \leq \mu(f) \leq \frac{p(k+\omega\ell)}{(\omega-1)\ell}.$$

Moreover, if there exists N such that, for any $n \geq N$, $(P_n, Q_n) = 1$, then

$$\frac{k+\omega\ell}{k+\ell} \leq \mu(f) \leq \max\left(\frac{k+\omega\ell}{k+\ell}, 1 + \frac{p(k+\ell)}{(\omega-1)\ell}\right).$$

If $U = \varepsilon$, that is $k = 0$, then

$$\omega \leq \mu(f) \leq p\frac{\omega}{\omega-1}.$$

Furthermore, if $\omega - 1 \geq \sqrt{p}$ and $(P_n, Q_n) = 1$, then $\mu(f) = \omega$. All this explains why, in many cases, the general upper bound we obtained in Theorem 6.1.2 can be significatively improved.

Proof of Theorem 6.1.2. By construction, $\omega = 1 + 1/\ell$ and $k + \ell \leq d$. By Theorem 6.2.1, it immediately follows that $\mu(f) \leq p^{s+1}d$. □

6.2.4 An equivalent condition for coprimality of P_n and Q_n

We have seen in Remark 6.2.1 that, in the case where the numerator P_n and the denominator Q_n of our rational approximations are relatively prime, the bound for the irrationality exponent obtained in Theorem 6.1.2 can be significatively improved. This serves as a motivation for this section, which is devoted to the coprimality of P_n and Q_n.

First, let us recall the following result, which is an easy consequence of the fact that the greatest common divisor of two polynomials, defined over a field $\mathbb{K}$, also belongs to $\mathbb{K}$.

Lemma 6.2.6. *Let $P, Q \in \mathbb{F}_q[T]$. Then $(P, Q) = 1$ over $\mathbb{F}_q[T]$ if, and only if, $(P, Q) = 1$ over $\overline{\mathbb{F}}_p[T]$.*

Let $n \in \mathbb{N}^*$ and $Q_n(T) = T^{kp^n-1}(T^{\ell p^n} - 1) \in \mathbb{F}_q[T]$. Since we work in caracteristic p, we have that

$$Q_n(T) = T^{kp^n-1}(T^\ell - 1)^{p^n}.$$

Now, let P be an arbitrary polynomial with coefficients in $\mathbb{F}_q$. Then $(P, Q_n) = 1$ if, and only if, $(P(T), T) = 1$ and $(P(T), T^\ell - 1) = 1$. In other words, $(P, Q_n) = 1$, if, and only if, $P(0) \neq 0$ and $P(a) \neq 0$ for all $a \in \overline{\mathbb{F}}_p$ such that $a^\ell = 1$.

Therefore, we easily obtain the following lemma, which will simplify the study of the coprimality of polynomials P_n and Q_n, by using some properties of P_{U_n} and P_{V_n}.

Lemma 6.2.7. *Let $n \in \mathbb{N}^*$ and P_n, Q_n defined in Eq. (6.6) and (6.5). Then, $(P_n, Q_n) = 1$ over $\mathbb{F}_q(T)$ if, and only if, we have*

(i) $P_{U_n}(0) \neq P_{V_n}(0)$,

(ii) for any $a \in \overline{\mathbb{F}}_p$, such that $a^\ell = 1$, $P_{V_n}(a) \neq 0$.

6.3 Matrix associated with morphisms

The purpose of this section is to give an approach which will allow one to compute the polynomials $P_{U_n}(T)$ and $P_{V_n}(T)$, described in the previous section. In particular, we show that, if $\alpha \in \overline{\mathbb{F}}_p$, the sequences $(P_{U_n}(\alpha))_{n\geq 1}$ and $(P_{V_n}(\alpha))_{n\geq 1}$ are ultimately periodic. Lemma 6.2.7 implies that we have to test the coprimality of P_n and Q_n only for a finite number of index n.

Let $U = a_0 a_1 \cdots a_{k-1}$ be a finite word on $\mathcal{A}_m$ and let $i \in \mathcal{A}_m$. We let $\mathcal{P}_U(i)$ denote the set of positions of i in the word $\overline{U}$; or, simply, $\mathcal{P}_i$ if there is no doubt about U.

Définition 6.3.1. *We associate with U the row vector $v_U(T) = (\beta_{U,j}(T))_{0\leq j\leq m-1}$ where, for any $j \in \mathcal{A}_m$, $\beta_{U,j}$ is defined as:*

$$\beta_{U,j}(T) = \begin{cases} \sum_{l\in\mathcal{P}_j} T^l, \text{ if } j \text{ occurs in } U, \\ 0, \text{ otherwise.} \end{cases} \tag{6.9}$$

Example 6.3.1. We consider $U = 1020310 \in \mathcal{A}_5^*$. Then $\mathcal{P}_0 = \{0, 3, 5\}$, $\mathcal{P}_1 = \{1, 6\}$, $\mathcal{P}_2 = \{4\}$, $\mathcal{P}_3 = \{2\}$ and $\mathcal{P}_4 = \emptyset$. The vector associated with U is

$$v_U(T) = (1 + T^3 + T^5, T + T^6, T^4, T^2, 0).$$

We also recall that $P_U(T) = T^6 + 2T^4 + 3T^2 + T$ (see Definition 6.2.1) and we observe that

$$P_U(T) = v_U(T) \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

Définition 6.3.2. *Let $\sigma : \mathcal{A}_m \mapsto \mathcal{A}_m^*$ be a morphism. We associate with σ the $m \times m$ matrix $M_\sigma(T)$ with coefficients in $\mathbb{F}_p[T]$ defined by:*

$$M_\sigma(T) = (\beta_{\sigma(i),j}(T))_{0 \leq i,j \leq m-1}.$$

Example 6.3.2. Let $\sigma : \mathcal{A}_3 \mapsto \mathcal{A}_3^*$, $\sigma(0) = 010, \sigma(1) = 2101$ and $\sigma(2) = 00211$. Then

$$M_\sigma(T) = \begin{pmatrix} T^2+1 & T & 0 \\ T & T^2+1 & T^3 \\ T^4+T^3 & T+1 & T^2 \end{pmatrix}.$$

It is not difficult to see that such matrices satisfy some interesting general properties as claimed in the following remarks.

***Remark* 6.3.1.** The matrix $M_\sigma(1)$ is the reduction modulo p of the so-called incidence matrix associated with the morphism σ. This matrix satisfy some very nice properties and has been the subject of considerable study (see for instance [105]).

***Remark* 6.3.2.** Let σ_1 and σ_2 be two p-morphisms over $\mathcal{A}_m$. We have that

$$M_{\sigma_1 \circ \sigma_2}(T) = M_{\sigma_2}(T^p) M_{\sigma_1}(T).$$

Now, our main goal is to prove that, if $\alpha \in \overline{\mathbb{F}}_p$, the sequences $(P_{U_n}(\alpha))_{n \geq 1}$ and $(P_{V_n}(\alpha))_{n \geq 1}$ are ultimately periodic. This will be the subject of Proposition 6.3.2. In order to prove this, we will need the following auxiliary results.

Lemma 6.3.1. *Let $\sigma : \mathcal{A}_m \mapsto \mathcal{A}_m^*$ be a p-morphism and $U = a_0 \cdots a_{k-1} \in \mathcal{A}_m^*$. For any $n \in \mathbb{N}$ we denote $U_n = \sigma^n(U) = \sigma^n(a_0) \cdots \sigma^n(a_{k-1})$. We have*

$$P_{U_n}(T) = v_U(T^{p^n}) R_n(T),$$

where, for any $n \in \mathbb{N}$,

$$R_n(T) = \begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \\ \vdots \\ P_{\sigma^n(m-1)}(T) \end{pmatrix}.$$

Proof. Since $U_n = \sigma^n(a_0) \cdots \sigma^n(a_{k-1})$, we infer from the fact that σ is a p-morphism, that

$$P_{\sigma^n(U)}(T) = P_{\sigma^n(a_0)}(T) T^{(k-1) \cdot p^n} + P_{\sigma^n(a_1)}(T) T^{(k-2) \cdot p^n} + \cdots + P_{\sigma^n(a_{k-1})}(T).$$

Hence there exists a vector $S_n(T) = (s_0(T^{p^n}), s_1(T^{p^n}), \ldots, s_{m-1}(T^{p^n}))$, where $s_i(T)$, $0 \leq i \leq m-1$, are some polynomials with coefficients 0 or 1, such that

$$P_{\sigma^n(U)}(T) = S_n(T) \begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \\ \vdots \\ P_{\sigma^n(m-1)}(T) \end{pmatrix}$$

and $S_n(T) = S_0(T^{p^n})$. For $n = 0$, the equality above becomes

$$P_U(T) = S_0(T) \begin{pmatrix} 0 \\ 1 \\ \vdots \\ m-1 \end{pmatrix}.$$

By definitions 6.3.1 and 6.2.1, we deduce that $S_0(T) = v_U(T)$. This ends the proof. □

Example 6.3.3. Let σ be a 3-morphism over $\mathcal{A}_2$ and $U = 10100$. Then, for any $n \in \mathbb{N}$,

$$\sigma^n(U) = \sigma^n(1)\sigma^n(0)\sigma^n(1)\sigma^n(0)\sigma^n(0) \text{ and } v_U(T) = (T^3 + T + 1, T^4 + T^2).$$

Hence,

$$\begin{aligned} P_{\sigma^n(U)}(T) =& P_{\sigma^n(1)}(T)T^{4\cdot 3^n} + P_{\sigma^n(0)}(T)T^{3\cdot 3^n} + \\ &+ P_{\sigma^n(1)}(T)T^{2\cdot 3^n} + P_{\sigma^n(0)}(T)T^{3^n} + P_{\sigma^n(0)}(T) \\ =& (T^{3\cdot 3^n} + T^{3^n} + 1, T^{4\cdot 3^n} + T^{2\cdot 3^n}) \begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \end{pmatrix} \\ =& v_U(T^{3^n})R_n(T). \end{aligned}$$

Lemma 6.3.2. *Let $n \in \mathbb{N}$ and let σ be a p-morphism over $\mathcal{A}_m$. We have*

$$R_{n+1}(T) = M_\sigma(T^{p^n})R_n(T),$$

where $M_\sigma(T)$ is the matrix associated with σ, as in Definition 6.3.2.

Proof. Let σ be defined as follows

$$\begin{cases} \sigma(0) &= a_0^{(0)}a_1^{(0)}\cdots a_{p-1}^{(0)} \\ \sigma(1) &= a_0^{(1)}a_1^{(1)}\cdots a_{p-1}^{(1)} \\ \vdots & \vdots \\ \sigma(m-1) &= a_0^{(m-1)}a_1^{(m-1)}\cdots a_{p-1}^{(m-1)}, \end{cases}$$

where $a_i^{(j)} \in \mathcal{A}_m$, for any $i \in \{0, 1, \ldots, p-1\}$ and $j \in \{0, 1, \ldots, m-1\}$. Then, for any $j \in \{0, 1, \ldots, m-1\}$ and $n \in \mathbb{N}^*$, we have

$$\sigma^{n+1}(j) = \sigma^n(\sigma(j)) = \sigma^n(a_0^{(j)}a_1^{(j)}\cdots a_{p-1}^{(j)}) = \sigma^n(a_0^{(j)})\sigma^n(a_1^{(j)})\cdots\sigma^n(a_{p-1}^{(j)}).$$

Hence

$$P_{\sigma^{n+1}(j)}(T) = P_{\sigma^n(a_0^{(j)})}(T)T^{(p-1)p^n} + \cdots + P_{\sigma^n(a_{p-1}^{j})}(T).$$

It follows by Lemma 6.3.1 that

$$\begin{pmatrix} P_{\sigma^{n+1}(0)}(T) \\ P_{\sigma^{n+1}(1)}(T) \\ \vdots \\ P_{\sigma^{n+1}(m-1)}(T) \end{pmatrix} = (\beta_{\sigma(i),j}(T^{p^n}))_{0\le i,j\le m-1} \begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \\ \vdots \\ P_{\sigma^n(m-1)}(T) \end{pmatrix},$$

that is, $R_{n+1}(T) = M_\sigma(T^{p^n})R_n(T)$, which ends the proof.

□

***Remark* 6.3.3.** In particular, if $n = 0$ in the previous identity, we obtain the following equalities

$$\begin{pmatrix} P_{\sigma(0)}(T) \\ P_{\sigma(1)}(T) \\ \vdots \\ P_{\sigma(m-1)}(T) \end{pmatrix} = (\beta_{\sigma(i),j}(T))_{0\leq i,j\leq m-1} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ m-1 \end{pmatrix} = M_\sigma(T)R_0(T),$$

and this being true for any p-morphism σ defined over $\mathcal{A}_m$. Therefore, we observe that given a matrix M of this form, we can find only one morphism whose matrix is M.

Notice also that, if φ is a coding defined over $\mathcal{A}_m$, we have a similar identity

$$\begin{pmatrix} P_{\varphi(\sigma(0))}(T) \\ P_{\varphi(\sigma(1))}(T) \\ \vdots \\ P_{\varphi(\sigma(m-1))}(T) \end{pmatrix} = (\beta_{\varphi(\sigma(i)),j}(T))_{0\leq i,j\leq m-1} \begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \vdots \\ \varphi(m-1) \end{pmatrix},$$

The following corollaries immediately yield.

Corollary 6.3.1. *Let $n \in \mathbb{N}^*$ and let σ be a p-morphism defined on $\mathcal{A}_m$. We have*

$$R_n(T) = M_\sigma(T^{p^{n-1}})M_\sigma(T^{p^{n-2}})\cdots M_\sigma(T) \begin{pmatrix} 0 \\ 1 \\ \vdots \\ m-1 \end{pmatrix},$$

where $M_\sigma(T)$ is the matrix associated with σ, as in Definition 6.3.2.

Corollary 6.3.2. *Let σ be a p-morphism defined on $\mathcal{A}_m$. Then for any $n \in \mathbb{N}^*$ we have*

$$M_{\sigma^n}(T) = M_\sigma(T^{p^{n-1}})M_\sigma(T^{p^{n-2}})\cdots M_\sigma(T)$$

where $M_\sigma(T)$ is the matrix associated with σ, as in Definition 6.3.2.

Corollary 6.3.3. *Let $a \in \mathbb{F}_p$. Then for any $n \in \mathbb{N}$ we have $M_{\sigma^n}(a) = M_\sigma^n(a)$.*

Proposition 6.3.1. *Let p be a prime and $\sigma : \mathcal{A}_m \mapsto \mathcal{A}_m^*$ be a p-morphism. Let $\alpha \in \mathbb{F}_r$, where $r = p^t$, $t \in \mathbb{N}^*$. Then for any positive integer k we have*

$$M_{\sigma^{kt}}(\alpha) = (M_{\sigma^t}(\alpha))^k. \tag{6.10}$$

Proof. We argue by induction on k. Obviously, this is true for $k = 1$. We suppose that Eq. (6.10) is satisfied for k and we prove it for $k + 1$. Using Corollary 6.3.2 and the fact that $\alpha^r = \alpha$ we obtain that

$$\begin{aligned} M_{\sigma^{(k+1)t}}(\alpha) &= \underbrace{M_\sigma(\alpha^{p^{kt+t-1}})\cdots M_\sigma(\alpha^{p^t})}\, M_\sigma(\alpha^{p^{t-1}})\cdots M_\sigma(\alpha) \\ &= \underbrace{M_\sigma(\alpha^{p^{kt-1}})\cdots M_\sigma(\alpha)}\, \underbrace{M_\sigma(\alpha^{p^{t-1}})\cdots M_\sigma(\alpha)} \\ &= M_{\sigma^{kt}}(\alpha)M_{\sigma^t}(\alpha) = (M_{\sigma^t}(\alpha))^{k+1}. \end{aligned}$$

□

Proposition 6.3.2. *Let p be a prime and $U = a_{k-1}\cdots a_0 \in \mathcal{A}_m^*$. Let $\sigma : \mathcal{A}_m \mapsto \mathcal{A}_m^*$ be a p-morphism and $\varphi : \mathcal{A}_m \mapsto \mathbb{F}_q$ a coding. Let $\alpha \in \mathbb{F}_r$, where $r = p^t$, $r \in \mathbb{N}^*$. Then the sequence $(P_{\varphi(\sigma^n(U))}(\alpha))_{n\geq 0}$ is ultimately periodic.*

Proof. First, notice that, as in Lemma 6.3.1, we have

$$P_{\varphi(\sigma^n(U))}(T) = v_U(T^{p^n}) \begin{pmatrix} P_{\varphi(\sigma^n(0))}(T) \\ P_{\varphi(\sigma^n(1))}(T) \\ \vdots \\ P_{\varphi(\sigma^n(m-1))}(T) \end{pmatrix}.$$

By Remark 6.3.3, we have that

$$\begin{pmatrix} P_{\varphi(\sigma^n(0))}(T) \\ P_{\varphi(\sigma^n(1))}(T) \\ \vdots \\ P_{\varphi(\sigma^n(m-1))}(T) \end{pmatrix} = M_{\sigma^n}(T) \begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \vdots \\ \varphi(m-1) \end{pmatrix}.$$

Hence

$$P_{\varphi(\sigma^n(U)}(\alpha) = v_U(\alpha^{p^n}) M_{\sigma^n}(\alpha) \begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \vdots \\ \varphi(m-1) \end{pmatrix}.$$

Clearly, the sequence $(v_U(\alpha^{p^n}))_{n\geq 0}$ is periodic with a period less than or equal to t since $v_U(\alpha^{p^{n+t}}) = v_U(\alpha^{p^n})$, for any $n \in \mathbb{N}^*$. We now prove that the sequence $(M_{\sigma^n}(\alpha))_{n\geq 0}$ is ultimately periodic.
Since $\alpha \in \mathbb{F}_{p^t}$, we have $\alpha^{p^t} = \alpha$ and thus, by Corollary 6.3.2, for any k and $n \in \mathbb{N}$, we obtain

$$M_{\sigma^{n+kt}}(\alpha) = M_{\sigma^n}(\alpha) M_{\sigma^{kt}}(\alpha).$$

Therefore, by Proposition 6.3.1 we have, for any $k \in \mathbb{N}$,

$$M_{\sigma^{n+kt}}(\alpha) = M_{\sigma^n}(\alpha) M_{\sigma^{kt}}(\alpha) = M_{\sigma^n}(\alpha)(M_{\sigma^t}(\alpha))^k.$$

Since $M_{\sigma^t}(\alpha)$ is a $m \times m$ matrix with coefficients in a finite field, there exist two positive integers m_0 and n_0, $m_0 \neq n_0$ (suppose that $m_0 < n_0$) such that $M_{\sigma^t}(\alpha)^{m_0} = M_{\sigma^t}(\alpha)^{n_0}$. This implies that

$$M_{\sigma^{n+m_0 t}}(\alpha) = M_{\sigma^n}(\alpha)(M_{\sigma^t}(\alpha))^{m_0} = M_{\sigma^n}(\alpha)(M_{\sigma^t}(\alpha))^{n_0} = M_{\sigma^{n+n_0 t}}(\alpha)$$

and thus, the sequence $(M_{\sigma^n}(\alpha))_{n\geq 0}$ is ultimately periodic, with pre-period at most $m_0 t$ and period at most $(n_0 - m_0)t$. Since $(v_U(\alpha^{p^n}))_{n\geq 0}$ is periodic with period at most t, it follows that $(P_{\varphi(\sigma^n(U))}(\alpha))_{n\geq 0}$ is ultimately periodic (with pre-period at most $m_0 t$ and period at most $(n_0 - m_0)t$). This ends the proof. □

***Remark* 6.3.4.** The properties of matrix associated with mophisms, appearing in this section, are still true when replacing p-morphisms by p^s-morphisms, for any $s \in \mathbb{N}^*$.

6.4 Examples

In Theorem 6.1.2 we give a general upper bound for the irrationality exponent of algebraic Laurent series with coefficients in a finite field. In many cases, the sequence of rational approximations $(P_n/Q_n)_{n\geq 0}$ we construct turns out to satisfy the conditions (i) and (ii) of Lemma 6.2.7. This naturally gives rise to a much better estimate, as hinted in Remark 6.2.1. In this section, we illustrate this claim with few examples of algebraic Laurent series for which the irrationality exponent is exactly computed or at least well estimated.

Example 6.4.1. Let us consider the following equation over $\mathbb{F}_2(T)$

$$X^4 + X + \frac{T}{T^4+1} = 0. \tag{6.11}$$

This equation is related to the Mahler algebraic Laurent series, previously mentioned. Let $E_1 = \{\alpha \in \mathbb{F}_2((T^{-1})), \quad |\alpha| < 1\}$. We first notice that Eq. (6.11) has a unique solution f in E_1. This can be obtained by showing that the application

$$\begin{aligned} h: \quad E_1 &\longmapsto E_1 \\ X &\longmapsto X + \frac{T}{T^4+1} \end{aligned}$$

is well defined and is a contracting map from E_1 to E_1. Then the fixed point theorem in a complete metric space implies that the equation $h(X) = X$, that is, Eq. (6.11) has a unique solution in E_1. Let f denote this solution and let $f(T) := \sum_{i\geq 0} a_i T^{-i}$ (with $a_0 = 0$, since f belongs to E_1).

The second step is to find the morphisms that generate the sequence of coefficients of f, as in Cobham's theorem. Notice that there is a general method that allows one to obtain these morphisms when we know the algebraic equation. In this example, we try to describe the important steps of this method and we will give further details later.

By replacing f in Eq. (6.11) and using the fact that $f^4(T) = \sum_{i\geq 0} a_i T^{-4i}$ we easily obtain the following relations between the coefficients of f

$$a_{i+1} + a_i + a_{4i+4} + a_{4i} = 0, \tag{6.12}$$

$$a_1 = 0, a_2 = 0, a_3 = 1, \tag{6.13}$$

$$a_{i+4} + a_i = 0, \text{ if } i \not\equiv 0[4]. \tag{6.14}$$

By Eq. (6.13) and (6.14), we get that $a_{4i+1} = 0$, $a_{4i+2} = 0$ and $a_{4i+3} = 1$, for any $i \geq 0$. From Eq. (6.12) we deduce that

$$\begin{aligned} a_{16i+4} &= a_{16i+8} = a_{16i} + a_{4i}, \\ a_{16i+12} &= a_{16i+8} + 1 = a_{16i} + a_{4i} + 1. \end{aligned}$$

This implies that the 4-kernel of $\mathbf{a} := (a_i)_{i\geq 0}$ is the following set

$$K_4(\mathbf{a}) = \{(a_i)_{i\geq 0}, (a_{4i})_{i\geq 0}, (a_{16i})_{i\geq 0}, (0), (1)\}.$$

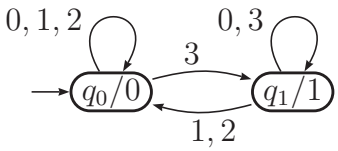


Figure 6.1: A 4-automaton recognizing **a**

Consequently, the 4-automaton generating **a** is

Once we have the automaton, there is a general approach to obtain the morphisms that generate an automatic sequence. More precisely, the proof of Cobham's theorem precisely describes this process. The reader may consult the original article of Cobham [50] or the monograph [20] (Theorem 6.3.2, page 175). Following this approach, we obtain that $\mathbf{a} = \sigma^\infty(0)$, where σ is defined as follows

$$\sigma(0) = 0001$$
$$\sigma(1) = 1001.$$

It is now possible to apply our approach described in the first part of this chapter. We will prove the following result.

Proposition 6.4.1. *One has*

$$\mu(f) = 3.$$

Notice that, Mahler's Theorem implies only that $\mu(f) \leq 4$ and Osgood's Theorem or Lasjaunias and de Mathan's Theorem cannot be applied since this Laurent series is clearly hyperquadratic.

Proof. We are going to introduce an infinite sequence of rational fractions $(P_n/Q_n)_{n\geq 1}$ converging to f. Since **a** begins with 0001, we denote $V = 0$ and for any $n \geq 1$, $V_n = \sigma^n(V)$. Hence **a** begins with $V_nV_nV_n$ for any nonnegative integer n.

Since $|V_n| = 4^n$, we set $Q_n(T) = T^{4^n} - 1$. In Section 6.2.3, we showed that there exists a polynomial $P_n(T) \in \mathbb{F}_2[T]$ such that:

$$\frac{P_n(T)}{Q_n(T)} = f_{V_n^\infty}(T).$$

The formal power series expansion of P_n/Q_n begins with

$$\sigma^n(0001)\sigma^n(0001)\sigma^n(0001)\sigma^n(0),$$

and we deduce that it begins with

$$\sigma^n(0001)\sigma^n(0001)\sigma^n(0001)0$$

while the sequence **a** begins with

$$\sigma^n(0001)\sigma^n(0001)\sigma^n(0001)1.$$

Hence, the first $3 \cdot 4^n$th digits of the series expansion of P_n/Q_n and f are the same, while the following coefficients are different. According to Remark 6.2.1, we deduce that:

$$3 \leq \mu(f) \leq 6,$$

or, if for any $n \geq 1$ we have $(P_n, Q_n) = 1$ then

$$\mu(f) = 3.$$

It thus remains to prove that $(P_n, Q_n) = 1$ for every positive integer n. Let $n \geq 1$. We now prove that $(P_n, Q_n) = 1$. By Lemma 6.2.7 we have to prove that $P_n(1) \neq 0$, *i.e.*, $P_{\sigma^n(0)}(1) \neq 0$. Remark 6.3.3 implies that

$$\begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \end{pmatrix} = M_{\sigma^n}(T) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence, $P_{\sigma^n(0)}(1) \neq 0$ if, and only if,

$$\begin{pmatrix} 1 & 0 \end{pmatrix} M_{\sigma^n}(1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \neq 0,$$

that is, if, and only if,

$$\begin{pmatrix} 1 & 0 \end{pmatrix} M_{\sigma}^n(1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \neq 0.$$

Indeed, by Corollary 6.3.3, we have $M_{\sigma^n}(1) = M_{\sigma}^n(1)$. Since

$$M_{\sigma}(1) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

we deduce that $M_{\sigma^n}(1) = M_{\sigma}(1)$ and then,

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \neq 0.$$

Consequently, we obtain that $(P_n, Q_n) = 1$. This ends the proof. □

***Remark* 6.4.1.** Let $f(T) = \sum_{i \geq 0} a_i T^{-i}$ be a formal power series with coefficients in a finite field $\mathbb{F}_q$, where q is a power of p, and suppose that there is $P(T) \in \mathbb{F}_q[T]$ such that $P(f) = 0$. As mentioned before, there is a general approach that allows one to find the automaton that generates the infinite sequence $\mathbf{a} := (a_i)_{i \geq 0}$. This consists of the following steps. First, there exists a polynomial Q with coefficients in $\mathbb{F}_q[T]$, of the form $Q(X) = \sum_{i \geq 0} B_i(T) X^{p^i}$, $B_0(T) \neq 0$ such that $Q(f) = 0$. This is known as an Ore's polynomial and the existence is due to Ore's theorem (for a proof see, for example, [20], Lemma 12.2.3, page 355). Hence, the first step is to find such a Ore's polynomial vanishing f (this is possible by passing P to the power of p as many times as we need). The second step is to find some recurrent relations between the terms

a_i in order to find the kernel of $\mathbf{a}$. This is possible thanks to the Frobenius morphism. The third step is the construction of the automaton generating $\mathbf{a}$. Notice that a sequence is k-automatic if, and only if, its k-kernel is finite. The proof of this well-known result is explicit and we refer the reader to [20] (Theorem 6.6.2, page 18). The last step is to find the morphisms generating $\mathbf{a}$, as described in Cobham's theorem. In order to do this, the reader may refer to the proof of Cobham's theorem, which is explicit as well.

The following examples present different computations of irrationality exponents of Laurent power series over finite fields. We do not give the algebraic equations of them because the computation is quite long, but, as in the previous example (where we find the morphisms if we know the equation), there is a general approach that allows one to compute the equation of a formal power series when we know the automaton generating the sequence of its coefficients. Indeed, by knowing the morphisms we can find the automaton (see the proof of Cobham's theorem), then knowing the automaton allows to find the kernel (see [20], Theorem 6.6.2, page 185) and also the relations between the coefficients. These relations allow one to find a polynomial that vanishes the formal power series (the reader may consult the proof of Christol's theorem in [46] or [20]-page 356, but also [75]-where a generalisation of Christol's theorem is given). More precisely the polynomial that we compute from these relations is also an Ore's polynomial. Finally, we have to factor this polynomial and to check which is the irreducible factor vanishing our algebraic Laurent series.

Example 6.4.2. We now consider the formal power series

$$f(T) = \sum_{i \geq 0} a_i T^{-i} \in \mathbb{F}_2[[T^{-1}]],$$

where the sequence $\mathbf{a} := (a_i)_{i \geq 0}$ is the image under the coding φ of the fixed point of the 8-uniform morphism σ, φ and σ being defined as follows

$$\begin{array}{lll} \varphi(0) = 1 & & \sigma(0) = 00000122 \\ \varphi(1) = 0 & \text{and} & \sigma(1) = 10120011 \\ \varphi(2) = 1 & & \sigma(2) = 12120021. \end{array}$$

Thus $\mathbf{a} = 11111011111 \cdots$.

Proposition 6.4.2. *One has* $\mu(f) = 5$.

Proof. We are going to introduce an infinite sequence of rational fractions $(P_n/Q_n)_{n \geq 0}$ converging to f. Since $\mathbf{a}$ begins with 00000, we denote $V = 0$ and, for any $n \geq 1$, $V_n = \sigma^n(0)$. Hence, $\mathbf{a}$ begins with V_n^5 for any $n \geq 1$. Now, we set $Q_n(T) = T^{8^n} - 1$. In Section 6.2.3 we showed that there exists a polynomial $P_n(T) \in \mathbb{F}_2[T]$ such that:

$$\frac{P_n(T)}{Q_n(T)} = f_{V_n^\infty}(T).$$

The formal power series expansion of P_n/Q_n begins with

$$\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)$$

and we deduce that it begins with

$$\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)0,$$

while the sequence $\mathbf{a}$ begins with

$$\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)\sigma^n(0)1.$$

Hence, the first $5 \cdot 8^n$th digits of the series expansion of P_n/Q_n and f are the same, while the following coefficients are different.

Using Remark 6.2.1, we deduce that

$$5 \leq \mu(f) \leq 10.$$

Moreover, if for any $n \geq 1$ we have $(P_n, Q_n) = 1$, then we obtain

$$\mu(f) = 5.$$

We now prove that $(P_n, Q_n) = 1$ for every $n \geq 1$. By Lemma 6.2.7 we have to prove that $P_n(1) \neq 0$, *i.e.*, $P_{\varphi(\sigma^n(0))}(1) \neq 0$. By Remark 6.3.3 we have that

$$\begin{pmatrix} P_{\varphi(\sigma^n(0))}(T) \\ P_{\varphi(\sigma^n(1))}(T) \\ P_{\varphi(\sigma^n(2))}(T) \end{pmatrix} = M_{\sigma^n}(T) \begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \varphi(2) \end{pmatrix}$$

and by Corollary 6.3.3

$$M_{\sigma^n}(1) = M_\sigma^n(1).$$

The matrix associated with σ, when $T = 1$, is equal to the incidence matrix of σ:

$$M_\sigma(1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Hence, by an easy computation, we obtain that, for any $j \geq 1$,

$$M_\sigma^{2j+1}(1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } M_\sigma^{2j}(1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and thus, for every $n \geq 1$, we have

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} M_\sigma^n(1) \begin{pmatrix} \varphi(0) \\ \varphi(1) \\ \varphi(2) \end{pmatrix} = 1 \neq 0.$$

Consequently, for every $n \geq 1$, $P_{\varphi(\sigma^n(0))}(1) \neq 0$ and thus $(P_n, Q_n) = 1$. □

Example 6.4.3. We now consider the formal power series

$$f(T) = \sum_{i \geq 0} a_i T^{-i} \in \mathbb{F}_3[[T^{-1}]],$$

where the sequence $\mathbf{a} := (a_i)_{i \geq 0}$ is the fixed point beginning with zero of the following 3-uniform morphism:

$$\begin{aligned} \sigma(0) &= 010 \\ \sigma(1) &= 102 \\ \sigma(2) &= 122. \end{aligned}$$

Thus $\mathbf{a} = 010102010\cdots$.

Proposition 6.4.3. *The irrationality exponent of the formal power series f satisfies*

$$2.66 \leq \mu(f) \leq 2.81.$$

In this case, we are not able to compute the exact value of the irrationality exponent but the lower bound we found shows that the degree of f is greater than or equal to 3. Hence our upper bound obviously improves on the one that could be deduced from the Mahler's theorem.

Proof. We are going to introduce an infinite sequence of rational fractions $(P_n/Q_n)_{n \geq 0}$ converging to f. Since $\mathbf{a}$ begins with 010102, we denote $V := 010102$ and for any $n \geq 1$, $V_n := \sigma^n(V)$. Hence $\mathbf{a}$ begins with

$$\sigma^n(010102)\sigma^n(010102)\sigma^n(0101)$$

for any $n \geq 1$. Now, we set $Q_n(T) = T^{2 \cdot 3^n} - 1$. There exists a polynomial $P_n(T) \in \mathbb{F}_3[T]$ such that:

$$\frac{P_n(T)}{Q_n(T)} = f_{V_n^\infty}(T).$$

The formal power series expansion of P_n/Q_n begins with

$$\sigma^n(010102)\sigma^n(010102)\sigma^n(0101)\sigma^n(0)$$

and we deduce that it begins with

$$\sigma^n(010102)\sigma^n(010102)\sigma^n(0101)0,$$

while the sequence $\mathbf{a}$ begins with

$$\sigma^n(010102)\sigma^n(010102)\sigma^n(0101)1.$$

Hence, the first $16 \cdot 3^n$th digits of the series expansion of P_n/Q_n and f are the same, while the following coefficients are different. By Remark 6.2.1, we deduce that

$$2.66 \leq \mu(f) \leq 4.81.$$

Furthermore, if for every $n \geq 1$ we have $(P_n, Q_n) = 1$, then

$$2.66 \leq \mu(f) \leq 2.81.$$

Let $n \geq 1$. We now prove that $(P_n, Q_n) = 1$. By Lemma 6.2.7, since

$$Q_n(T) = (T-1)^{3^n}(T+1)^{3^n},$$

we have to prove that, for all $n \geq 1$, $P_n(1) \neq 0$ and $P_n(-1) \neq 0$.

By definition of $(P_n(T))_{n\geq 0}$ (see Eq. (6.6)) we have

$$P_n(1) = P_{V_n}(1) \text{ and } P_n(-1) = P_{V_n}(-1).$$

Since $V_n = \sigma^n(010102) = \sigma^{n+1}(01)$ then,

$$P_{V_n}(T) = P_{\sigma^{n+1}(0)}(T)T^{3^{n+1}} + P_{\sigma^{n+1}(1)}(T).$$

Hence,

$$P_{V_n}(1) = P_{\sigma^{n+1}(0)}(1)+P_{\sigma^{n+1}(1)}(1) \text{ and } P_{V_n}(-1) = -P_{\sigma^{n+1}(0)}(-1)+P_{\sigma^{n+1}(1)}(-1).$$

Remark 6.3.3 implies that

$$\begin{pmatrix} P_{\sigma^n(0)}(T) \\ P_{\sigma^n(1)}(T) \\ P_{\sigma^n(2)}(T) \end{pmatrix} = M_{\sigma^n}(T) \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

If we now replace $T = 1$ (respectively $T = -1$), we obtain that $P_n(1) \neq 0$ (respectively $P_n(-1) \neq 0$) if, and only if,

$$\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} M_\sigma^n(1) \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \neq 0$$

(respectively,

$$\begin{pmatrix} -1 & 1 & 0 \end{pmatrix} M_\sigma^n(-1) \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \neq 0).$$

The matrix associated with σ is

$$M_\sigma(T) = \begin{pmatrix} T^2+1 & T & 0 \\ T & T^2 & 1 \\ 0 & T^2 & T+1 \end{pmatrix}.$$

Hence,

$$M_\sigma(1) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \text{ and } M_\sigma(-1) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Notice that $M_\sigma^2(\pm 1) = M_\sigma^4(\pm 1)$ and, by an easy computation, one can obtain that the sequence $(P_{V_n}(1))_{n\geq 0}$ is 2-periodic and $(P_{V_n}(-1))_{n\geq 0}$ is 1-periodic; more precisely, $(P_{V_n}(1))_{n\geq 0} = (12)^\infty$ and $(P_{V_n}(-1))_{n\geq 0} = (1)^\infty$. This proves that $P_{V_n}(1)$ and $P_{V_n}(-1)$ never vanish, which ends the proof.

□

Example 6.4.4. We now consider the formal power series

$$f(T) = \sum_{i \geq 0} a_i T^{-i} \in \mathbb{F}_5[[T^{-1}]],$$

where the sequence $\mathbf{a} := (a_i)_{i \geq 0}$ is the fixed point beginning with zero of the following 5-uniform morphism:

$$\begin{aligned}
\sigma(0) &= 00043 \\
\sigma(1) &= 13042 \\
\sigma(2) &= 14201 \\
\sigma(3) &= 32411 \\
\sigma(4) &= 00144.
\end{aligned}$$

Thus $\mathbf{a} = 0004300043 \cdots$.

Proposition 6.4.4. *One has*

$$\mu(f) = 3.4.$$

Notice that, after Mahler's Theorem, the degree of algebricity of f is greater than or equal than 4.

Proof. We are going to introduce an infinite sequence of rational fractions $(P_n/Q_n)_{n \geq 0}$ converging to f. We can remark that the sequence $\mathbf{a}$ begins with

$$00043000430004300144,$$

and thus, more generally, $\mathbf{a}$ begins with

$$\sigma^n(00043)\sigma^n(00043)\sigma^n(00043)\sigma^n(00)\sigma^n(1)$$

for every $n \geq 1$. Now, we set $Q_n(T) = T^{5^n} - 1$. There exists a polynomial $P_n(T) \in \mathbb{F}_5[T]$ such that:

$$\frac{P_n(T)}{Q_n(T)} = f_{\sigma^n(00043)^\infty}(T).$$

The formal power series expansion of P_n/Q_n begins with

$$\begin{aligned}
&\sigma^n(00043)\sigma^n(00043)\sigma^n(00043)\sigma^n(00043) = \\
&= \sigma^n(00043)\sigma^n(00043)\sigma^n(00043)\sigma^n(00)\sigma^n(0)\sigma^n(43)
\end{aligned}$$

and we deduce that it begins with

$$\sigma^n(00043)\sigma^n(00043)\sigma^n(00043)\sigma^n(00)0,$$

while the sequence $\mathbf{a}$ begins with

$$\sigma^n(00043)\sigma^n(00043)\sigma^n(00043)\sigma^n(00)1.$$

Hence, the first $17 \cdot 5^n$th digits of the series expansion of P_n/Q_n and f are the same, while the following coefficients are different.

According to Remark 6.2.1, we deduce that

$$3.4 \leq \mu(f) \leq 7.08.$$

Moreover, if for every $n \geq 1$ we have $(P_n, Q_n) = 1$, then we obtain

$$\mu(f) = 3.4.$$

Let $n \geq 1$. We are going to prove that $(P_n, Q_n) = 1$, for any $n \geq 1$. By Lemma 6.2.7, it is necessary and sufficient to prove that $P_n(1) \neq 0$, *i.e.*, $P_{\sigma^n(0)}(1) \neq 0$, which is equivalent to:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix} M_\sigma^n(1) \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \neq 0.$$

Since the matrix associated with σ is

$$M_\sigma(T) = \begin{pmatrix} T^4+T^3+T^2 & 0 & 0 & 1 & T \\ T^2 & T^4 & 1 & T^3 & T \\ T & T^4+1 & T^2 & 0 & T^3 \\ 0 & T+1 & T^3 & T^4 & T^2 \\ T^4+T^3 & T^42 & 0 & 0 & T+1 \end{pmatrix}.$$

we obtain that

$$M_\sigma(1) = \begin{pmatrix} 3 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 & 1 \\ 0 & 2 & 1 & 1 & 1 \\ 2 & 1 & 0 & 0 & 2. \end{pmatrix}.$$

It follows by a short computation (for instance, using Maple) that the sequence $(M_\sigma^n(1))_{n\geq 1}$ is 20-periodic and $(P_{V_n}(1))_{n\geq 0}$ is 4-periodic. Moreover $(P_{V_n}(1))_{n\geq 0} = (2134)^\infty$. Hence $P_n(1) \neq 0$, for any $n \geq 1$.

□

7

Subword complexity and Laurent series with coefficients in a finite field

Dans ce chapitre nous utilisons la complexité des facteurs des mots infinis pour introduire une notion de complexité pour les séries de Laurent à coefficients dans un corps fini. Nous démontrons dans un premier temps que l'inverse de la série de Laurent Π_q, analogue du nombre π dans ce contexte, a en général une complexité sous-linéaire, sauf dans le cas $q = 2$, où la complexité est quadratique. En particulier, nous en déduisons une nouvelle preuve de la transcendance de Π_2. Dans la deuxième partie de ce chapitre, nous étudions les propriétés de certaines classes de séries formelles associées à la hiérarchie induite par cette notion de complexité. Nous montrons plus précisément que la classe des séries de Laurent de complexité au plus polynômiale et celle des séries de Laurent d'entropie nulle forment un espace vectoriel sur le corps $\mathbb{F}_q(T)$, et sont laissées invariantes par diverses opérations classiques comme le produit de Hadamard. Nous observons, dans une dernière partie, qu'il semble difficile de majorer la complexité du produit usuel (produit de Cauchy) de deux séries de faible complexité. Ce travail fait l'objet d'un article soumis pour publication.

7.1 Introduction and motivations

The sequence of digits of the real number $\pi = 3.14159\cdots$ has baffled mathematicians for a long time. Though the decimal expansion of π has been

calculated to billions of digits, we do not even know, for example, if the digit 1 appears infinitely often. Actually, it is expected that, for any $b \geq 2$, the b-ary expansion of π should share some of the properties of a random sequence (see, for instance, [23]). More concretely, it is widely believed that π is normal, meaning that all blocks of digits of equal length occur in the b-ary representation of π with the same frequency, but we do not yet have a proof for this claim. A common way to describe the disorder of an infinite sequence $\mathbf{a}$ is to compute its subword complexity, which is the function that associates with each positive integer m the number of distinct blocks of length m occurring in the word $\mathbf{a}$. Let α be a positive real number and let $a_{-k}a_{-k+1}\cdots a_0.a_1a_2\cdots$ be the representation of α in an integer base $b \geq 2$. The complexity function of α is defined as follows

$$p(\alpha, b, m) = \mathrm{Card}\{(a_j, a_{j+1}, \ldots, a_{j+m-1}),\ j \in \mathbb{N}\},$$

for any positive integer m. Notice that if π were normal, then its complexity would be maximal, that is $p(\pi, b, m) = b^m$, for every $b \geq 2$ and $m \geq 1$. In this direction, similar questions have been asked about other well-known constants and it is for instance widely believed that for every $\alpha \in \{e, \log 2, \zeta(3), \sqrt{2}\}$, one should have

$$p(\alpha, b, m) = b^m,$$

for any $m \geq 1$ and $b \geq 2$.

In this chapter we use the familiar asymptotic notation of Landau. We write $f = O(g)$ if there exist two positive real numbers k and n_0 such that, for every $n \geq n_0$ we have $|f(n)| < k\,|g(n)|$. We also write $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$.

If α is a rational real number then for every integer $b \geq 2$ we have $p(\alpha, b, m) = O(1)$, where the constant depends on b and α. Moreover, there is a classical theorem of Morse and Hedlund [96] which implies that for every irrational real number α, we have

$$p(\alpha, b, m) \geq m + 1, \tag{7.1}$$

for all integers $m \geq 1$ and $b \geq 2$.

Concerning irrational algebraic numbers, the main result known to date in this direction is due to Adamczewski and Bugeaud [4]. These authors proved that the complexity of an irrational algebraic real number α satisfies

$$\lim_{m\to\infty} \frac{p(\alpha, b, m)}{m} = +\infty, \tag{7.2}$$

for any base $b \geq 2$. For more details about complexity of real numbers, see [1, 4, 7].

For classical transcendental constants, there is a more ambiguous situation and, to the best of our knowledge, the only result that improves the bound following from Inequality (7.1) was recently proved in [1]. It concerns the real number e and some other exponential periods. More precisely, Adamczewski

showed that if ξ is an irrational real number whose irrational exponent $\mu(\xi) = 2$, then

$$\lim_{m\to\infty} p(\xi, b, m) - m = +\infty,$$

for any base $b \geq 2$.

The present work is motivated by this type of question, but in the setting of Laurent series with coefficients in a finite field. In the sequel we let $\mathbb{F}_q(T)$, $\mathbb{F}_q[[T^{-1}]]$ and $\mathbb{F}_q((T^{-1}))$ denote respectively the field of rational functions, the ring of formal series and the field of Laurent series over the finite field $\mathbb{F}_q$, q being a power of a prime number p.

By analogy with the real numbers, the complexity of a Laurent series is defined as the subword complexity of its sequence of coefficients. Again, the theorem of Morse and Hedlund gives a complete description of the rational Laurent series; more precisely, they are the Laurent series of bounded complexity. Furthermore, there is a remarkable theorem of Christol [46] (see also [47]) that precisely describes the algebraic Laurent series over $\mathbb{F}_q(T)$ as follows. Let $f(T) = \sum_{n\geq -n_0} a_n T^{-n}$ be a Laurent series with coefficients in $\mathbb{F}_q$. Then f is algebraic over $\mathbb{F}_q(T)$ if, and only if, the sequence of coefficients $(a_n)_{n\geq 0}$ is p-automatic. More references on automatic sequences can be found in [20] (chapters 5 and 6). Furthermore, Cobham proved that the subword complexity of an automatic sequence is at most linear [50]. Hence, a straightforward consequence of those two results is the following.

Theorem 7.1.1. *Let $f \in \mathbb{F}_q((T^{-1}))$ be algebraic over $\mathbb{F}_q(T)$. Then we have*

$$p(f, m) = O(m).$$

The converse is obviously not true, since there are uncountably many Laurent series with linear complexity.

In contrast with real numbers, the situation is thus clarified in the case of algebraic Laurent series. Also, notice that (7.2) and Theorem 7.1.1 point out the fact that the subword complexities of algebraic elements in $\mathbb{F}_q((T^{-1}))$ and in $\mathbb{R}$ are quite different.

On the other hand, Carlitz introduced [38] functions in positive characteristic by analogy with the Riemann ζ function, the usual exponential and the logarithm function. Many values of these functions, including an analog of the real number π, were shown to be transcendental over $\mathbb{F}_q(T)$ (see [44, 69, 125, 135, 133]). In the first part of this chapter we focus on the analog of π, denoted, for each q, by Π_q, and we prove that its inverse has a "low" complexity. More precisely, we will prove the following result in Section 7.3.

Theorem 7.1.2. *Let q be a power of a prime number p. The complexity of the inverse of Π_q satisfies*

(a) if $q = 2$ then

$$p\left(\frac{1}{\Pi_2}, m\right) = \Theta(m^2);$$

(b) if $q \geq 3$ *then*

$$p\left(\frac{1}{\Pi_q}, m\right) = \Theta(m).$$

Since any algebraic Laurent series has at most linear complexity (by Theorem 7.1.1), the following corollary yields.

Corollary 7.1.1. Π_2 *is transcendental over* $\mathbb{F}_2(T)$.

The transcendence of Π_q over $\mathbb{F}_q(T)$ was first proved by Wade in 1941 (see [135]) using an analog of a classical method of transcendence in zero characteristic. Another proof was given by Yu in 1991 (see [133]), using the theory of Drinfeld modules. De Mathan and Chérif, in 1993 (see [44]), using tools from Diophantine approximation, proved a more general result, but in particular their result implied the transcendence of Π_q. Christol's theorem has also been used as a combinatorial criterion in order to prove the transcendence of Π_q. This is what is usually called an "automatic proof". The non-automaticity and also the transcendence, was first obtained by Allouche, in [12], via the so-called q-kernel. Notice that our proof here is based also by Christol's theorem, but we obtain the non-automaticity of Π_2 over $\mathbb{F}_2(T)$ as a consequence of the subword complexity.

Another motivation for this work comes from a paper of Beals and Thakur [26]. These authors proposed a classification of Laurent series by their space or time complexity. They showed that some classes of Laurent series have good algebraic properties (for instance, the class of Laurent series corresponding to any deterministic space class at least linear form a field). They also place some of Carlitz's analogs in the computational hierarchy. Furthermore, motivated by Theorems 7.1.1 and 7.1.2, we consider the classes of Laurent series of at most polynomial complexity $\mathcal{P}$ and zero entropy $\mathcal{Z}$ (see Section 7.4), which seem to be good candidates to enjoy some nice closure properties. In particular, we prove the following theorem.

Theorem 7.1.3. $\mathcal{P}$ *and* $\mathcal{Z}$ *are vector spaces over* $\mathbb{F}_q(T)$.

We will also show that both classes are closed under some usual operations such as the Hadamard product, the formal derivative and the Cartier operators. In particular, Theorem 7.1.3 provides a criterion of linear independence over $\mathbb{F}_q(T)$ (see Proposition 7.4.4).

7.2 Terminology and basic notions

In this section, we briefly recall some definitions and well-known results from combinatorics on words.

We say that a finite word V is a *subword* (or *factor*) of a finite word U if there exist some finite words A, B, possibly empty such that $U = AVB$, and we denote it by $V \triangleleft U$. Otherwise, $V \not\triangleleft U$. We say that X is a *prefix* of U, and

we denote by $X \prec_p U$ if there exists Y such that $U = XY$. We say that Y is a *suffix* of U, and we denote by $Y \prec_s U$ if there exists X such that $U = XY$.

Also, we say that a finite word V is a subword (or factor) of an infinite word $\mathbf{a} = (a_n)_{n\geq 0}$ if there exists a nonnegative integer j such that $V = a_j a_{j+1} \cdots a_{j+m-1}$. The integer j is called an occurrence of V.

Let U, V, W be three finite words over $\mathcal{A}$, V possibly empty. We denote

$$i(U,V,W) := \{AVB,\ A \prec_s U,\ B \prec_p W,\ A, B \text{ possibly empty}\},$$

and

$$i(U,V,W)^+ := \{AVB,\ A \prec_s U,\ B \prec_p W,\ A, B \text{ nonempty}\}.$$

7.2.1 Subword complexity and topological entropy

Let $\mathbb{F}_q$ be the finite field with q elements, where q is a power of a prime number p. In this chapter, we are interested in Laurent series with coefficients in $\mathbb{F}_q$. Let $n_0 \in \mathbb{N}$ and consider

$$f(T) = \sum_{n=-n_0}^{+\infty} a_n T^{-n} \in \mathbb{F}_q((T^{-1})).$$

Let m be a nonnegative integer. We define *the complexity* of f, denoted by $p(f,m)$, as being equal to the complexity of the infinite word $\mathbf{a} = (a_n)_{n\geq 0}$.

We also define *the entropy* of f, denoted by $h(f)$, as being equal to the entropy of the infinite word $\mathbf{a} = (a_n)_{n\geq 0}$.

We now give an useful tool in order to obtain bounds on the subword complexity function (for a proof see, for example, [20], p. 300–302).

Lemma 7.2.1. *Let $\boldsymbol{a}$ be an infinite word over an alphabet $\mathcal{A}$. We have the following properties:*

(i) $p(\boldsymbol{a}, m) \leq p(\boldsymbol{a}, m+1) \leq \operatorname{card} \mathcal{A} \cdot p(\boldsymbol{a}, m)$*, for every integer* $m \geq 0$*;*

(ii) $p(\boldsymbol{a}, m+n) \leq p(\boldsymbol{a}, m)p(\boldsymbol{a}, n)$*, for all integers* $m, n \geq 0$*.*

7.2.2 Morphisms

The main definitions about morphisms are recalled in the first part of this thesis.

For example, the Fibonacci word $\mathbf{f} = 0100101001001\cdots$ is an infinite word generated by iterating the morphism: $\sigma(0) = 01$ and $\sigma(1) = 0$. More precisely, $\mathbf{f} = \sigma^\infty(0)$ is the unique fixed point of σ.

We recall that a morphism σ is said to be uniform of length $m \geq 2$ if $|g(x)| = m$. Notice that a word generated by an uniform morphism of length m is m-automatic (see, for example, [50]). In particular, its complexity is $O(1)$ if the word is ultimately periodic; otherwise, it is $\Theta(m)$.

The *order of growth* of a letter x is the function $|\sigma^n(x)|$, for $n \geq 0$. In general, this function grows asymptotically like $n^{a_x} b_x^n$. A morphism is said

to be *polynomially diverging* if there exists $b > 1$ such that, for any letter x, the order of growth of x is $n^{a_x}b^n$ and $a_x \geq 1$ for some x. A morphism is *exponentially diverging* if every letter x has the order of growth $n^{a_x}b_x^n$ with $b_x > 1$ and not all b_x are equal. For more details the reader may refer to [101].

7.3 The analog of Π

In 1935, Carlitz [38] introduced for function fields in positive characteristic an analog of the exponential function defined over $\mathcal{C}_\infty$, which is the completion of the algebraic closure of $\mathbb{F}_q((T^{-1}))$ (this is the natural analog of the field of complex numbers). In order to get good properties analogous to the complex exponential, the resulting analog, $z \to e_C(z)$, satisfies

$$e_C(0) = 0,\ d/dz(e_C(z)) = 1 \text{ and } e_C(Tz) = Te_C(z) + e_C(z)^q.$$

This is what we call the Carlitz exponential and the action $u \to Tu + u^q$ leads to the definition of the Carlitz $\mathbb{F}_q[T]$-module, which is in fact a particular case of a Drinfeld module. The Carlitz exponential, $e_C(z)$, may be defined by the following infinite product

$$e_C(z) = z \prod_{a \in \mathbb{F}_q[T],\, a \neq 0} (1 - \frac{z}{a\widetilde{\Pi}_q})$$

where

$$\widetilde{\Pi}_q = (-T)^{\frac{q}{q-1}} \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j-1}}\right)^{-1}.$$

Since $e^z = 1$ if, and only if, $z \in 2\pi i\mathbb{Z}$ and since $e_C(z)$ was constructed by analogy such that $e_C(z) = 0$ if, and only if, $z \in \widetilde{\Pi}_q\mathbb{F}_q[T]$ (in other words the kernel of $e_C(z)$ is $\widetilde{\Pi}_q\mathbb{F}_q[T]$), we get a good analog $\widetilde{\Pi}_q$ of $2\pi i$. In order to obtain a good analog of the real number π, we take its one unit part and hence we obtain

$$\Pi_q = \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j-1}}\right)^{-1}.$$

We mention that we use the notation that appears in [125] (page 32 and 47) and we consider the analog of π defined by the formula above (as explained in [125], page 48 and page 365). For more details about analogs given by the theory of Carlitz modules, and in particular about the exponential function or its fundamental period $\widetilde{\Pi}_q$, we refer the reader to the monographs [69] (pages 51, 362) and [125] (pages 32, 47, 365).

If we look for the Laurent series expansion of Π_q, we obtain that

$$\Pi_q = \prod_{j=1}^{\infty} \left(1 - \frac{1}{T^{q^j-1}}\right)^{-1} = \sum_{n \geq 0} a_n T^{-n},$$

where a_n is defined as the number of partitions of n whose parts take values in $I = \{q^j - 1, j \geq 1\}$, taken modulo p. To compute the complexity of Π_q,

we would like to find a closed formula or some recurrence relations for the sequence of partitions $(a_n)_{n\geq 0}$. This question seems quite difficult and we are not able to solve it at this moment.

However, it was shown in [12] that the inverse of Π_q has the following simple Laurent series expansion

$$\frac{1}{\Pi_q} = \prod_{j=1}^{\infty}\left(1-\frac{1}{T^{q^j-1}}\right) = \sum_{n=0}^{\infty} p_n T^{-n}$$

where the sequence $\mathfrak{p}_q = (p(n))_{n\geq 0}$ is defined as follows

$$p_n = \begin{cases} 1 \text{ if } n=0; \\ (-1)^{\operatorname{card} J} \text{ if there exists a nonempty set } J\subset\mathbb{N}^* \text{ such that } n=\sum_{j\in J}(q^j-1); \\ 0 \text{ if there is no set } J\subset\mathbb{N}^* \text{ such that } n=\sum_{j\in J}(q^j-1). \end{cases} \tag{7.3}$$

***Remark* 7.3.1.** In [37], the authors show, using a well-known result of Fraenkel [65], that a nonnegative integer n can be written of the form $n=\sum_{i\geq 0} a_i(q^i-1)$, where $a_i\in\{0,1,\ldots,k-1\}$, if, and only if, n is a multiple of $q-1$. If it is the case, then n has a unique representation of this form.

7.3.1 Proof of Part (a) of Theorem 7.1.2

In this subsection we study the sequence $\mathfrak{p}_2 = (p_n^{(2)})_{n\geq 0}$, defined by the formula (7.3) in the case where $q=2$. More precisely:

$$p_n^{(2)} = \begin{cases} 1 \text{ if } n=0 \text{ or if there exists } J\subset\mathbb{N}^* \text{ such that } n=\sum_{j\in J}(2^j-1); \\ 0 \text{ otherwise.} \end{cases} \tag{7.4}$$

In order to simplify the notation, in the rest of this subsection we set $p_n := p_n^{(2)}$ so that $\mathfrak{p}_2 = p_0p_1p_2\cdots$. For every $n\geq 1$, we let W_n denote the factor of $\mathfrak{p}_2$ that occurs between positions 2^n-1 and $2^{n+1}-2$, that is:

$$W_n := p_{2^n-1}\cdots p_{2^{n+1}-2}.$$

We also set $W_0 := 1$. Observe that $|W_n| = 2^n$. With this notation the infinite word $\mathfrak{p}_2$ can be factorized as:

$$\mathfrak{p}_2 = \underbrace{1}_{W_0}\underbrace{10}_{W_1}\underbrace{1100}_{W_2}\underbrace{11011000}_{W_3}\cdots = W_0W_1W_2\cdots.$$

***Remark* 7.3.2.** The sequence $\mathfrak{p}_2$ is related to von Neumann's sequence (see [16], Theorem 2). In [16], the authors proved that $\mathfrak{p}_2$ is the fixed point of the morphism σ defined by $\sigma(1)=110$ and $\sigma(0)=0$. Note that, using Pansiot's theorem it is possible to prove that the complexity of $\mathfrak{p}_2$ is quadratic. In

an unpublished note [15], Allouche showed in an elementary way that the complexity of the sequence $\mathfrak{p}_2$ satisfies

$$p(\mathfrak{p}_2, m) \geq Cm \log m,$$

for some strictly positive constant C and every positive integer m.

In [16], they also showed that, for every $n \geq 2$, we have

$$W_n = 1W_1W_2 \cdots W_{n-1}0. \tag{7.5}$$

Since the subword W_n ends with 0, we can define U_n by $W_n := U_n0$, for every $n \geq 1$. Thus, $U_1 = 1$, $U_2 = 110$. For every $n \geq 1$, we have

$$U_{n+1} = U_nU_n0. \tag{7.6}$$

Lemma 7.3.1. *For every $n \geq 2$, there exists a word Z_n such that $W_n = 1Z_n10^n$ and $0^{n-1} \not\triangleleft Z_n$ (in other words W_n ends with exactly n zeros and Z_n does not contain blocks of 0 of length larger than $n-2$). This is equivalent to saying that $U_n = 1Z_n10^{n-1}$ and $0^{n-1} \not\triangleleft Z_n$.*

Proof. We argue by induction on n. For $n = 2$, $W_2 = 1100$ ends with two zeros and obviously there are no other zeros. We assume that W_n ends with n zeros and does not contain any other block of zeros of length greater than $n-2$. We show that this statement holds for $n+1$. By (7.6),

$$W_{n+1} = U_{n+1}0 = U_nU_n00.$$

As U_n ends exactly with $n-1$ zeros (by the induction hypothesis), then W_{n+1} also ends with $n+1$ zeros. Now, $U_n = 1Z_n10^{n-1}$ and $0^{n-1} \not\triangleleft Z_n$; so we have $W_{n+1} = 1Z_n10^{n-1}1Z_n10^{n-1}00 = 1Z_{n+1}10^{n+1}$, where $Z_{n+1} := Z_n10^{n-1}1Z_n$. Since $0^{n-1} \not\triangleleft Z_n$, then $0^n \not\triangleleft Z_{n+1}$. This completes the proof. □

Lemma 7.3.2. *For every $n \geq 1$, let $A_n := \{U_n^20^k,\ k \geq 1\}$. Then $\mathfrak{p}_2 \in A_n^{\mathbb{N}}$.*

Proof. Let $n \geq 1$. By definition of W_n and U_n and using the relation (7.5), the infinite word $\mathfrak{p}_2$ can be factorized as:

$$\mathfrak{p}_2 = \underbrace{1W_1W_2 \cdots W_{n-1}}_{U_n}\underbrace{U_n0}_{W_n}\underbrace{U_{n+1}0}_{W_{n+1}}\underbrace{U_{n+2}0}_{W_{n+2}} \cdots . \tag{7.7}$$

We prove that for every positive integer k, there exist a positive integer r and $k_1, k_2, \ldots, k_r \in \mathbb{N}^*$ such that:

$$U_{n+k} = U_n^20^{k_1}U_n^20^{k_2} \cdots U_n^20^{k_r}. \tag{7.8}$$

We argue by induction on k. For $k = 1$, we have $U_{n+1} = U_nU_n0 = U_n^20$. We suppose that the relation (7.8) is true for k and we show it for $k+1$. By (7.6)

$$U_{n+k+1} = U_{n+k}U_{n+k}0 = U_n^20^{k_1}U_n^20^{k_2} \cdots U_n^20^{k_r}U_n^20^{k_1}U_n^20^{k_2} \cdots U_n^20^{k_r+1}.$$

By Eq. (7.7), this ends the proof. □

***Remark* 7.3.3.** In fact, in Eq. (7.8), one can easily see that $k_r = k$ and $k_i < k$, for $i < r$. We also have the following expression for W_{n+k}:

$$W_{n+k} = U_n^2 0^{k_1} U_n^2 0^{k_2} \cdots U_n^2 0^{k+1}.$$

Fix $m \in \mathbb{N}, m \geq 2$. Then, there is a unique integer n such that:

$$2^{n-1} < m \leq 2^n. \tag{7.9}$$

Lemma 7.3.3. *Let $m \in \mathbb{N}$. All distinct words of length m of $\mathfrak{p}_2$ occur in the prefix:*

$$P_m = W_0 W_1 \cdots W_m.$$

Proof. Let m, n be some positive integers satisfying (7.9).

W_m because

We show that all distinct words of length m occur in the prefix

$$P_m = W_0 W_1 W_2 \cdots W_{n-1} W_n \cdots W_m = U_n \underbrace{U_n 0}_{W_n} \underbrace{U_n U_n 00}_{W_{n+1}} \underbrace{U_n U_n 0 U_n U_n 000}_{W_{n+2}} \cdots W_m;$$

the second identity following from (7.5) and (7.6).

Notice that we cannot choose a shorter prefix because the word 0^m first occurs in W_m.

By Remark 7.3.3, W_i ends with $U_n U_n 0^{i-n+1}$, for every $i \geq n+1$ and there are no other block of zeros of length greater than $i - n + 1$. Hence, all the words $U_n U_n 0^k$, with $0 \leq k \leq m - n + 1$, are factors of P_m. More precisely, if

$$B_n := \{U_n U_n 0^k, 0 \leq k \leq m - n + 1\} \text{ then } P_m \in B_n^*.$$

After the prefix P_m, it is not possible to see new different subwords of length m. Indeed, suppose that there exists a word F of length m such that $F \triangleleft W_{m+1} W_{m+2} W_{m+3} \cdots$ and $F \not\triangleleft P_m$. Then, by Lemma 7.3.2 and by the remark above, Since $1 + |U_n| = 2^n \geq m$ and using Lemma 7.3.2, F must occur in the words $U_n 0^k U_n$, with $k \geq m - n + 1$. But since U_n ends with $n - 1$ zeros (by Lemma 7.3.1), F must be equal to 0^m or $0^i R_i$, where $i \geq m - n + 2$ and $R_i \prec_p U_n$, or $F \triangleleft U_n$. But all these words already appear in P_m. This contradicts our assumption. □

An upper bound for $p(\frac{1}{\Pi_2}, m)$

In this part we prove the following result:

$$p(\mathfrak{p}_2, m) \leq \frac{(m - \log_2 m)(m + \log_2 m + 2)}{2} + 2m. \tag{7.10}$$

In order to find all different factors of length m that occur in $\mathfrak{p}_2$, it suffices, by Lemmas 7.3.2 and 7.3.3, to consider factors appearing in the word $U_n U_n$ and in the sets $i(U_n, 0^k, U_n)$, where $1 \leq k \leq m - n$.

In the word U_nU_n we can find at most $|U_n|$ distinct words of length m. Since $|U_n| = 2^n - 1$ and $2^{n-1} < m \leq 2^n$, the number of factors of length m that occur in U_nU_n is at most 2^n, so at most $2m$.

Also, it is not difficult to see that $\left|i(U_n, 0^k, U_n)\right| \cap \mathcal{A}^m \leq m - k + 1$. The total number of subwords occurring in all these sets, for $1 \leq k \leq m-n$, is less than or equal to:

$$\sum_{k=1}^{m-n} (m-k+1) = (m+1)(m-n) - \frac{(m-n)(m-n+1)}{2}.$$

Counting all these words and using the fact that $2^{n-1} < m \leq 2^n$, we obtain that:

$$p(\mathfrak{p}_2, m) \leq 2m + \frac{(m-n)(m+n+1)}{2} < \frac{(m-\log_2 m)(m+\log_2 m+2)}{2} + 2m$$

as claimed.

A lower bound for $p(\frac{1}{\Pi_2}, m)$

In this part we prove the following result:

$$p(\mathfrak{p}_2, m) \geq \frac{(m-\log_2 m)(m-\log_2 m+1)}{2}. \tag{7.11}$$

By Lemma 7.3.3, we have to look for distinct words of length m occuring in $W_nW_{n+1}\cdots W_m$.

In order to prove this proposition, we use the final blocks of 0 from each W_i. These blocks are increasing (as we have shown in Lemma 7.3.1). First, in the word W_m we find for the first time the word of length m: 0^m.

In the set $i(W_{m-1}, \varepsilon, W_m)$, we find two distinct words of length m that do not occur previously (10^{m-1} and $0^{m-1}1$) since there are no other words containing blocks of zeros of length $m-1$ in $i(W_k, \varepsilon, W_{k+1})$, for $k < m-1$.

More generally, fix k such that $n \leq k \leq m-2$. Since

$$W_kW_{k+1} = \underbrace{1Z_k10^k}_{W_k}\underbrace{1Z_{k+1}10^{k+1}}_{W_{k+1}},$$

in $i(W_k, \varepsilon, W_{k+1})$ we find $m-k+1$ words of length m of the form $\alpha_k0^k\beta_k$. More precisely, the words we count here are the following $S_{m-k-1}10^k$, $S_{m-k-2}10^k1$, $S_{m-k-3}10^k1T_1$, ..., $S_10^k1T_{m-k-2}$, 0^k1T_{m-k-1}, where $S_i \prec_s Z_k$ and $T_i \prec_p Z_{k+1}$, $|S_i| = |T_i| = i$, for every integer i, $1 \leq i \leq m-k-1$.

All these words do not occur previously, that is in $i(W_s, \varepsilon, W_{s+1})$, for $s < k$, since there are no blocks of zeros of length k before the word W_k (according to Lemma 7.3.1). Also, in $i(W_s, \varepsilon, W_{s+1})$, for $s > k$, we focus on the words $\alpha_s0^s\beta_s$ and hence they are different from all the words seen before (because $k < s$).

Consequently, the total number of subwords of length m of the form $\alpha_k0^k\beta_k$ considered before, is equal to

$$1 + 2 + \ldots + (m-n+1) = \frac{(m-n+1)(m-n+2)}{2}.$$

Since $2^{n-1} < m \leq 2^n$ we obtain the desired lower bound.

Proof of Part (a) of Theorem 7.1.2. Follows from inequalities (7.10) and (7.11). □

A consequence of Theorem 7.1.2 and Theorem 7.1.1 is the following result on transcendence.

Corollary 7.3.1. *Let $\mathbb{K}$ be a finite field and let $(p_n^{(2)})_{n\geq 0}$ be the sequence defined in (7.4). Let us consider the associated formal power series over $\mathbb{K}$:*

$$f(T) := \sum_{n\geq 0} p_n^{(2)} T^{-n} \in \mathbb{K}[[T^{-1}]].$$

Then f is transcendental over $\mathbb{K}(T)$.

Notice that, if $\mathbb{K} = \mathbb{F}_2$ then the formal power series f coincide with $1/\Pi_2$ and hence Corollary 7.3.1 implies Corollary 7.1.1.

7.3.2 Proof of Part (b) of Theorem 7.1.2

In this part we study the sequence $\mathfrak{p}_q = (p_n^{(q)})_{n\geq 0}$ defined by the formula (7.3) in the case where $q \geq 3$. In the following, we will consider the case $q = p^n$, where $p \geq 3$.

Proposition 7.3.1. *Let $q \geq 3$. For every positive integer m, we have*

$$p(\mathfrak{p}_q, m) \leq (2q+4)m + 2q - 3.$$

In particular, this proves Part (b) of Theorem 7.1.2. Indeed, we do not have to find a lower bound for the complexity function, as the sequence $\mathfrak{p}_q$ is not ultimately periodic (see Remark 7.3.5) and thus, by Morse and Hedlund's theorem we have

$$p(\mathfrak{p}_q, m) \geq m + 1,$$

for any $m \geq 0$.

In order to simplify the notation, we set in the sequel $p_n := p_n^{(q)}$ so that $\mathfrak{p}_q = p_0 p_1 p_2 \cdots$.

For every $n \geq 1$, we let W_n denote the factor of $\mathfrak{p}_q$ defined in the following manner

$$W_n := p_{q^n-1} \cdots p_{q^{n+1}-2}.$$

Let us fix $W_0 := 0^{q-2} = \underbrace{00\cdots 0}_{q-2}$ and $\alpha_0 := q-2$. Thus $W_0 = 0^{\alpha_0}$.

In other words, W_n is the factor of $\mathfrak{p}_q$ occurring between positions $q^n - 1$ and $q^{n+1} - 2$. Notice that $|W_n| = q^n(q-1)$.

With this notation the infinite word $\mathfrak{p}_q$ may be factorized as follows

$$\mathfrak{p}_q = 1\underbrace{00\cdots 0}_{W_0}\underbrace{(-1)00\cdots 0}_{W_1}\underbrace{(-1)\cdots 00100\cdots 0}_{W_2}(-1)00\cdots.$$

Now, we prove some lemmas that we use in order to bound from above the complexity function of $\mathfrak{p}_q$.

First, we can deduce from Remark 7.3.1 the following properties of $\mathfrak{p}_q$:

$$\text{for any } k,n\in\mathbb{N}^* \text{ such that } k\in\Big[2(q^n-1),q^{n+1}-2\Big], \text{ we have } p_k=0; \quad (7.12)$$

$$\text{for any } k,n\in\mathbb{N}^* \text{ such that } k<q^n-1, \text{ we have } p_k=-p_{k+(q^n-1)}. \quad (7.13)$$

If $W=a_1a_2\cdots a_l\in\{0,1,-1\}^l$ then set $\widehat{W}:=(-a_1)(-a_2)\cdots(-a_l)$.

Lemma 7.3.4. *For every $n\geq 1$, we have the following*

$$W_n=(-1)\widehat{W_0}\widehat{W_1}\cdots\widehat{W_{n-1}}0^{\alpha_n}$$

with $\alpha_n=(q^{n+1}-1)-2(q^n-1)$.

Proof. Obviously, the word W_n begins with -1 since $p_{q^n-1}=-1$. In order to prove the relation above it suffices to split W_n into subwords as follows

$$W_n=\underbrace{p_{q^n-1}}_{-1}\underbrace{0\cdots 0}_{W_0'}\underbrace{p_{(q^n-1)+(q-1)}p_{(q^n-1)+(q^2-2)}}_{W_1'}\underbrace{p_{(q^n-1)+(q^2-1)}\cdots p_{(q^n-1)+(q^3-2)}}_{W_2'}$$

$$\cdots\underbrace{p_{(q^n-1)+(q^{n-1}-1)}\cdots p_{(q^n-1)+(q^n-2)}}_{W_{n-1}'}\underbrace{p_{2(q^n-1)}\cdots p_{q^{n+1}-2}}_{0^{\alpha_n}}.$$

Since $p_{(q^n-1)+k}=-p_k$, for every $k<q^n-1$ (by (7.13)), we obtain that $W_i'=\widehat{W_i}$, for $0\leq i\leq n-1$. The relation (7.12) ends the proof. □

Since the subword W_n ends with 0^{α_n}, we can define U_n as prefix of W_n such that $W_n:=U_n0^{\alpha_n}$, for every $n\geq 1$. Notice that $|U_n|=q^n-1$.

Lemma 7.3.5. *For every $n\geq 1$, we have $U_{n+1}=U_n\widehat{U_n}0^{\alpha_n}$.*

Proof. By Lemma 7.3.4, $U_n=(-1)\widehat{W_0}\widehat{W_1}\cdots\widehat{W_{n-1}}$. Consequently:

$$U_{n+1}=\underbrace{(-1)\widehat{W_0}\widehat{W_1}\cdots\widehat{W_{n-1}}}_{U_n}\widehat{W_n}=U_n\widehat{U_n0^{\alpha_n}}=U_n\widehat{U_n}0^{\alpha_n}.$$

□

***Remark* 7.3.4.** Since $q\geq 3$ we have $\alpha_n\geq|U_n|$ for every $n\geq 1$. Moreover $(\alpha_n)_{n\geq 1}$ is a positive and increasing sequence.

Lemma 7.3.6. *For every $n\geq 1$, let $A_n:=\{U_n,\widehat{U_n},0^{\alpha_i},\ i\geq n\}$. Then $\mathfrak{p}_q\in A_n^{\mathbb{N}}$.*

Proof. Let $n \geq 1$. By definition of W_n and U_n, the infinite word $\mathfrak{p}_q$ can be factorized as:

$$\mathfrak{p}_q = \underbrace{1W_0W_1\cdots W_{n-1}}_{V_n} W_nW_{n+1}\cdots .$$

By Lemma 7.3.4, since $U_n = (-1)\widehat{W_0}\widehat{W_1}\cdots\widehat{W_{n-1}}$ then the prefix $V_n = \widehat{U_n}$.

Also, by Lemma 7.3.5 $W_{n+1} = U_n\widehat{U_n}0^{\alpha_n}0^{\alpha_{n+1}}$, $W_{n+2} = U_n\widehat{U_n}0^{\alpha_n}\widehat{U_n}U_n0^{\alpha_n+\alpha_{n+1}+\alpha_{n+2}}$. Keeping on this procedure, W_n can be written as a concatenation of U_n, $\widehat{U_n}$ and 0^{α_i}, $i \geq n$. More precisely, $\mathfrak{p}_q$ can be written in the following manner

$$\mathfrak{p}_q = \widehat{U_n}\underbrace{U_n0^{\alpha_n}}_{W_n}\underbrace{U_n\widehat{U_n}0^{\alpha_n+\alpha_{n+1}}}_{W_{n+1}}\underbrace{U_n\widehat{U_n}0^{\alpha_n}\widehat{U_n}U_n0^{\alpha_n+\alpha_{n+1}+\alpha_{n+2}}}_{W_{n+2}}\cdots .$$

□

Proof of Proposition 7.3.1. Let $m \in \mathbb{N}$. Then there exists a unique positive integer n, such that:

$$q^{n-1} - 1 \leq m < q^n - 1.$$

By Lemma 7.3.6 and the Remark 7.3.4, between the words U_n and $\widehat{U_n}$ (when they do not occur consecutively), there are only blocks of zeros of length greater than $\alpha_n \geq |U_n| = q^n - 1$ and thus greater than m. Hence, all distinct factors of length m appear in the following words: $U_n\widehat{U_n}$, $\widehat{U_n}U_n$, $0^{\alpha_n}U_n$, $0^{\alpha_n}\widehat{U_n}$, $U_n0^{\alpha_n}$ and $\widehat{U_n}0^{\alpha_n}$.

In $U_n\widehat{U_n}$ we may find at most $|U_n\widehat{U_n}| - m + 1 = 2|U_n| - m + 1$ factors at length m. In $\widehat{U_n}U_n$ we may find at most $m - 1$ new different factors of length m. More precisely, they form the set $i(\widehat{U_n}, \varepsilon, U_n)^+$.

In $0^{\alpha_n}U_n$ (respectively $0^{\alpha_n}\widehat{U_n}$, $U_n0^{\alpha_n}$, $\widehat{U_n}0^{\alpha_n}$) we may find at most m (respectively $m - 1$) new different factors (they belong to $i(0^{\alpha_n}, \varepsilon, U_n)^+ \cup \{0^m\}$, respectively $i(0^{\alpha_n}, \varepsilon, \widehat{U_n})^+$, $i(U_n, \varepsilon, 0^{\alpha_n})^+$ and $i(\widehat{U_n}, \varepsilon, 0^{\alpha_n})^+$).

Consequently, the number of such subwords is at most $2|U_n| + 4m - 3$. Since $U_n = q^n - 1 = q(q^{n-1} - 1) + q - 1 \leq qm + q - 1$ we obtain that:

$$p(\mathfrak{p}_q, m) \leq 2(qm + q - 1) + 4m - 3 \leq (2q + 4)m + 2q - 3.$$

□

***Remark* 7.3.5.** It is not difficult to prove that $\mathfrak{p}_q$ is not ultimately periodic. Indeed, recall that $\mathfrak{p}_q = W_0W_1W_2\cdots$. Using Theorem 7.3.4 and the Remark 7.3.4,

$$\mathfrak{p}_q = A_10^{l_1}A_20^{l_2}\cdots A_i0^{l_i}\cdots , \tag{7.14}$$

where A_i, $i \geq 1$, are finite words such that $A_i \neq 0^{|A_i|}$ and $(l_i)_{i\geq 1}$ is a strictly increasing sequence.

***Remark* 7.3.6.** This part concerns the case where $q \geq 3$. If the characteristic of the field is 2, that is, if $q = 2^n$, where $n \geq 2$, then, in the proof we have $-1 = 1$, but the structure of $\mathfrak{p}_q$ remains the same. We will certainly have a lower complexity, but $\mathfrak{p}_q$ is still of the form (7.14), and thus $p(\mathfrak{p}_q, m) \leq (2q + 4)m + 2q - 3$.

7.4 Closure properties of two classes of Laurent series

The subword complexity offers a natural way to classify Laurent series with coefficients in a finite field. In this section we study some closure properties of the following classes:

$$\mathcal{P} = \{f \in \mathbb{F}_q((T^{-1})), \text{there exists } K \text{ such that } p(f,m) = O(m^K)\},$$

and

$$\mathcal{Z} = \{f \in \mathbb{F}_q((T^{-1})), \text{ such that } \ h(f) = 0\}.$$

Clearly, $\mathcal{P} \subset \mathcal{Z}$. We recall that h denotes the topological entropy, as defined in Section 7.2.1. We have already seen, in Theorem 7.1.1, that all algebraic Laurent series belong to $\mathcal{P}$. Also, by Theorem 7.1.2, $1/\Pi_q$ belongs to $\mathcal{P}$. Hence, $\mathcal{P}$, and more generally $\mathcal{Z}$, seem to be two relevant sets for this classification.

The main result we will prove in this section is Theorem 7.1.3. We will also prove theat $\mathcal{P}$ and $\mathcal{Z}$ are closed under usual operations such as the Hadamard product, the formal derivative and the Cartier operators.

7.4.1 Proof of Theorem 7.1.3

The proof of Theorem 7.1.3 is a straightforward consequence of Propositions 7.4.1 and 7.4.3 below.

Proposition 7.4.1. *Let f and g be two Laurent series belonging to $\mathbb{F}_q((T^{-1}))$. Then, for every integer $m \geq 1$, we have:*

$$\frac{p(f,m)}{p(g,m)} \leq p(f+g,m) \leq p(f,m)p(g,m).$$

Proof. Let $f(T) := \sum_{i \geq -i_1} a_i T^{-i}$ and $g(T) := \sum_{i \geq -i_2} b_i T^{-i}$, $i_1, i_2 \in \mathbb{N}$.

By definition of the complexity of Laurent series (see Section (7.2.1)), for every $m \in \mathbb{N}$:

$$p(f(T) + g(T), m) = p(\sum_{i \geq 0} c_i T^{-i}, m),$$

where $c_i := (a_i + b_i) \in \mathbb{F}_q$. Thus we may suppose that

$$f(T) := \sum_{i \geq 0} a_i T^{-i} \text{ and } g(T) := \sum_{i \geq 0} b_i T^{-i}.$$

We let $\mathbf{a} := (a_i)_{i \geq 0}$, $\mathbf{b} := (b_i)_{i \geq 0}$ and $\mathbf{c} := (c_i)_{i \geq 0}$.

For the sake of simplicity, throughout this part, we set $x(m) := p(f,m)$ and $y(m) := p(g,m)$. Let $\mathcal{L}_{f,m} := \{U_1, U_2, \ldots, U_{x(m)}\}$ (resp. $\mathcal{L}_{g,m} := \{V_1, V_2, \ldots, V_{y(m)}\}$) be the set of different factors of length m of the sequence of coefficients of f (resp. of g). As the sequence of coefficients of

the Laurent series $f+g$ is obtained by the termwise addition of the sequence of coefficients of f and the sequence of coefficients of g, we deduce that:

$$\mathcal{L}_{f+g,m} \subseteq \{U_i + V_j,\ 1 \leq i \leq x(m),\ 1 \leq j \leq y(m)\}$$

where $\mathcal{L}_{f+g,m}$ is the set of all distinct factors of length m occurring in $\mathbf{c}$, and where the sum of two words with the same length $A = a_1 \cdots a_m$ and $B = b_1 \cdots b_m$ is defined as

$$A + B = (a_1 + b_1) \cdots (a_m + b_m)$$

(each sum being considered over $\mathbb{F}_q$). Consequently, $p(f+g,m) \leq p(f,m)p(g,m)$.

We shall now prove the first inequality using Dirichlet's principle.

Notice that if $x(m) < y(m)$ the inequality is obvious.

Now assume that $x(m) \geq y(m)$. Notice that if we extract $x(m)$ subwords of length m from $\mathbf{b}$, there is at least one word which appears at least $\left\lceil \frac{x(m)}{y(m)} \right\rceil$ times.

For every fixed m, there exist exactly $x(m)$ different factors of $\mathbf{a}$. The subwords of $\mathbf{c}$ will be obtained adding factors of length m of $\mathbf{a}$ with factors of length m of $\mathbf{b}$.

Consider all distinct factors of length m of $\mathbf{a}$: $U_1, U_2, \ldots, U_{x(m)}$, that occur in positions $i_1, i_2, \ldots, i_{x_m}$. Looking at the same positions in $\mathbf{b}$, we have $x(m)$ factors of length m belonging to $\mathcal{L}_{g,m}$. Since $x(m) \geq y(m)$, by the previous remark, there is one word W which occur at least $\left\lceil \frac{x(m)}{y(m)} \right\rceil$ times in $\mathbf{b}$.

Since we have $U_i + W \neq U_j + W$ if $U_i \neq U_j$, the conclusion follows immediately. □

***Remark* 7.4.1.** In fact, the first inequality may also be easily obtained from the second one, but we chose here to give a more intuitive proof. Indeed, if we denote $f := h_1 + h_2$, $g := -h_2$, where $h_1, h_2 \in \mathbb{F}_q((T^{-1}))$, the first relation follows immediately, since $p(h_2, m) = p(-h_2, m)$, for any $m \in \mathbb{N}$.

***Remark* 7.4.2.** If $f \in \mathbb{F}_q((T^{-1}))$ and $a \in \mathbb{F}_q[T]$ then, obviously, there exists a constant C (depending on the degree of the polynomial a) such that, for any $m \in \mathbb{N}$,

$$p(f + a, m) \leq p(f, m) + C.$$

***Remark* 7.4.3.** Related to Proposition 7.4.1, one can naturally ask if it is possible to attain the bounds in Proposition 7.4.1. By Remark 7.4.1, it suffices to show that this is possible for one inequality. In the sequel, we give an example of two Laurent series of linear complexity whose sum has quadratic complexity.

Let α and β be two irrational numbers such that 1, α and β are linearly independent over $\mathbb{Q}$. For any $i \in \{\alpha, \beta\}$ we consider the following rotations:

$$R_i : \mathbb{T}^1 \to \mathbb{T}^1 \qquad x \to \{x + i\},$$

where $\mathbb{T}^1$ is the circle $\mathbb{R}/\mathbb{Z}$, identified to the interval $[0, 1)$.

We may partition $\mathbb{T}^1$ in two intervals I_i^0 and I_i^1, delimited by 0 and $1-i$. We let ν_i denote the coding function:

$$\nu_i(x) = \begin{cases} 0 & \text{if } x \in I_i^0; \\ 1 & \text{if } x \in I_i^1. \end{cases}$$

We define $\mathbf{a} := (a_n)_{n\geq 0}$ such that, for any $n \geq 0$,

$$a_n = \nu_\alpha(R_\alpha^n(0)) = \nu_\alpha(\{n\alpha\})$$

and $\mathbf{b} := (a_n)_{n\geq 0}$ such that, for any $n \geq 0$,

$$b_n = \nu_\beta(R_\beta^n(0)) = \nu_\beta(\{n\beta\}).$$

Let us consider $f(T) = \sum_{n\geq 0} a_n T^{-n}$ and $g(T) = \sum_{n\geq 0} b_n T^{-n}$ be two elements of $\mathbb{F}_3((T^{-1}))$. We will prove that, for any $m \in \mathbb{N}$, we have

$$p(f+g, m) = p(f, m)p(g, m). \tag{7.15}$$

We thus provide an example of two infinite words whose sum has a maximal complexity, in view of Proposition 7.4.1.

A sequence of the form $(\nu(R_\alpha^n(x)))_{n\geq 0}$ is a particular case of a rotation sequence. It is not difficult to see that the complexity of the sequence $\mathbf{a}$ satisfies $p(\mathbf{a}, m) = m + 1$ for any $m \in \mathbb{N}$ and hence $\mathbf{a}$ is Sturmian. For a complete proof, the reader may consult the monograph [104], but also the original paper of Morse and Hedlund [96], where they prove that every Sturmian sequence is a rotation sequence.

Let $m \in \mathbb{N}$. Let

$$\mathcal{L}_{\mathbf{a},m} := \{U_1, U_2, \ldots, U_{m+1}\}$$

and resp.

$$\mathcal{L}_{\mathbf{b},m} := \{V_1, V_2, \ldots, V_{m+1}\}$$

be the set of distinct factors of length m that occur in $\mathbf{a}$, resp. in $\mathbf{b}$.

In order to prove the relation (7.15), we show that

$$\mathcal{L}_{\mathbf{a}+\mathbf{b},m} = \{U_i + V_j, 1 \leq i, j \leq m+1\}. \tag{7.16}$$

Let $I := [0, 1)$. It is well-known (see, for example, Proposition 6.1.7 in [104]) that, using the definition of the sequence $\mathbf{a}$ (resp. of $\mathbf{b}$), we can split I in $m+1$ intervals of positive length $J_1, J_2, \ldots, J_{m+1}$ (resp. $L_1, L_2, \ldots, L_{m+1}$) corresponding to $U_1, U_2, \ldots, U_{m+1}$ (resp. $V_1, V_2, \ldots, V_{m+1}$) such that:

$$\{n\alpha\} \in J_k \text{ if, and only if, } a_n a_{n+1} \cdots a_{n+m-1} = U_k$$

$$(\text{resp. } \{n\beta\} \in L_k \text{ if, and only if, } b_n b_{n+1} \cdots b_{n+m-1} = V_k).$$

In other words, $\{n\alpha\} \in J_k$ (resp. $\{n\beta\} \in L_k$) if, and only if, the factor U_k (resp. V_k) occurs in $\mathbf{a}$ (resp. $\mathbf{b}$) at the position n.

Now we use the well-known Kronecker's theorem which asserts that the sequence of fractional parts $(\{n\alpha\},\{n\beta\})_{n\geq 0}$ is dense in the square $[0,1)^2$ since by assumption 1, α and β are linearly independent over $\mathbb{Q}$.

In particular, this implies that, for any pair $(i,j)\in\{0,1,\ldots,m+1\}^2$, there exists a positive integer n such that $(\{n\alpha\},\{n\beta\})\in J_i\times L_k$. This is equivalently to saying that, for any pair of factors $(U_i,V_j)\in\mathcal{L}_{\mathbf{a},m}\times\mathcal{L}_{\mathbf{b},m}$, there exists n such that $U_i=a_na_{n+1}\cdots a_{n+m-1}$ and $V_k=b_nb_{n+1}\cdots b_{n+m-1}$. This proves Equality (7.16) and more precisely, since we are in characteristic 3, we have the following equality

$$\operatorname{Card}\mathcal{L}_{\mathbf{a}+\mathbf{b},m}=\operatorname{Card}\mathcal{L}_{\mathbf{a},m}\cdot\operatorname{Card}\mathcal{L}_{\mathbf{b},m}=(m+1)^2.$$

We mention that the idea of our construction here already appears in Theorem 7.6.6 in [20].

We point out the following consequence of Proposition 7.4.1.

Corollary 7.4.1. *Let $f_1,f_2,\ldots,f_l\in\mathbb{F}_q((T^{-1}))$. Then for every $m\in\mathbb{N}$ and for every integer $i\in[1;l]$ we have the following*

$$\frac{p(f_i,m)}{\prod_{j\neq i,1\leq j\leq l}p(f_j,m)}\leq p(f_1+f_2+\cdots+f_l,m)\leq\prod_{1\leq j\leq l}p(f_j,m).$$

Notice that the bounds in these inequalities can be attained, just generalizing the construction above (choose l Sturmian sequences of irrational slopes $\alpha_1,\alpha_2,\ldots,\alpha_l$, such that $1,\alpha_1,\alpha_2,\ldots,\alpha_l$ are linearly independent over $\mathbb{Q}$).

We shall next prove that the sets $\mathcal{P}$ and $\mathcal{Z}$ are closed under multiplication by rational functions. Let us begin with a particular case, that is the multiplication by a polynomial.

Proposition 7.4.2. *Let $b(T)\in\mathbb{F}_q[T]$ and $f(T)\in\mathbb{F}_q((T^{-1}))$. Then there is a positive constant M (depending only on $b(T)$), such that for all $m\in\mathbb{N}$:*

$$p(bf,m)\leq M\,p(f,m).$$

Proof. Let

$$b(T):=b_0T^r+b_1T^{r-1}+\cdots+b_r\in\mathbb{F}_q[T]$$

and

$$f(T):=\sum_{i\geq -i_0}a_iT^{-i}\in\mathbb{F}_q((T^{-1})),\ i_0\in\mathbb{N}.$$

Then

$$\begin{aligned}b(T)f(T)&=b(T)\left(\sum_{i=-i_0}^{-1}a_iT^{-i}+\sum_{i\geq 0}a_iT^{-i}\right)\\&=b(T)\left(\sum_{i=-i_0}^{-1}a_iT^{-i}\right)+b(T)\left(\sum_{i\geq 0}a_iT^{-i}\right).\end{aligned}\tag{7.17}$$

Now, the product

$$\begin{aligned} b(T)(\sum_{i\geq 0} a_i T^{-i}) &= T^r(b_0 + b_1 T^{-1} + \cdots + b_r T^{-r})(\sum_{i\geq 0} a_i T^{-i}) \\ &:= T^r(\sum_{j\geq 0} c_j T^{-j}) \end{aligned} \tag{7.18}$$

where the sequence $\mathbf{c} := (c_j)_{j\geq 0}$ is defined as follows

$$c_j = \begin{cases} b_0 a_j + b_1 a_{j-1} + \cdots + b_j a_0 \text{ if } j < r \\ b_0 a_j + b_1 a_{j-1} + \cdots + b_r a_{j-r} \text{ if } j \geq r. \end{cases}$$

According to the definition of complexity (see Section 7.2.1) and using (7.17) and (7.18), for every $m \in \mathbb{N}$, we have

$$p\left(b(t)f(T), m\right) = p\left(b(T)(\sum_{i\geq 0} a_i T^{-i}), m\right) = p\left((\sum_{j\geq r} c_j T^{-j}), m\right).$$

Our aim is to count the number of words of the form $c_j c_{j+1} \cdots c_{j+m-1}$, when $j \geq r$. By definition of $\mathbf{c}$, we notice that for $j \geq r$ these words depend only on $a_{j-r}a_{j-r+1}\cdots a_{j+m-1}$ and on $b_0, b_1, \cdots, b_r$, which are fixed. The number of words $a_{j-r}a_{j-r+1}\cdots a_{j+m-1}$ is exactly $p(f, m+r)$. By Lemma 7.2.1 we obtain

$$p(f, m+r) < p(f, r)p(f, m) = Mp(f, m),$$

where $M = p(f, r)$. More precisely, we may bound M from above by q^r, since this is the number of all possible words of length r over an alphabet of q letters. □

Proposition 7.4.3. *Let $r(T) \in \mathbb{F}_q(T)$ and $f(T) = \sum_{n\geq -n_0} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$. Then for every $m \in \mathbb{N}$, there is a positive constant M, depending only on $r(T)$ and n_0, such that:*

$$p(rf, m) \leq Mp(f, m).$$

Proof. Let $f(T) := \sum_{i\geq -i_0} a_i T^{-i} \in \mathbb{F}_q((T^{-1}))$, $i_0 \in \mathbb{N}$ and $m \in \mathbb{N}$. By Proposition 7.4.1, we have

$$p(r(T)f(T), m) \leq p\left(r(T)(\sum_{i=-i_0}^{-1} a_i T^{-i}), m\right) \cdot p\left(r(T)(\sum_{i\geq 0} a_i T^{-i}), m\right).$$

Proposition 7.4.2 implies that

$$p\left(r(T)(\sum_{i=-i_0}^{-1} a_i T^{-i}), m\right) \leq R$$

where R does not depend on m. Thus, we may assume that $f(T) = \sum_{i\geq 0} a_i T^{-i}$.

We divide the proof of Proposition 7.4.3 into five steps.

Step 1. Since $r(T) \in \mathbb{F}_q(T)$, the sequence of coefficients of r is ultimately periodic. Thus, there exist two positive integers $S, L \in \mathbb{N}^*$ and $p_1 \in \mathbb{F}_q[T]$ (with degree equal to $S-1$) et $p_2 \in \mathbb{F}_q[T]$ (with degree equal to $L-1$) such that

$$r(T) = \frac{P(T)}{Q(T)} = \frac{p_1(T)}{T^{S-1}} + \frac{p_2(T)}{T^{S+L-1}}(1 + T^{-L} + T^{-2L} + \cdots).$$

Hence

$$\begin{aligned} r(T)f(T) &= \underbrace{\frac{1}{T^{S-1}}p_1(T)f(T)}_{g(T)} + \underbrace{p_2(T)\frac{1}{T^{S+L-1}}f(T)(1 + T^{-L} + T^{-2L}\cdots)}_{h(T)} \\ &:= \sum_{n\geq 0} f_n T^{-n}. \end{aligned} \tag{7.19}$$

We let $\mathbf{d} = (d(n))_{n\geq 0}$ denote the sequence of coefficients of $g(T)$ and by $\mathbf{e} = (e_n)_{n\geq 0}$ the sequence of coefficients of $h(T)$. Clearly $\mathbf{f} := (f_n)_{n\geq 0}$ is such that $f_n = d_n + e_n$, for every $n \in \mathbb{N}$.

Fix $m \in \mathbb{N}$. Our aim is to bound from above $p(\mathbf{f}, m)$. First, assume that m is a multiple of L and set $m = kL$, where $k \in \mathbb{N}$.

In order to bound the complexity of $\mathbf{f}$, we will consider separately the sequences $\mathbf{e}$ and $\mathbf{d}$.

Step 2. We now study the sequence $\mathbf{e}$, defined in (7.19).

In order to describe the sequence $\mathbf{e}$, we shall first study the product

$$f(T)(1 + T^{-L} + T^{-2L} + \cdots) = (\sum_{i\geq 0} a_i T^{-i})(1 + T^{-L} + T^{-2L} + \cdots) := \sum_{j\geq 0} c_j T^{-j}.$$

Expanding this product, it is not difficult to see that $c_l = a_l$ if $l < L$ and $c_{kL+l} = a_l + a_{l+L} + \cdots + a_{kL+l}$, for $k \geq 1$ and $0 \leq l \leq L-1$.

By definition of c_n, $n \in \mathbb{N}$, we can easily obtain

$$c_{n+L} - c_n = a_{n+L}.$$

Consequently, for all $s \in \mathbb{N}$, we have

$$c_{n+sL} - c_n = a_{n+sL} + a_{n+(s-1)L} + \cdots + a_{n+L}. \tag{7.20}$$

Our goal is now to study the subwords of $\mathbf{c}$ with length $m = kL$.

Let $j \geq 0$ and let $c_j c_{j+1} c_{j+2} \cdots c_{j+kL-1}$ be a finite factor of length $m = kL$. Using identity (7.20), we may split the factor above in k words of length L as follows

$$\begin{aligned} c_j c_{j+1} c_{j+2} \cdots c_{j+kL-1} = &\underbrace{c_j c_{j+1} \cdots c_{j+L-1}}_{D_1} \underbrace{c_{j+L} c_{j+L+1} \cdots c_{j+2L-1}}_{D_2} \cdots \\ &\cdots \underbrace{c_{j+(k-1)L} c_{j+(k-1)L+1} \cdots c_{j+kL-1}}_{D_k} \end{aligned}$$

where the words D_i, $2 \leq i \leq k$ depend only on D_1 and $\mathbf{a}$. More precisely, we have

$$\begin{aligned} D_2 &= (c_j + a_{j+L})(c_{j+1} + a_{j+L+1}) \cdots (c_{j+L-1} + a_{j+2L-1}) \\ &\vdots \\ D_k &= (c_j + a_{j+L} + \cdots + a_{j+(k-1)L})(c_{j+1} + a_{j+L+1} + \cdots + a_{j+(k-1)L+1}) \cdots \\ &\quad (c_{j+L-1} + a_{j+2L-1} + \cdots + a_{j+kL-1}). \end{aligned}$$

Consequently, the word $c_j c_{j+1} c_{j+2} \cdots c_{j+m-1}$ depends only on D_1, which is a factor of length L, determined by $r(T)$, and on the word $a_{j+L} \cdots a_{j+kL-1}$, factor of length $kL - L = m - L$ of $\mathbf{a}$.

Now, let us return to the sequence $\mathbf{e}$. We recall that

$$\sum_{n \geq 0} e_n T^{-n} = \frac{p_2(T)}{T^{S+L-1}} \sum_{j \geq 0} c_j T^{-j}. \tag{7.21}$$

Using a similar argument as in the proof of Proposition 7.4.2 and using the identity (7.21), a factor of the form $e_j e_{j+1} \cdots e_{j+m-1}$, $j \in \mathbb{N}$, depends only on the coefficients of p_2, which are fixed, and on $c_{j-L+1} \cdots c_{j-1} c_j \cdots c_{j+m-1}$. Hence, the number of distinct factors of the form $e_j e_{j+1} \cdots e_{j+m-1}$ depends only on the number of distinct factors of the form $a_{j+1} a_{j+2} \cdots a_{j+(k-1)L}$ and on the number of factors of length L that occur in $\mathbf{c}$.

Step 3. We now describe the sequence $\mathbf{d}$, defined in (7.19).

Doing the same proof as for Proposition 7.4.2, we obtain that the number of words $d_j \cdots d_{j+m-1}$, when $j \in \mathbb{N}$, depends only on the coefficients of p_1, which are fixed, and on the number of distinct factors $a_{j-S+1} \cdots a_j \cdots a_{j+m-1}$.

Step 4. We now give an upper bound for the complexity of $\mathbf{f}$, when m is a multiple of L.

According to steps 2 and 3, the number of distinct factors of the form $f_j f_{j+1} \cdots f_{j+m-1}$, $j \in \mathbb{N}$, depends on the number of distinct factors of the form $a_{j-S+1} a_{j+2} \cdots a_{j+m-1}$ and on the number of factors of length L that occur in $\mathbf{c}$.

Consequently,

$$p(rf, m) \leq p(f, m + S - 1) q^L,$$

and by Lemma 7.2.1

$$p(f, m + S - 1) \leq p(f, m) p(f, S - 1) \leq q^{S-1} p(f, m).$$

Finally,

$$p(rf, m) \leq q^{L+S-1} p(f, m).$$

Step 5. We now give an upper bound for the complexity of $\mathbf{f}$, when m is not a multiple of L.

In this case, let us suppose that $m = kL + l$, $1 \leq l \leq L - 1$. Using Lemma 7.2.1 and according to Step 4:

$$\begin{aligned} p(rf, m) &= p(rf, kL + l) \leq p(rf, kL) p(rf, l) \leq p(rf, kL) p(rf, L - 1) \\ &\leq q^{L-1} p(rf, kL) \leq q^{S+2L-2} p(f, m). \end{aligned}$$

□

As a consequence of Propositions 7.4.1 and 7.4.3, we give a criterion of linear independence over $\mathbb{F}_q(T)$ for two Laurent series in function of their complexity.

Proposition 7.4.4. *Let $f, g \in \mathbb{F}_q((T^{-1}))$ be two irrational Laurent series such that:*

$$\lim_{m\to\infty} \frac{p(f,m)}{p(g,m)} = \infty.$$

Then f and g are linearly independent over the field $\mathbb{F}_q(T)$.

Proof. We argue by contradiction. Assume there exist polynomials $A(T)$, $B(T)$, $C(T)$ over $\mathbb{F}_q$, not all zeros, such that:

$$A(T)f(T) + B(T)g(T) + C(T) = 0.$$

Next use the fact that addition with a rational function and multiplication by a rational function do not increase the asymptotic order of complexity. Indeed, since $A(T) \neq 0$ because $g(T) \notin \mathbb{F}_q(T)$, we would have

$$f(T) + \frac{C(T)}{A(T)} = -\frac{B(T)}{A(T)}g(T).$$

However, Propositions 7.4.1 and 7.4.3 would imply that the complexity of the left-hand side of this inequality is asymptotically larger than the one of the right-hand side. □

Let us now give an example of two Laurent series that are linearly independent over $\mathbb{F}_q(T)$. Their sequences of coefficients are generated by non-uniform morphisms and we study their subword complexity in function of the order of growth of letters, using a classical result of Pansiot [101]. Notice that, the following sequences are non-automatic and hence, the associated Laurent series are transcendental over $\mathbb{F}_q(T)$.

Example 7.4.1. Consider the infinite word $\mathbf{a} = 000100010001110\cdots$; $\mathbf{a} = (a_n)_{n\geq 0} = \sigma^\infty(0)$ where $\sigma(0) = 0001$ and $\sigma(1) = 11$. If we look at the order of growth of 0 and 1 we have $|\sigma^n(0)| = 3^n + 5\cdot 2^{n-2}$ and $|\sigma^n(1)| = 2^n$. Hence, the morphism σ is exponentially diverging (see the Section (7.2.2)). Consequently, by Pansiot's theorem mentioned above, $p(\mathbf{a}, m) = \Theta(m \log m)$. Next, consider $\mathbf{b} = 010110101111010\cdots$; $\mathbf{b} = (b_n)_{n\geq 0} = \phi^\infty(0)$, where $\phi(0) = 0101$ and $\phi(1) = 11$. It is not difficult to see that ϕ is polynomially diverging (see Section (7.2.2)) since $|\phi^n(0)| = (n+1)2^n$ and $|\phi(1)^n| = 2^n$. By Pansiot's theorem, $p(\mathbf{b}, m) = \Theta(m \log\log m)$.

Now we consider the formal series whose coefficients are the sequences generated by the morphisms above:

$$f(T) = \sum_{n\geq 0} a_n T^{-n} = \frac{1}{T^3} + \frac{1}{T^7} + \frac{1}{T^{11}} + \frac{1}{T^{12}} + \cdots \in \mathbb{F}_q[[T^{-1}]]$$

and

$$g(T) = \sum_{n \geq 0} b_n T^{-n} = \frac{1}{T^1} + \frac{1}{T^3} + \frac{1}{T^4} + \frac{1}{T^6} + \cdots \in \mathbb{F}_q[[T^{-1}]].$$

Since $\lim_{m \to \infty} p(f,m)/p(g,m) = +\infty$, Proposition 7.4.4 implies that f and g are linearly independent over $\mathbb{F}_q(T)$.

7.4.2 Other closure properties

In this section we prove that both classes $\mathcal{P}$ and $\mathcal{Z}$ are closed under various natural operations such as the Hadamard product, the formal derivative and the Cartier operators.

Hadamard product

Let $f(T) := \sum_{n \geq -n_1} a_n T^{-n}$, $g(T) := \sum_{n \geq -n_2} b_n T^{-n}$ be two Laurent series in $\mathbb{F}_q((T^{-1}))$. The Hadamard product of f and g is defined as follows

$$f \odot g = \sum_{n \geq -\min(n_1, n_2)} a_n b_n T^{-n}.$$

As in the case of addition of two Laurent series (see Proposition 7.4.1) one can easily obtain the following.

Proposition 7.4.5. *Let f and g be two Laurent series belonging to $\mathbb{F}_q((T^{-1}))$. Then, for every $m \in \mathbb{N}$, we have*

$$\frac{p(f,m)}{p(g,m)} \leq p(f \odot g, m) \leq p(f,m)p(g,m).$$

The proof is similar to the one of Proposition 7.4.1. The details are left to the reader. Note that, as in the case of addition, it is possible to attain the bounds in Proposition 7.4.5 (see Remark 7.4.3).

Formal derivative

As an easy application of Proposition 7.4.5, we present here the following result. First, let us recall the definition of the formal derivative.

Definition 7.4.1. *Let $n_0 \in \mathbb{N}$ and consider $f(T) = \sum_{n=-n_0}^{+\infty} a_n T^{-n} \in \mathbb{F}_q((T^{-1}))$. The formal derivative of f is defined as follows*

$$f'(T) = \sum_{n=-n_0}^{+\infty} (-n \mod p) a_n T^{-n+1} \in \mathbb{F}_q((T^{-1})).$$

We prove the following result.

Proposition 7.4.6. *Let $f(T) \in \mathbb{F}_q((T^{-1}))$ and k be a positive integer. If $f^{(k)}$ is the derivative of order k of f, then there exists a positive constant M, such that, for all $m \in \mathbb{N}$, we have*

$$p(f^{(k)}, m) \leq M \; p(f,m).$$

Proof. The derivative of order k of f is almost the Hadamard product of the Laurent series by a rational function. By definition of $p(f,m)$, we may suppose that $f(T) := \sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_q[[T^{-1}]]$. Then

$$f^{(k)}(T) = \sum_{n\geq k} ((-n)(-n-1)\cdots(-n-k+1)a_n)T^{-n-k} := T^{-k}\sum_{n\geq k} b_n a_n T^{-n},$$

where $b_n := (-n)(-n-1)\cdots(-n-k+1) \mod p$. Since $b_{n+p} = b_n$, the sequence $(b_n)_{n\geq 0}$ is periodic of period p. Hence, let $g(T)$ denote the formal power series whose coefficients are precisely given by $(b_n)_{n\geq 0}$. Thus there exists a positive constant M such that:

$$p(g,m) \leq M.$$

By Proposition 7.4.5,

$$p(f^{(k)},m) \leq p(g,m)p(f,m) \leq Mp(f,m),$$

which completes the proof. □

Cartier's operators

In the fields of positive characteristic, there is a natural operator, the so-called "Cartier operator" that plays an important role in many problems in algebraic geometry and arithmetic in positive characteristic [46, 45, 56, 112]. In particular, if we consider the field of Laurent series with coefficients in $\mathbb{F}_q$, we have the following definition.

Definition 7.4.2. *Let $f(T) = \sum_{i\geq 0} a_i T^{-i} \in \mathbb{F}_q[[T^{-1}]]$ and $0 \leq r < q$. The Cartier operator Λ_r is a linear transformation such that:*

$$\Lambda_r(\sum_{i\geq 0} a_i T^{-i}) = \sum_{i\geq 0} a_{qi+r} T^{-i}.$$

The classes $\mathcal{P}$ and $\mathcal{Z}$ are closed under this operator. More precisely, we prove the following result.

Proposition 7.4.7. *Let $f(T) \in \mathbb{F}_q[[T^{-1}]]$ and $0 \leq r < q$. Then there is M such that, for every $m \in \mathbb{N}$ we have the following*

$$p(\Lambda_r(f),m) \leq qp(f,m)^q.$$

Proof. Let $\mathbf{a} := (a_n)_{n\geq 0}$ be the sequence of coefficients of f and $m \in \mathbb{N}$. In order to compute $p(\Lambda_r(f),m)$, we have to look at factors of the form

$$a_{qj+r}a_{qj+q+r}\cdots a_{qj+(m-1)q+r},$$

for all $j \in \mathbb{N}$. But these only depend on factors of the form

$$a_{qj+r}a_{qj+r+1}\cdots a_{qj+(m-1)q+r}.$$

Using Lemma 7.2.1, we obtain that:

$$p(\Lambda_r(f),m) \leq p(f,(m-1)q+1) \leq qp(f,m-1)^q \leq qp(f,m)^q.$$

□

7.5 Cauchy product of Laurent series

In the previous section, we proved that $\mathcal{P}$ and $\mathcal{Z}$ are vector spaces over $\mathbb{F}_q(T)$. This naturally raises the question whether or not these classes form a ring, *i.e.*, whether they are closed under the usual Cauchy product. There are actually some particular cases of Laurent series with low complexity whose product still belongs to $\mathcal{P}$.

In this section we discuss the case of automatic Laurent series. However, we are not able to prove whether $\mathcal{P}$ or $\mathcal{Z}$ are or not rings or fields.

7.5.1 Products of automatic Laurent series

A particular case of Laurent series stable by multiplication is the class of k-automatic series, k being a positive integer:

$$\mathrm{Aut}_k = \{f(T) = \sum_{n\geq 0} a_n T^{-n} \in \mathbb{F}_q((T^{-n})),\ \mathbf{a} = (a_n)_{n\geq 0} \text{ is } k\text{-automatic}\}.$$

Since any k-automatic sequence has at most linear complexity, $\mathrm{Aut}_k \subset \mathcal{P}$. A theorem of Allouche and Shallit [19] states that the set Aut_k is a ring. In particular, this implies that, if f and g belong to Aut_k, then $p(fg, m) = O(m)$.

Notice also that, in the case where k is a power of p, the characteristic of the field $\mathbb{F}_q((T^{-1}))$, the result follows from Christol's theorem.

***Remark* 7.5.1.** However, we do not know whether or not this property is still true if we replace Aut_k by $\cup_{k\geq 2}\mathrm{Aut}_k$. More precisely, if we consider two Laurent series f and g, which are respectively k-automatic and l-automatic, k and l being multiplicatively independent, we do not know if the product fg still belongs to $\mathcal{P}$. In the sequel, we give a particular example of two such Laurent series for which we prove that their product is still in $\mathcal{P}$.

We now focus on the product of Laurent series of the form:

$$f(T) = \sum_{n\geq 0} T^{-d^n} \in \mathbb{F}_q((T^{-1})).$$

It is not difficult to prove that $p(f, m) = O(m)$. The reader may refer to [67] for more general results concerning the complexity of lacunary formal power series. The fact that the complexity of f is at most linear is also implied by the fact that $f \in \mathrm{Aut}_d$. Notice also that f is transcendental over $\mathbb{F}_q(T)$ if q is not a power of d. This is an easy consequence of Christol's theorem and a theorem of Cobham [49].

In this section we will prove the following result.

Theorem 7.5.1. *Let d and e be two multiplicatively independent positive integers (that is $\frac{\log d}{\log e}$ is irrational) and let $f(T) = \sum_{n\geq 0} T^{-d^n}$ and $g(T) = \sum_{n\geq 0} T^{-e^n}$ be two Laurent series in $\mathbb{F}_q((T^{-1}))$. Then:*

$$p(fg, m) = O(m^4).$$

Let $h(T) := f(T)g(T)$. Then $h(T) = \sum_{n\geq 0} a_n T^{-n}$ where the sequence $\mathbf{a} = (a_n)_{n\geq 0}$ is defined as follows

$$a_n := (\text{the number of pairs } (k,l) \in \mathbb{N}^2 \text{ that verify } n = d^k + e^l) \mod p.$$

The main clue of the proof is the following consequence of the theory of S-unit equations (see [2] for a proof).

Lemma 7.5.1. *Let d and e be two multiplicatively independent positive integers. There is a finite number of solutions $(k_1, k_2, l_1, l_2) \in \mathbb{N}^4$, $k_1 \neq k_2$, $l_1 \neq l_2$, that satisfy the equation:*

$$d^{k_1} + e^{l_1} = d^{k_2} + e^{l_2}.$$

Obviously, we have the following consequence concerning the sequence $\mathbf{a} = (a_n)_{n\geq 0}$:

Corollary 7.5.1. *There exists a positive integer N such that, for every $n \geq N$ we have $a_n \in \{0, 1\}$. Moreover, $a_n = 1$ if, and only if, there exists one unique pair $(k, l) \in \mathbb{N}^2$ such that $n = d^k + e^l$.*

We now prove Theorem 7.5.1. For the sake of simplicity, we consider $d = 2$ and $e = 3$, but the proof is exactly the same in the general case.

Proof. Let $\mathbf{b} := (b_n)_{n\geq 2}$ and $\mathbf{c} := (c_n)_{n\geq 2}$ be the sequences defined as follows

$$b_n = \begin{cases} 1 & \text{if there exists a pair } (k,l) \in \mathbb{N}^2 \text{ such that } n = 2^k + 3^l,\ 2^k > 3^l; \\ 0 & \text{otherwise,} \end{cases}$$

$$c_n = \begin{cases} 1 & \text{if there exists a pair } (k,l) \in \mathbb{N}^2 \text{ such that } n = 2^k + 3^l,\ 2^k < 3^l; \\ 0 & \text{otherwise.} \end{cases}$$

Let us denote by $h_1(T) := \sum_{n\geq 2} b_n T^{-n}$ and resp. $h_2(T) := \sum_{n\geq 2} c_n T^{-n}$ the Laurent series associated with $\mathbf{b}$ and $\mathbf{c}$. Using Corollary 7.5.1, there exists a polynomial $P \in \mathbb{F}_q[T]$, with degree less than N, such that h can be written as follows

$$h(T) = h_1(T) + h_2(T) + P(T).$$

By Remark 7.4.2, there is $C \in \mathbb{R}$ such that, for any $m \in \mathbb{N}$:

$$p(h, m) \leq p(h_1 + h_2, m) + C.$$

In the sequel, we will show that $p(h_1, m) = p(h_2, m) = O(m^2)$. Theorem 7.5.1 will then follow by Proposition 7.4.1.

We now study the subword complexity of the sequence of coefficients $\mathbf{b} := (b_n)_{n\geq 2}$. The proof is similar to the proof of Theorem 7.1.2. The complexity of the sequence $\mathbf{c}$ can be treated in essentially the same way as for $\mathbf{b}$.

Step 1. For all $n \geq 1$, we let W_n denote the factor of $\mathbf{b}$ that occurs between positions $2^n + 1$ and 2^{n+1}, that is

$$W_n := b_{2^n+3^0} b_{2^n+2} b_{2^n+3^1} \cdots b_{2^{n+1}}.$$

We also set $W_0 := 1$.

Observe that $|W_n| = 2^n$.

With this notation the infinite word $\mathbf{b}$ can be factorized as:

$$\mathbf{b} = \underbrace{1}_{W_0}\underbrace{10}_{W_1}\underbrace{1010}_{W_2}\underbrace{10100000}_{W_3}\cdots = W_0W_1W_2\cdots . \tag{7.22}$$

Step 2. Let $n \geq 1$ and m_n be the greatest integer such that $2^n + 3^{m_n} \leq 2^{n+1}$. This is equivalent to saying that m_n is such that

$$2^n + 3^{m_n} < 2^{n+1} < 2^n + 3^{m_n+1}.$$

Notice also that $m_n = n\lfloor \log_3 2 \rfloor$.

With this notation we have (for $n \geq 5$)

$$W_n = 1010^5 1 \cdots 10^{\alpha_i} \cdots 10^{\alpha_{m_n}} 10^{\beta_n},$$

where $\alpha_i = 2 \cdot 3^{i-1} - 1$, for $1 \leq i \leq m_n$, and $\beta_n = 2^n - 3^{m_n} \geq 0$.

Let us denote by U_n the prefix of W_n such that $W_n := U_n 0^{\beta_n}$.

Notice that $(m_n)_{n\geq 0}$ is an increasing sequence. Hence $(\alpha_{m_n})_{n\geq 0}$ is increasing. Consequently, $U_n \prec_p U_{n+1}$ and more generally, $U_n \prec_p W_i$, for every $i \geq n + 1$.

Step 3. Let $M \in \mathbb{N}$. Our aim is to give an upper bound for the number of distinct factors of length M occurring in $\mathbf{b}$. In order to do this, we will show that there exists an integer N such that all these factors occur either in

$$W_0W_1 \cdots W_N$$

or

$$A_0 := \{Z \in \mathcal{A}^M;\ Z \text{ is of the form } 0^j P \text{ or } 0^i 10^j P,\ P \prec_p U_N,\ i, j \geq 0, \}.$$

Let $N = \lceil \log_2(M + 1) \rceil + 3$. Doing a simple computation we obtain that $\alpha_{m_N} \geq M$. Notice also that, for any $i \geq N$ we have

$$\alpha_{m_i} \geq M.$$

This follows since $(\alpha_{m_n})_{n\geq 0}$ is an increasing sequence.

Let V be a factor of length M of $\mathbf{b}$. Suppose that V does not occur in the prefix $W_0W_1 \cdots W_N$. Then, by (7.22), V must occur in $W_NW_{N+1} \cdots$. Hence, V must appear in some W_i, for $i \geq N + 1$, or in $\bigcup_{i\geq N} i(W_i, \varepsilon, W_{i+1})$.

Let us suppose that V occurs in $\bigcup_{i\geq N} i(W_i, \varepsilon, W_{i+1})$. Since W_i ends with $0^{\alpha_{m_i}} 10^{\beta_i}$, with $\alpha_{m_i} \geq M$, and since W_{i+1} begins with U_N and $|U_N| = 3^{m_N} + 1 \geq M$, we have

$$\mathcal{A}^M \cap \Big(\bigcup_{i\geq N} i(W_i, \varepsilon, W_{i+1})\Big) \subset A_0.$$

Hence, if V occurs in $\bigcup_{i\geq N} i(W_i, \varepsilon, W_{i+1})$ then $V \in A_0$.

Let us suppose that V occurs in some W_i, for $i \geq N+1$. By definition of W_i and α_i, for $i \geq N+1$ and by the fact that we have

$$W_i = 1010^5 1 \cdots 10^{\alpha_{m_N}} 10^{\alpha_{m_N+1}} \cdots 10^{\alpha_{m_i}} 10^{\beta_i} = U_N 0^{\alpha_{m_N+1}} \cdots 10^{\alpha_{m_i}} 10^{\beta_i}.$$

By assumption, V does not occur in $W_0W_1\cdots W_N$; hence V cannot occur in U_N which by definition is a prefix of W_N. Consequently, V must be of the form 0^r10^s, $r, s \geq 0$. Indeed, since $\alpha_{m_N} \geq M$, all blocks of zeros that follow after U_N (and before the last digit 1 in W_i) are all longer than M. But the words of the form 0^r10^s, $r, s \geq 0$ belong also to A_0.

Hence, we proved that if V does not occur in the prefix $W_0W_1\cdots W_N$, then V belongs to A_0, as desired.

Step 4. In the previous step we showed that all distinct factors of length M occur in the prefix $W_0W_1\cdots W_N$ or in the set A_0.

Since

$$|W_0W_1\cdots W_N| = \sum_{i=0}^{N} 2^i = 2^{N+1} - 1$$

and since $N = \lceil \log_2(M+1) \rceil + 3$ we have

$$2^{N+1} - 1 \leq 2^{\log_2(M+1)+5} - 1 = 32M + 31,$$

and the number of distinct factors that occur in $W_0W_1\cdots W_N$ is less or equal to $32M + 31$.

Also, by an easy computation, we obtain that the cardinality of the set A_0 is

$$\operatorname{Card} A_0 = \frac{M^2}{2} + \frac{3M}{2}.$$

Finally, $p(\mathbf{b}, m) = p(h_1, m) = O(m^2)$. In the same manner, one could prove that $p(h_2, m) = O(m^2)$. This achieves the proof of Theorem 7.5.1, in view of Proposition 7.4.1. □

7.5.2 A more difficult case

Set

$$\theta(T) := 1 + 2\sum_{n\geq 1} T^{-n^2} \in \mathbb{F}_q((T^{-1})),\ q \geq 3.$$

The function $\theta(T)$ is related to the classical Jacobi theta function. One can easily prove that:

$$p(\theta, m) = \Theta(m^2).$$

In particular this implies the transcendence of $\theta(T)$ over $\mathbb{F}_q(T)$, for any $q \geq 3$. Notice that this also implies the transcendence over $\mathbb{Q}(T)$ of the same Laurent series, but viewed as an element of $\mathbb{Q}((T^{-1}))$. Since $\theta(T) \in \mathcal{P}$, it would be interesting to know whether or not $\theta(T)^2$ belongs also to $\mathcal{P}$. Notice that

$$\theta(T)^2 = \sum_{n\geq 1} r_2(n) T^{-n}$$

where $r_2(n)$ is the number of representations of n as sum of two squares of integers mod p. In the rich bibliography concerning Jacobi theta function (see, for instance, [59, 73]), there is the following well-known formula

$$r_2(n) = 4(d_1(n) - d_3(n)) \mod p$$

where $d_i(n)$ denotes the number of divisors of n congruent to i modulo 4, for $i \in \{1, 3\}$, and we may easily deduce that $r_2(n)$ is a multiplicative function of n. Recall that we would like to study the subword complexity of $r_2(n)_{n\geq 0}$, that is the number of distinct factors of the form $r_2(j)r_2(j+1)\cdots r_2(j+m-1)$, when $j \in \mathbb{N}$. Hence, it would be useful to describe some additive properties of $r_2(n)_{n\geq 0}$; for instance, it would be interesting to find some relations between $r_2(j+N)$ and $r_2(j)$, for some positive integers j and N. This seems to be a rather difficult question about which we are not able to say anything conclusive.

7.6 Conclusion

It would be also interesting to investigate the following general question: *is it true that Carlitz's analogs of classical constants all have a "low" complexity (i.e., polynomial or subexponential)?*

The first clue in this direction are the examples provided by Theorems 7.1.1 and 7.1.2. Notice also that a positive answer would reinforce the differences between $\mathbb{R}$ and $\mathbb{F}_q((T^{-1}))$ as hinted in our Introduction. When investigating these problems, we need, in general, the Laurent series expansions of such functions. In this context, one has to mention the work of Berthé [29, 30, 27, 28], where some Laurent series expansions of Carlitz's functions are described.

A Other examples of products of Laurent series

Let $f(T) = \sum_{i=1}^{\infty} a_i\, T^{-i} \in \mathbb{F}_q[[T^{-1}]]$ be a formal power series whose coefficients take only the values 0 and 1. In this part, we give some sufficient conditions on the infinite word $\mathbf{a}$ in order to guarantee that $f^2(T)$ has at most linear complexity.

Since we consider Laurent series with coefficients 0 an 1, a relevant notion is the 1-distribution function, whose definition is recalled below.

Definition A.1. *Let $\boldsymbol{a} := (a_i)_{i\geq 0}$ be an infinite sequence over the binary alphabet $\{0,1\}$. The function $G : \mathbb{N} \to \mathbb{N}$ is called the 1-distribution function of $\boldsymbol{a}$ if $G(n)$ denotes the nth occurrence of 1 in $\boldsymbol{a}$. By convention $G(0) = 0$. The function $g(n) = G(n) - G(n-1)$, defined for $n \geq 0$, is called* the gap function. *If g is strictly increasing then $\boldsymbol{a}$ is said to be gap increasing.*

In [67], the author proved that the subword complexity of a gap increasing word $\mathbf{a}$ satisfies the following inequalities:

$$m + 1 \leq p(\mathbf{a}, m) \leq \lceil m/2 \rceil \lceil m/2 \rceil + \lceil m/2 \rceil + 1,$$

for every $m \geq 0$. Furthermore, both bounds are optimal.

Our aim is to prove the following result.

Theorem A.1. *Let $f(T) = \sum_{i\geq 0} a_i T^{-i}$ be a formal power series such that $\boldsymbol{a} := (a_i)_{i\geq 0} \in \{0,1\}^{\infty}$. Let $G : \mathbb{N} \to \mathbb{N}$ be the 1-distribution function of $\boldsymbol{a}$. If*

$$\lim_{n\to\infty} \frac{G(n)}{G(n-1)} = l > 2,$$

then $f(T)$ and $f^2(T)$ have at most linear complexity.

Proposition A.1. *Let $f(T) = \sum_{i\geq 0} a_i T^{-i}$ be a formal power series such that $\boldsymbol{a} := (a_i)_{i\geq 0} \in \{0,1\}^{\infty}$. Let $G : \mathbb{N} \to \mathbb{N}$ be the 1-distribution function of $\boldsymbol{a}$. We assume that the following hypothesis is verified:*

$$\begin{array}{c} \text{if } G(m) - G(n) = G(m') - G(n'), \text{ then either} \\ (m = m' \text{ and } n = n') \text{ or } (m = n \text{ and } m' = n'). \end{array} \quad \text{(H)}$$

Then, $f^2(T) = \sum_{k\geq 0} b_k T^{-k}$ is such that $\boldsymbol{b} := (b_k)_{k\geq 0} \in \{0,1,2\}^{\mathbb{N}}$.

***Remark* A.1.** This proposition must be understood as follows. Let us consider the Laurent series $f(T) = \sum_{n\geq 0} T^{-G(n)}$. Notice that to know the coefficients of the formal power series $f^2(T)$ is equivalent to know the number of ordered ways of writing n as a sum of two terms $G(k)$ and $G(l)$, $k, l \geq 0$. The hypothesis (H) means that there are at most two ways of writing an integer n as a sum of two terms $G(k)$ and $G(l)$.

Proof of Proposition A.1. By definition, $f^2(T) = \sum_{k\geq 0} b_k T^{-k}$, where the sequence $(b_k)_{k\geq 0}$ is the convolution product

$$b_k = a_0 a_k + a_1 a_{k-1} + \cdots + a_k a_0.$$

Since $a_i \in \{0,1\}$, for any $i \in \mathbb{N}$, then b_k may be defined by the following formula:

$$b_k = \#\underbrace{\{i \leq k \text{ such that } a_i = a_{k-i} = 1\}}_{B_k}.$$

More precisely, we will prove that

$$b_k = \begin{cases} 1, \text{ if there exists } i_1 \text{ such that } k = 2G(i_1) \\ 2, \text{ if there exist } i_1 \text{ and } i_2, i_1 < i_2 \text{ such that } k = G(i_1) + G(i_2) \\ 0, \text{ otherwise.} \end{cases}$$

First, notice that if k is odd, then b_k is even. Indeed, in that case there is no i such that $i = k - i$ and hence, each time i belongs to B_k, the index $k - i$, which is different from i, also belongs to B_k. If k is even and $a_{k/2} = 1$ then b_k is odd (B_k contains i and $k - i$, for $i \neq k/2$ and, since $a_{k/2} = 1$, then $k/2$ also belongs to the set B_k). If $a_{k/2} = 0$, then, b_k is even because each time $i \in B_k$, we have $k - i \in B_k$ and $i \neq k - i$.

Next, we will show that the sequence $(b_k)_{k\geq 0}$ only takes values in $\{0,1,2\}$. Let us suppose that there exists k such that $b_k \geq 3$. We have to distinguish different cases.

If k is odd then $b_k \geq 4$. This means that there exists $i \neq j$, $i + j \neq k$, such that $a_i = a_j = a_{k-i} = a_{k-j} = 1$. In other words, there exist four different integers i_1, i_2, i_3, i_4 such that $i = G(i_1)$, $j = G(i_2)$, $k - i = G(i_3)$, $k - j = G(i_4)$. Hence, we obtain that $G(i_1) - G(i_2) = G(i_4) - G(i_3)$, which contradicts Hypothesis (H).

If k is even and $a_{k/2} = 1$, then, b_k is odd and, if we suppose that b_k is greater than or equal to 3, there exist two different integers $i \neq k/2$ such that $a_i = a_{k-i} = a_{k/2} = 1$. This implies that there are three numbers i_1, i_2, i_3 such that any two of them are distinct and $i = G(i_1)$, $k - i = G(i_2)$ and $k/2 = G(i_3)$. Hence $G(i_1) - G(i_3) = G(i_3) - G(i_2)$, which contradicts (H). The case where k is even and $a_{k/2} = 0$ is similarly treated.

It is now clear that $b_k = 1$ if, and only if, the word

$$a_0 a_1 \cdots a_k = \overbrace{00\cdots0}^{G(i_1)}1\overbrace{0\cdots00}^{G(i_1)}$$

and so $k = 2G(i_1)$.

At last, we have to find the integers k such that $b_k = 2$. As we have done before, we remark that $b_k = 2$ if, and only if, the word

$$a_0 a_1 \cdots a_k = \overbrace{\underbrace{00\cdots0}_{G(i_1)}100\cdots00}^{G(i_2)}1\underbrace{0\cdots0}_{G(i_1)}.$$

Thus $k = G(i_1) + G(i_2)$ and $i_2 > i_1$ by definition of the function G. This completes the proof. □

***Remark* A.2.** (H) is a necessary and sufficient condition for $f^2(T)$ to remain on the alphabet $\{0, 1, 2\}$. Indeed, if the equation $G(m)-G(n) = G(m')-G(n')$ has a non trivial solution, then, the digit 3 or 4 appears as a coefficient of the formal power series $f^2(T)$.

Proof of Theorem A.1(sketch). The hypothesis concerning the limit implies that, for n large enough, $G(n) \geq 2G(n-1)$. This immediately implies that the complexity of f is at most linear.

On the other hand, this also implies that the equation $G(m) - G(n) = G(m')-G(n')$ has only trivial solutions (m, n, m', n') with $m = n$ an $m' = n'$ or $m = m'$ and $n = n'$ for $\max(m, n, m', n')$ large enough. Indeed, let (m, n, m', n') be a solution of $G(m) - G(n) = G(m') - G(n')$ and let us suppose that $m > n$ and $m > m'$. Since $n < m$ and $m' < m$ we have $G(n) \leq G(m-1)$ and $G(m') \leq G(m-1)$. This implies that $G(m) = G(n) + G(m') - G(n') \leq 2G(m-1) - G(n') < 2G(m-1)$ and this is not true for m large enough.

Following the same idea of the proposition above, one could prove that the sequence of coefficients of $f^2(T) = \sum_{n\geq 0} b_n T^{-n}$ is defined, for n large enough, by

$$b_n = \begin{cases} 1 \text{ if there exists } k > 1 \text{ such that } n = 2G(k) \\ 2 \text{ if there exist } m \text{ and } l, l \neq m \text{ such that } n = G(m) + G(l) \\ 0 \text{ otherwise.} \end{cases}$$

Next, we split $\mathbf{b} = (b_n)_{n\geq 0}$ into subwords W_n defined by

$$W_n = b_{G(n)} b_{G(n)+1} b_{G(n)+2} \cdots b_{G(n+1)-1}.$$

Thus, we can write $\mathbf{b}$ using this "canonical" decomposition as $\mathbf{b} = W_0 \cdots W_n \cdots$. Notice that the length of W_n is $|W_n| = G(n+1) - G(n)$.

Now, fix $M \in \mathbb{N}$, large enough. There exists N such that $G(N) \leq M < G(N+1)$. It is not difficult to prove that all different factors of length M occur in the prefix $W_0 W_1 \cdots W_{N+3}$. Roughly speaking, this can be explained by the fact that in $W_{N+4} W_{N+5} \cdots$ the different factors which appear are formed by blocks of zeros with length greater that $G(N+1)$, hence greater than M. On the other hand, one has

$$\begin{aligned} W_{N+3} &= b_{G(N+3)} \cdots b_{G(N+3)+G(N+1)} 00 \cdots 0 b_{G(N+3)+G(N+3)} 0 \cdots 0 b_{G(N+4)-1} \\ &= b_{G(N+3)} \cdots 2 \underbrace{00 \cdots 0}_{\text{more than } M \text{ times}} 2 \underbrace{00 \cdots 0}_{\text{more than } M \text{ times}} 1 \underbrace{00 \cdots 0}_{\text{more than } M \text{ times}} b_{G(N+4)-1}, \end{aligned}$$

and thus, all possible words of length M of the form $0^{m_1}10^{m_2}$ and $0^{m_1}20^{m_2}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = M - 1$, occur in W_{N+3}.

The length of the prefix $W_0W_1\cdots W_{N+3}$ is $G(N+3)$. Using the fact that

$$\lim_{n\to\infty}\frac{G(n)}{G(n-1)}=l,$$

we have, for n large enough, $G(n)\leq lG(n-1)$. Thus, we obtain that

$$p(f^2,M)\leq(l+1)^3M$$

and, hence, $p(f^2,M)=O(M)$. □

We now apply the technique already appeared in the proof of the previous theorem in order to bound up the complexity of the square of a well-known Laurent series, whose 1-distribution function is given by the Fibonacci sequence. Note that this result is not covered by Theorem A.1.

Proposition A.2. *Let $f(T)=\sum_{n\geq 2}T^{-F_n}$ where F_n denotes the $n-$th Fibonacci number. Then $f(T)$ and $f^2(T)$ have at most linear subword complexity.*

Proof. The first claim immediately follows. Indeed, the sequence of coefficients of f is the infinite binary word $\mathbf{a}=a_0a_1a_2\cdots a_n\cdots$ defined by

$$a_n=\begin{cases}1, \text{ if there exists } k\geq 1 \text{ such that } n=F(k),\\ 0, \text{ otherwise.}\end{cases}$$

The gap function of $\mathbf{a}$ is $g(n)=F(n+1)-F(n)=F(n-1)$ and we can deduce from [67] that $p(\mathbf{a},m)=\Theta(m)$.

If $f(T)=\sum_{n\geq 0}a_nT^{-n}$, then the coefficients of $f^2(T)=\sum_{n\geq 0}b_nT^{-n}$ are defined by the convolution product

$$b_n=a_0a_n+a_1a_{n-1}+\cdots+a_na_0, \text{ for } n\in\mathbb{N}.$$

Since $a_i\in\{0,1\}$, for all $i\in\mathbb{N}$, we may define b_n by the following formula:

$$b_n=\#\underbrace{\{i\leq n \text{ such that } a_i=a_{n-i}=1\}}_{B_n}.$$

***Remark* A.3.** From a combinatorial point of view, b_n is the number of 1 which occur at the same position in the word $a_0a_1\cdots a_n$ and in its mirror image $a_na_{n-1}\cdots a_1a_0$.

In the sequel, we show that b_n may be defined by the following formula:

$$b_n=\begin{cases}3 \text{ if there exists } k>1 \text{ such that } n=2F(k)\\ 2 \text{ if there exist } m \text{ and } 1<l<m,\ l\neq m-3 \text{ such that } n=F(m)+F(l)\\ 1 \text{ if } n=2\\ 0 \text{ otherwise.}\end{cases}$$

First of all, we prove that the sequence $\mathbf{b} = (b_n)_{n\geq 1}$ takes values in $\{0,1,2,3\}$. Let us assume by contradiction that there is a positive integer n such that $b_n \geq 4$. By definition of b_n, there would exist i, j, $i \neq j$, $n \neq i+j$, such that $a_i = a_j = a_{n-i} = a_{n-j} = 1$. In other words, there would exist four integers $k < l < s < m$ such that:

$$F(l) - F(k) = F(m) - F(s).$$

This would implies that $F(m) = F(s) + F(l) - F(k)$. Since $s \leq m-1$ and $l \leq m-2$, we would have that $F(l) \leq F(m-2)$ and $F(s) \geq F(m-1)$. Thus, $F(m) \leq F(m-1)+F(m-2)-F(k) < F(m-1)+F(m-2)$, which would provide a contradiction with the well-known relation $F(m) = F(m-1) + F(m-2)$.

***Remark* A.4.** If we consider the equation $F(m) + F(s) = F(l) + F(k)$, for $m, s, l, k \geq 2$, then, one can prove that the only non trivial solution (m, s, k, l) is on the form $(m, m-3, m-1, m-1)$ (and of course $(m-3, m, m-1, m-1)$, $(m-1, m-1, m, m-3)$, $(m-1, m-1, m-3, m)$), for $m \geq 3$.

By the previous remark, $b_n = 3$ if, and only if, the word

$$a_0a_1\cdots a_n = \underbrace{\overbrace{\underbrace{00\cdots0}_{F(m-3)}100\cdots00}^{F(m-1)}10\cdots001}_{F(m)}\underbrace{00\cdots00}_{F(m-3)}.$$

Thus $b_n = 3$ if, and only if, there is an integer m such that $n = F(m) + F(m-3) = 2F(m-1)$. Using a similar argument, we deduce that $b_n = 2$ if, and only if, the word

$$a_0a_1\cdots a_n = \overbrace{\underbrace{00\cdots0}_{F(l)}100\cdots00}^{F(m)}1\underbrace{0\cdots0}_{F(l)}.$$

Thus, $n = F(m)+F(l)$, with $l < m$ and $l \neq m-3$. (Indeed, if $l = m-3$ then, by the previous case $b_n = 3$.)

We remark that b_n is the number of ordered ways of writing n as a sum of two Fibonacci numbers. For example, $b_n = 3$ if $n = F(k-2) + F(k+1) = F(k)+F(k) = F(k+1)+F(k-2)$. The infinite sequence $\mathbf{b} = (b_n)_{n\geq 0}$ appears as A121549 in Sloane's On-Line Encyclopedia of Integer Sequences [115].

Let now consider the infinite word $\mathbf{b} = b_0b_1b_2\cdots b_n\cdots$ over the alphabet $\{0,1,2,3\}$. We show that $\mathbf{b}$ has a linear subword complexity.

For $n \geq 3$, we let W_n denote the word of length $F(n-1)$ that occurs at position $F(n)+1 = F(n)+F(2) = F(n)+F(1)$ in $\mathbf{b}$, *i.e.*,

$$W_n = b_{F(n)+1}\cdots b_{F(n)+F(n-1)}.$$

Using this notation, the infinite word $\mathbf{b}$ can be written as follows

$$\mathbf{b} = 001\underbrace{2}_{W_3}\underbrace{32}_{W_4}\underbrace{322}_{W_5}\underbrace{23202}_{W_6}\underbrace{22302002}_{W_7}\underbrace{2220300200002}_{W_8}\cdots.$$

For every $n \geq 9$ one can easily prove that

$$W_n = 222020^{F(3)-1}20^{F(4)-1}20^{F(5)-1}20\cdots020^{F(n-5)-1}30^{F(n-4)-1}20^{F(n-3)-1}2.$$

Indeed, by definition of $\mathbf{b}$, $b_{F_n+F_k} = 2$ if $k < n$, $k \neq n-3$ and $b_{F_n+F_{n-3}} = 3$. The number of zeros between $b_{F(n)+F(k)}$ and $b_{F(n)+F(k+1)}$ is $F(k+1) - F(k) - 1 = F(k-1) - 1$.

Let us fix a positive integer M. Without loss of generality, we may assume that $M \geq 9$. Then there is a positive integer N such that $F(N) \leq M < F(N+1)$. We now want to count the number of subwords of length M occurring in $\mathbf{b}$.

The main idea is to show that all the words of length M occur in the prefix $001W_3W_4\cdots W_{N+6}$.

Notice that we cannot choose a shorter prefix because the word $\underbrace{000\cdots0}_{M-1}3$ first occurs in W_m. It is not difficult to see that, after the prefix

$$001W_3W_4\cdots W_{N+6},$$

all the words of length M have already been seen before.

Hence, we obtain

$$p(\mathbf{b}, M) \leq 3 + F(2) + F(3) + \cdots + F(N+5).$$

On the other hand, by induction on n, we can obtain the following identity

$$\sum_{i=0}^{N} F(i) = F(N+2) - 1.$$

Hence $p(\mathbf{b}, M) \leq F(N+6)$ and a simple computation now gives us

$$F(N+6) = 13F(N) + 8F(N-1).$$

Hence $p(\mathbf{b}, M) \leq 13F(N)+8F(N-1) < 21F(N) \leq 21M$ and since the infinite word $\mathbf{b}$ is not eventually periodic, the subword complexity function is at most linear. □

Bibliographie

[1] B. Adamczewski, On the expansion of some exponential periods in an integer base, *Math. Ann.* **346** (2010), 107–116.

[2] B. Adamczewski, J. Bell, Function fields in positive characteristic : expansions and Cobham's theorem, *J. Algebra* **319** (2008), 2337–2350.

[3] B. Adamczewski, Y. Bugeaud, Real and p-adic expansions involving symmetric patterns, *Int. Math. Res. Not.*, (2006).

[4] B. Adamczewski, Y. Bugeaud, On the complexity of algebraic numbers I. Expansions in integer bases, *Ann. of Math.* **165** (2007), 547–565.

[5] B. Adamczewski, Y. Bugeaud, Nombres réels de complexité sous-linéaire : mesures d'irrationalité et de transcendance, à parraître dans *J. Reine Angew. Math.*.

[6] B. Adamczewski, J. Bell, On vanishing coefficients of algebraic power series over fields of positive characteristic, preprint.

[7] B. Adamczewski, Y. Bugeaud, F. Luca, Sur la complexité des nombres algébriques, *C. R. Math. Acad. Sci. Paris* **339** (2004), 11–14.

[8] B. Adamczewski, J. Cassaigne, Diophantine properties of real numbers generated by finite automata, *Compositio Math.* **142** (2006), 1351–1372.

[9] B. Adamczewski, T. Rivoal, Irrationality measures for some automatic reals numbers, *Math. Proc. Cambridge Phil. Soc.* **147** (2009), 659–678.

[10] J.-P. Allouche, Sur le développement en fraction continue de certaines séries formelles, *C.R. Acad. Sciences* **307** (1988), 631–633.

[11] J.-P. Allouche, Note sur un article de Sharif et Woodcock, *Séminaire de Théorie des Nombres de Bordeaux*, Série II**1** (1989), 163–187.

[12] J.-P. Allouche, Sur la transcendance de la série formelle Π, *J. Théor. Nombres Bordeaux* **2** (1990), 103–117.

[13] J.-P. Allouche, q-regular sequences and other generalizations of q-automatic sequences, *Lecture Notes in Computer Science* **583** (1992), 15–23.

[14] J.-P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc.* **1** (1994), 133–143.

[15] J.-P. Allouche, communication privée.

[16] J.-P. Allouche, J. Bétréma, J.Shallit, Sur des points fixes de morphismes du monoide libre, *RAIRO Inform. Theor. App.* **23** (1989), 235–249.

[17] J.-P. Allouche, D. Gouyou-Beauchamps, G. Skordev, Transcendence of binomial and Lucas' formal power series, *J. Algebra* **210** (1998), 577–592.

[18] J.-P. Allouche, J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in *Sequences and their applications (Singapore 1998)*, 1–16, *Springer*, London, 1999.

[19] J.-P. Allouche, J. Shallit, The ring of k-regular sequences II, *Theoret. Comput. Sci.* **307** (2003), 3–29.

[20] J.-P. Allouche, J. Shallit, *Automatic Sequences : Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[21] J.-P. Allouche, D. Thakur, Automata and transcendance of Tate period in finite characteristic, *Proc. Amer. Math. Soc.* **127** (1999), 1309–1312.

[22] S. V. Avgustinovich, The number of different subwords of given length in the Morse-Hedlund sequence, *Sibirsk. Zh. Issled. Oper.* **1** (1994), 3–7, 103. In Russian. English translation in A.D. Korshunov, editor, *Discrete Analysis and Operations Research*, Kluwer (1996), 1–5.

[23] D. H. Bailey, P. B. Borwein, S. Plouffe, On the Rapid Computation of Various Polylogarithmic Constants, *Math. Comp.* **66** (1997), 903–913.

[24] L. Baum, M. Sweet, Continued fractions of algebraic power series in characteristic 2, *Ann. of Math.* **103** (1976), 593-610.

[25] L. Baum, M. Sweet, Badly approximable power series in characteristic 2, *Ann. of Math.* **105** (1977), 573–580.

[26] R. M. Beals and D. S. Thakur, Computational classification of numbers and algebraic properties, *Int. Math. Res. Not.* **15** (1998), 799–818.

[27] V. Berthé, De nouvelles preuves preuves "automatiques" de transcendance pour la fonction zeta de Carlitz, *Journées Arithmétiques de Genève, Asterisque* **209** (1992), 159–168.

[28] V. Berthé, Fonction zeta de Carlitz et automates, *J. Theor. Nombres Bordeaux* **5** (1993), 53–77.

[29] V. Berthé, Automates et valeurs de transcendance du logarithme de Carlitz, *Acta Arith. LXVI* 4 (1994), 369–390.

[30] V. Berthé, Combinaisons linéaires de $\zeta(s)/\Pi^s$ sur $F_q(x)$, pour $1 \leq s \leq q-2$, *J. Number Theory* **53** (1995), 272–299.

[31] Y. Bilu, *The many faces of the subspace theorem [After Adamczewski, Bugeaud, Corvaja, Zannier...]*, Séminaire Bourbaki, Astérisque, 2008.

[32] A. W. Bluher, A. Lasjaunias, Hyperquadratic power series of degree four, *Acta Arith.* **124** (2006), 257–268.

[33] E. Borel, Sur les chiffres décimaux de $\sqrt{2}$ et divers problèmes de probabilités en chaîne, *C. R. Acad. Sci. Paris* **230** (1950), 591–593.

[34] S. Brlek, Enumeration of facteurs in the Thue-Morse word, *Discrete Appl. math.* **24** (1989), 83–96.

[35] M. W. Buck, D. P. Robbins, The continued fraction expansion of an algebraic power series satisfying a quartic equation, *J. Number Theory* **50** (1995), 335–344.

[36] Y. Bugeaud, J. H. Evertse, On two notions of complexity of algebraic numbers, *Acta Arith.* **133** (2008), 221–250.

[37] H. Cameron, D. Wood, Pm Numbers, Ambiguity, and Regularity, *RAIRO Inform. Theor. App.* **27** (1993), 261–275.

[38] L. Carlitz, On certain functions connected with polynomials in a Galois field, *Duke Math. J.* **1** (1935), 137–168.

[39] J. Cassaigne, Special factors of sequences with linear subword complexity, *Developments in language theory* (Magdeburg, 1995), 25–34, World Sci. Publishing, 1996.

[40] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc.* **4** (1997), 67–88.

[41] J. Cassaigne, Constructing infinite words of intermediate complexity. *Developments in language theory* (Kyoto 2002), volume 2450 of Lecture Notes in Comp. Sci., pages 173–184, Springer, 2002.

[42] J. W. S. Cassels, *An introduction to diophantine approximation*, Hafner Publishing Company, New York, 1972.

[43] C.-Yu Chang, J. Yu, Determination of algebraic relations among special values in positive characteristic, *Adv. Math.* **216** (2007), 321–345.

[44] H. Chérif and B. de Mathan, Irrationality measures of Carlitz zeta values in characteristic p, *J. Number Theory* **44** (1993), 260–272.

[45] G. Christol, Opération de Cartier et vecteurs de Witt, *Séminaire Delange-Pisot-Poitou* **12** (1972), 13.1–13.7.

[46] G. Christol, Ensembles presque périodiques k-reconnaissables, *Theoret. Comput. Sci.* **9** (1979), 141–145.

[47] G. Christol, T. Kamae, M. Mendès France, G. Rauzy, Suites algébriques, automates et substitutions, *Bull. Soc. Math. France* **108** (1980), 401–419.

[48] A. Cobham, On the Hartmanis-Stearns problem for a class of tag machines, *Conference Record of 1968 Ninth Annual Symposium on Switching and Automata Theory*, Schenectady, New York (1968), 51–60.

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

[50] A. Cobham, Uniform tag sequences, *Math. Systems Theory* **6** (1972), 164–192.

[51] P. Corvaja, U. Zannier, Diophantine equations with power sums and universal Hilbert sets, *Indag. Math. (N.S.)* **9** (1998), 317–332.

[52] P. Corvaja, U. Zannier, A Subspace Theorem approach to integral points on curves, *C. R. Acad. Sci. Paris Ser. I* **334** (2002), 267–271.

[53] P. Corvaja, U. Zannier, Some new applications of the subspace theorem, *Compositio Math.* **131** (2002), 319–340.

[54] P. Deligne, Intégration sur un cycle évanescent, *Inventiones Math.* **76** (1984), 129–143.

[55] J. Denef, L. Lipshitz, Algebraic power series and diagonals, *J. Number Theory* **26** (1987), 46–67.

[56] H. Derksen, A Skolem-Mahler-Lech theorem in positive characteristic and finite automata, *Invent. Math.* **168** (2007), 175–224.

[57] H. Derksen, D. Masser, Linear equations over multiplicative groups, recurrences, and mixingI, preprint.

[58] L. G. P. Dirichlet, Verallgemeinerung eines Satzes aus der Lehre von den Kettenbr0chen nebst einigen Anwendungen auf die Theorie der Zahlen, *S.-B.Preuss. Akad. Wiss.* (1842), 93–95.

[59] D. Duverney, Sommes de deux carrés et irrationalité de valeurs de fonctions thêta, *C. R. Math. Acad. Sci. Paris* **320** (1995), 1041–1044.

[60] F. J. Dyson, The approximation to algebraic numbers by rationals, *Acta Math.* **79** (1947), 225–240.

[61] A. Ehrenfeucht, K. P. Lee et G. Rozenberg, Subword complexities of various classes of deterministic developmental languages without interaction, *Theoret. Comput. Science* **1** (1975), 59–75.

[62] S. Eilenberg, *Automata, languages and machines*, vol. A, Academic Press, New York, 1974.

[63] P. Flajolet, Analytic models and ambiguity of context-free languages, *Theoret. Comput. Sci.* **49** (1987), 283–309.

[64] S. Ferenczi, C. Mauduit, Transcendence of numbers with a low complexity expansion, *J. Number Theory* **67** (1997), 146–161.

[65] A. Fraenkel, Systems of numeration, *Amer. Math. Monthly* **92** (1985), 105–114.

[66] H. Furstenberg, Algebraic functions over finite fields, *J. Algebra* **7** (1967), 271–277.

[67] I. Gheorghiciuc, The subword complexity of finite and infinite binary words, *Adv. in Appl. Math.* **39** (2007), 237–259.

[68] B. P. Gill, An analogue for Algebraic Functions of the Thue-Siegel Theorem, *Ann. of Math.* **31**, 207–218, 1930.

[69] D. Goss, *Basic Structures of Function Field Arithmetic*, Springer-Verlag, Berlin, 1996.

[70] J. Hadamard, Théorème sur les séries entières, *Acta Math.* **22** (1899), 55–63.

[71] H. Hahn, Uber die nichtarchimedische Grobensysteme, in *Gesammelte Abhandlungen* **I**, Springer, Vienna, 1995.

[72] T. Harase, Algebraic elements in formal power series rings, *Israel J. Math.* **63** (1988), 281–288.

[73] G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford Sci. Publ., Oxford Univ. Press, 1989.

[74] I. Kaplanasky, Maximal fields with valuations, *Duke Math. J.* **9** (1942), 303–321.

[75] K. Kedlaya, Finite automata and algebraic extensions of function fields, *J. Théor. Nombres Bordeaux* **18** (2006), 379–420.

[76] M. Kontsevich, D. Zagier, Periods. In : *Mathematics unlimited-2001 and beyond*, pp. 771–808, Springer-Verlag, 2001.

[77] P. Kůrka, *Topological and symbolic dynamics*, volume 11 of *Cours spécialisés* SMF, 2003.

[78] A. Lasjaunias, Diophantine approximation and continued fractions for algebraic power series in positive characteristic, *J. Number Theory* **65** (1997), 206–225.

[79] A. Lasjaunias, A Survey of Diophantine approximation in Fields of Power Series, *Monatsh. Math.* **130** (2000), 211-229.

[80] A. Lasjaunias, Diophantine approximation in positive characteristic, *Proc. of a Conference on Analytic Number Theory and Surronding Areas* RIMS Kokyuroku **1511** (2004), 156–165.

[81] A. Lasjaunias, Algebraic Continued Fractions in $\mathbb{F}_q((T^{-}1))$ and reccurent sequences in $\mathbb{F}_q$, *Acta Arith.* **133** (2008), 251–265.

[82] A. Lasjaunias, Continued fractions for hyperquadratic power series over a finite field, *Finite Fields Appl.* **14** (2008), 329–350.

[83] A. Lasjaunias, On Robbin's example of a continued fraction for a quartic power series over $\mathbb{F}_{13}$, *J. Number Theory* **128** (2008), 1109–1115.

[84] A. Lasjaunias, B. de Mathan, Thue's theorem in positive characteristic, *J. Reine Angew. Math* **473** (1996), 195–206.

[85] A. Lasjaunias, B. de Mathan, Differential equations and diophantine approximation in positive characteristic, *Monatsh. Math.* **128** (1999), 1–6.

[86] J. Liouville, Nouvelle démonstration d'un théoréme sur les irrationelles algébriques, *C. R. Acad. Sci. Paris* **18** (1844), 910–911.

[87] M. Lothaire, *Combinatorics on words*, Cambridge University Press, Cambridge, 1997.

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

[89] M. Lothaire, *Applied Combinatorics on words*, Encyclopedia of Mathematics and its Applications 105, Cambridge University Press, Cambridge, 2005.

[90] A. de Luca, S. Varrichio, Some combinatorial properties of the Thue-Morse sequence and a problem in semi-groups, *Theoret. Comput. Sci.* **63** (1989), 333–348.

[91] K. Mahler, On a theorem of Liouville in fields of positive characteristic, *Canad. J. Math* **1** (1949), 397–400.

[92] E. Maillet, *Introduction à la théorie des nombres transcendants et des propriétés arithmétiques des fonctions*, Gauthier-Villars, Paris, 1906.

[93] C. Mauduit, Multiplicative properties of the Thue-Morse sequence, *Period. Math. Hung.* **43**, 137–153, 2001.

[94] M. Mkaouar, Sur le développement en fraction continue de la série de Baum et Sweet, *Bull. Soc. Math. France* **123** (1995), 361–374.

[95] W. Mills, D. Robbins, Continued fractions for certain algebraic power series, *J. Number Theory* **23** (1986), 388–404.

[96] M. Morse, G. A. Hedlund, Symbolic dynamics, *Amer. J. Math.* **60** (1938), 815–866.

[97] M. Morse, G. A. Hedlund, Symbolic dynamics II, *Amer. J. Math.* **62** (1940), 1–42.

[98] B. Mossé, Reconnaissabilité des substitutions et complexité des suites automatiques, *Bull. Soc. Math. France* **124** (1996), 329–346.

[99] C. Osgood, An effective lower bound on the "Diophantine approximation" of algebraic functions by rational functions, *Mathematika* **20** (1973), 4–15.

[100] C. Osgood, Effectives bounds on the "Diophantine approximation" of algebraic functions over fields of arbitrary characteristic and applications to differential equations, *Nederl. Akad. Wetensch. Proc. Ser. A* **78** (1975) 105–119.

[101] J. J. Pansiot, Subword Complexities and Iteration, *Bulletin of EATCS* **26** (1985), 55–62.

[102] M. A. Papanikolas. Tannakian duality for Anderson-Drinfeld motives and algebraic independence of Carlitz logarithms, *Invent. Math.* **171** (2008), 123–174.

[103] F. Pellarin, Aspects de l'indépendance algébrique en caractéristique non nulle (d'après Anderson, Brownawell, Denis, Papanikolas, Thakur, Yu, et al.), Séminaire Bourbaki 2006/2007, exposé no. 973, *Astérisque* **317** (2008), 205–242.

[104] N. Pytheas Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Math., 1974, Springer-Verlag, Berlin, 2002.

[105] M. Queffélec, *Substitution Dynamical Systems. Spectral Analysis*, Lecture Notes in Math. **1294**, Springer-Verlag, Berlin, 1987.

[106] K. F. Roth, Rational approximations to algebraic numbers, *Mathematika* **2** (1955), 1–20 ; corrigendum, 169.

[107] O. Salon, Suites automatiques à multi-indices, *Sém. Théor. Nombres Bordeaux (1)* exposé numéro 4 (1986-1987), 4.01-4.36 (avec un appendice de J. Shallit).

[108] O. Salon, Suites automatiques à multi-indices et algébricité, *C. R. Acad. Sci. Paris*, Série I, Math. **305** (1987), 501–504.

[109] W. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics **785**, Springer, Berlin, 1980.

[110] W. Schmidt, On continued fractions and Diophantine approximation in power series fields, *Acta Arith.* **95** (2000), 139–166.

[111] J. Shallit, Simple continued fractions for some irrational numbers, *J. Number Theory* **11** (1979), 209–217.

[112] H. Sharif, C. F. Woodcock, Algebraic functions over a field of positive characteristic and Hadamard products, *J. Lond. Math. Soc.* **37** (1988), 395–403.

[113] H. Sharif and C. F. Woodcock, On the transcendence of certain series, *J. Algebra* **121** (1989), 364–369.

[114] C. Siegel, Approximation algebraischer Zahlen, *Math. Z.* **10** (1921), 173–213.

[115] N. Sloane, The On-Line Encyclopedia of Integer Sequences, Published electronically at `http://www.research.att.com/~njas/sequences`.

[116] R. P. Stanley, Differentiably finite power series, *European J. Combin.* **1** (1980), 175–188.

[117] T. Tapsoba, Automates calculant la complexité de suites automatiques, *J. Théor. Nombres Bordeaux* **6** (1994), 127–134.

[118] D. Thakur, Continued fraction for the exponential for $\mathbb{F}_q[T]$, *J. Number Theory* **41** (1992), 150–155.

[119] D. Thakur, Automata-style proof of Voloch's result on transcendence, *J. Number Theory* **58** (1996), 60–63.

[120] D. Thakur, Exponential and continued fractions *J. Number Theory* **59** (1996), 248–261.

[121] D. Thakur, Transcendence of Gamma Values for $\mathbb{F}_q(T)$, *Ann. Math.* **144** (1996), 181–188.

[122] D. Thakur, Patterns of continued fractions for the analogues of e and related numbers in function field case *J. Number Theory* **66** (1997), 129–147.

[123] D. Thakur, Diophantine approximation exponents and continued fractions for algebraic power series, *J. Number Theory* **79** (1999), 284–291.

[124] D. Thakur, *Diophantine approximation in finite characteristic* Algebra, Arithmetic and Geometry with Applications, Ed. C. Christensen et al, Springer 2003, 757–765.

[125] D. Thakur, *Function Field Arithmetic*, World Scientfic, Singapore, 2004.

[126] D. Thakur, Approximation Exponents for function fields, In *Analytic Number Theory* –Essays in Honor of Klaus Roth, Ed. by W. Chen, T. Gowers, H. Halbertam, W. Schmidt, R. Vaughn, Camb. U. Press (2009), 421-435.

[127] A. Thue, Uber unendliche Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **7**, 1–22, 1906. Reprinted in *Selected Mathematical Papers of Axel Thue* (Ed. T. Nagell). Oslo : Universitetsforlaget, 139–158, 1977.

[128] A. Thue, Uber Annaherungswerte algebraischer Zahlen, *J. Reine Angew. Math.* **135** (1909), 284–305.

[129] A. Thue, Uber die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Norske vid. Selsk. Skr. Mat. Nat. Kl.* **1**, 1–67, 1912. Reprinted in *Selected Mathematical Papers of Axel Thue* (Ed. T. Nagell). Oslo : Universitetsforlaget, 413–478, 1977.

[130] A. M. Turing, On computable numbers, with an application to the Entscheidungsproblem, *Proc. London Math. Soc.* **2** 42 (1937), 230–265, (Corrigendum 43 (1937), 544–546).

[131] S. Uchiyama, On the Thue-Siegel-Roth theorem. III., *Proc. Japan Acad.* **36** (1960), 1–2.

[132] J. Y. Yao, Critères de non-automaticité, et leurs applications, *Acta Arith.* **83** (1997), 237–248.

[133] J. Yu, Transcendence and Special Zeta Values in Characteristic p, *Ann. of Math.* **134** (1991), 1–23.

[134] J. F. Voloch, Diophantine approximation in positive characteristic, *Period. Math. Hungar.* **19** (1988), 217–225.

[135] L. I. Wade, Certain quantities transcendental over $GF(p^n, x)$, *Duke Math. J.* **8** (1941), 701–720.

[136] L. I. Wade, Two types of function field trasncendental numbers, *Duke Math. J.* **11** (1944), 755–758.

[137] L. I. Wade, Transcendence properties of Carlitz Ψ-functions, *Duke Math. J.* **13** (1946), 79–85.

[138] M. Waldschmidt, Un demi-siècle de transcendance, in *Development of Mathematics 1950–2000*, Birkhäuser, Basel, (2000), 1121–1186.