Université Pierre et Marie Curie

École doctorale de sciences mathématiques de Paris centre

Thèse de doctorat

Discipline : Mathématiques

présentée par

Shuo LI

Des résultats autour des suites automatiques

dirigée par Jean-Paul ALLOUCHE

Soutenue le 22 Janvier 2020 devant le jury composé de :

M. Jean-Paul Allouche	Institut de mathématiques de Jussieu-PRG	directeur
Mme Elise Janvresse	Université de Picardie Jules Verne	examinatrice
Mme Sophie Morier-Genoud	Institut de mathématiques de Jussieu-PRG	examinatrice
M. Michel Rigo	Université de Liège	rapporteur

Institut de mathématiques de Jussieu-
Paris Rive gauche. UMR 7586.
Boîte courrier 247
4 place Jussieu
75 252 Paris Cedex 05

Université Pierre et Marie Curie.
École doctorale de sciences
mathématiques de Paris centre.
Boîte courrier 290
4 place Jussieu
75 252 Paris Cedex 05

Remerciements

Premièrement je voudrais exprimer mes sincères remerciements à mon directeur, Jean-Paul Allouche, pour son encadrement, ses patiences et ses conseils pendant les quatre dernières années. Sans lui, je n'aurais pas pu découvrir le monde des automates dans lequel j'ai trouvé mon premier intérêt de recherche. Grâce à sa personnalité libérale et ses supports altruistes, j'ai pu passer mon temps "capricieusement" sur des questions "peu accessibles".

Ensuite, j'aimerais remercier Srecko Brlek de m'accueillir à Montréal pendant un trimètre de 2018. Il m'a ouvert une autre porte vers la combinatoire des mots. J'ai bien profité le temps à Montréal et j'ai pu beaucoup parler avec les gens qui travaillent dans les autres branches de combinatoire.

Un grand merci à Michel Rigo et à Pelantová Edita d'avoir accepté de rapporter cette thèse, rédigé des rapports et donné des conseils de modification. Et je tiens à exprimer ma gratitude à Elise Janvresse et à Sophie Morier-Genoud qui complètent le jury de thèse.

Finalement, j'aimerais remercier mes familles, le groupe des amis et tous les gens qui m'ont apporté des idées en mathématiques. Sans les supports de leur côté, je ne pourrais pas finaliser ces travaux.

Résumé

Cette thèse est consacrée à l'étude des propriétés des suites automatiques. Cette dernière est une notion premièrement introduite et étudiée par les mathématiciens et les informaticiens théoriciens en combinatoire, notamment en théories des langages, mais elle a aussi des applications intéressantes en théorie des nombres. Dans cette thèse, on travaille sur quatre sujets concernant les aspects mathématiques et informatiques, liés aux suites automatiques.

Dans le Chapitre 1, on donne une introduction aux suites automatiques ainsi que des résultats récents autour de ce sujet. Dans le Chapitre 2, on étudie le prolongement méromorphe des séries de Dirichlet du type

$$f(s) = \sum_{x \in \mathbf{N}^k} \frac{a_x}{p(x)^s}$$

sur $\mathbf{C}$, avec $(a_x)_{n \in \mathbf{N}^k}$ une suite automatique de dimension k, et p un polynôme elliptique qui ne s'annule pas sur $\mathbf{N}^k$. Et aussi, des produits infinis sont calculés comme conséquences de ce résultat. Dans le Chapitre 3, on trouve une expression explicite pour toute suite automatique complètement multiplicative. Dans le Chapitre 4, on considère les séries formelles définies par des produits infinis du type,

$$\sum_{n=1}^{\infty} a_n x^n = \prod_{i=1}^{\infty} p(x^{q^i})$$

avec p un polynôme à coefficients dans $\mathbf{Q}$ et $p(0) = 1$, et q un entier plus grand que 1. On démontre que pour q et d fixés, il n'y a qu'un nombre fini des polynômes de dégrée au plus d tel que la série infinie obtenue par la définition précédente soit q-automatique. Dans le Chapitre 5, on étudie la longueur palindromique des suites automatiques, et on trouve toute les suites ayant la même longueur palindromique que celle de Thue-Morse.

Mots-clés

Suites automatiques, Fonctions complètement multiplicatives, Fonction Mahlérienne, Longueur palindromique.

Abstract

In this thesis we are interested in automatic sequences, which is a notion introduced and initially studied by combinatorists, and by people working on language theory. Meanwhile these sequences also appear to have some interesting properties in number theory. In this thesis, we deal with some topics in mathematics and computer science, related to automatic sequences.

In Chapter 1, we give an introduction to automatic sequences and a brief overview of recent works on related topics. In Chapter 2, we study the meromorphic continuation of Dirichlet series of type

$$f(s) = \sum_{x \in \mathbf{N}^k} \frac{a_x}{p(x)^s}$$

over $\mathbf{C}$, where $(a_x)_{n \in \mathbf{N}^k}$ is an automatic sequence of dimension k, and p is an elliptic polynomial non-vanishing over $\mathbf{N}^k$. And some infinite products are calculated as consequences of this result. In Chapter 3, we give a formal expression to all completely multiplicative automatic sequences. In Chapter 4, we study formal powers series defined by infinite products of type

$$\sum_{n=1}^{\infty} a_n x^n = \prod_{i=1}^{\infty} p(x^{q^i})$$

where p is a polynomial with coefficients over $\mathbf{Q}$ and $p(0) = 1$, and q is an integer larger than 1. We prove that for given integer d, there are finitely many polynomials of degree d such that the sequence defined as above is automatic. In Chapter 5, we study the palindromic length of automatic sequences and find all sequences having the same palindromic length as the one of Thue-Morse.

Keywords

Automatic sequences, completely multiplicative functions, Mahler functions, Palindromic length.

Chapter 0

This thesis consists of 5 chapters following this one. The last four chapter correspondent to four articles that are either on the ArXiv or submitted (or both).

Contents

Chapter 1

Introduction to automatic sequences

Introduction to automatic sequences

Sequence is a fundamental notion in mathematics. To describe a sequence, apart from listing all elements one by one, we may want to find some relations between its elements. Periodic sequences are good examples in this case, as they are highly ordered and easy to describe. While there are other sequences which are more complicated but still preserve relations between elements, they may be called as "programmable sequences". In this thesis we study a specific type of programmable sequences, called automatic sequences.

Automatic sequences can be defined by a finite automaton, which is a notion from language theory. However, there are some equivalent definitions from other branches of mathematics such as combinatorics as well as number theory. In the first section, we will present the related definitions of automatic sequences and applications in each field, and show how automaticity can be used to build a link between two different fields.

At the end of each sub-section of this chapter we introduce some new results obtained by the author. All these results are based, directly or indirectly, but strongly, on the mechanism of substitutions (of words, functions or matrices..). It is this mechanism that gives automatic sequences a complex but computational structure, which allows us to find formal expressions for some "disordered" phenomena.

1.1 First definition and applications in number theory

In this section we will introduce one of the definitions of automatic sequences and show why this kind of sequences are interesting in number theory.

Definition A sequence $(a_n)_{n\in\mathbf{N}}$ is called k-automatic if and only if the set

$$Ker((a_n)_{n\in\mathbf{N}}) = \left\{(a_{k^l n+i})_{n\in\mathbf{N}} | l \in \mathbf{N}, 0 \leq i < k^l\right\}$$

is finite. This set will be called the k-kernel of the sequence $(a_n)_{n\in\mathbf{N}}$.

Example Let us consider a sequence defined in such way: we begin with a word of one letter 1, next step we change the letter 1 to the string 10, in the third step we change 1 to 10 and 0 to 01... At each step we change 1 to 10 and 0 to 01 letter by letter so that we obtain an infinite sequence. The process is presented as follows:

1
10
1001
10010110
1001011001101001...

At the end we get a sequence like

10010110011010010110...

This sequence is called the Thue-Morse sequence, and it is easy to show that this sequence is 2-automatic. In fact, we have the following relations:

$$a_{2n} = a_n$$

$$a_{2n+1} = 1 - a_n$$

for all $n \geq 0$. If we let $(a_n)_{n\in\mathbf{N}}$ denote the Thue-Morse sequence, the kernel of this sequence is $Ker((a_n)_{n\in\mathbf{N}}) = \{(a_{2n+1})_{n\in\mathbf{N}}, (a_n)_{n\in\mathbf{N}}\}$.

Example Another example is the Rudin-Shapiro sequence, which can be defined by the number of the factors 11 in the binary expansion of n: the sequence $(g(n))_{n\in\mathbf{N}}$ satisfies the relation

$$g(n) = (-1)^{a_{2,11}(n)}$$

for each n, where $a_{2,11}(n)$ counts the number of 11 in the binary expansion of n. The first terms of the Rudin-Shapiro sequence are

$$+1, +1, +1, -1, +1, +1, -1, +1, +1, +1, +1, -1, -1, -1, +1, -1, ...$$

Because of the definition, $(g(n))_{n\in\mathbf{N}}$ satisfies the following recurrence:

$$g(2n+1) = (-1)^n g(n)$$

$$g(2n) = g(n).$$

So the 2-kernel of this sequence is

$$Ker(g(n)_{n\in\mathbf{N}}) = \left\{-g(n)_{n\in\mathbf{N}}, g(n)_{n\in\mathbf{N}}, (-1)^n g(n)_{n\in\mathbf{N}}, (-1)^{n+1} g(n)_{n\in\mathbf{N}}\right\}.$$

1.1.1 Infinite products and Dirichlet series

Thanks to the finiteness of the k-kernel, there are several infinite sums or products associated with a k-automatic sequence presenting interesting properties. One of the most famous results about automatic sequences is the Woods–Robbins product identity [44] [37], which is written as:

$$\left(\frac{1}{2}\right)^{(-1)^{a(0)}} \left(\frac{3}{4}\right)^{(-1)^{a(1)}} \left(\frac{5}{6}\right)^{(-1)^{a(3)}} ... = 2^{-\frac{1}{2}}$$

where $(a(m))_{m\in\mathbf{N}}$ is the Thue-Morse sequence. This result can be obtained and generalized by evaluating a function of two variables defined by $f(a,b) = \prod_{n\geq 1}(\frac{n+a}{n+b})^{(-1)^{a_n}}$ at some rational points, and using the finiteness of $Ker((a_n))_{n\in\mathbf{N}}$.

Theorem 1.1. *([3]) Let us define $g(x) = \frac{f(\frac{x}{2},\frac{x+1}{2})}{x+1}$ then we have*

$$(1+x)g(x) = \frac{g(\frac{x}{2})}{g(\frac{x+1}{2})}, x \neq -1, -2, -3...$$

Corollary 1.1.

$$\prod_{n\geq 0}\left(\frac{4n+1}{4n+3}\right)^{(-1)^{a_n}} = \frac{1}{2}$$

$$\prod_{n\geq 1}\left(\frac{(2n-1)(4n+1)}{(2n+1)(4n-1)}\right)^{(-1)^{a_n}} = 2$$

For more results see [3]. However, the W-R identity can also be obtained by considering the associated Dirichlet series:

$$f(s) = \sum_{n\geq 1}\frac{a_n}{(n+1)^s}$$

and calculating the derivative at 0. Before doing this, let us introduce an important theorem on general automatic Dirichlet series:

Theorem 1.2. *([2]) Let $(a_n)_{n\geq 0}$ be an automatic sequence with values in $\mathbb{C}$. Then the Dirichlet series*

$$\sum_{n=0}^{\infty}\frac{a_n}{(n+1)^s} \qquad \sum_{n=1}^{\infty}\frac{a_n}{(n)^s}$$

have meromorphic continuations to the whole complex plane, whose poles (if any) are located on a finite number of left semi-lattices.

As a consequence, if we can calculate the derivative of those functions on some points over their convergence domain, then we can build some equalities between an infinite product and a real number.

Corollary 1.2. *Let q, r be two integers larger then 1, and let $(s_q(m))_{m\in\mathbf{N}}$ and $(u_j(m))_{m\in\mathbf{N}}$ be sequences defined respectively as*

$$\begin{cases} s_q(0) = 0 \\ s_q(qm+a) = s_q(m) + a, 0 \leq a \leq q-1; \end{cases}$$

$$u_j(m) = \begin{cases} r-1 & \text{if } s_q(m) \equiv j \mod r \\ -1 & \text{otherwise.} \end{cases}$$

Then we have the equality

$$\prod_{m=0}^{\infty} \left(\frac{m+1}{q[m/q]+1} \right)^{u_j(m)} = q^{j+(1-r)/2},$$

where $[a]$ represents the largest integer smaller than a. In particular, taking $q = r = 2$ and $j = 0$, we have the W-R identity.

The Chapter 2 of this thesis aims to generalize this result to multi-index automatic sequences, a notion generalized from Definition 1 and studied in [40].

Definition We let $\underline{x}$ denote an n-tuple $(x_1, x_2...x_n)$ and let $d \geq 2$ be an integer. A sequence $(a_{\underline{n}})_{\underline{n}\geq\underline{0}}$ with values in the set $\mathcal{A}$ is called $d-$automatic if and only if its $d-$kernel $\mathcal{N}_d(a)$ is finite, where the $d-$kernel of the sequence$(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is the set of subsequences defined by

$$\mathcal{N}_d(\underline{a}) = \left\{ (m_1, m_2...m_n) \longmapsto a_{(d^k m_1 + l_1, d^k m_2 + l_2, \ldots, d^k m_n + l_n)}; k \geq 0, \underline{(0)} \leq \underline{l} \leq \underline{(d^k - 1)} \right\}.$$

Remarking that constant sequences are a kind of particular automatic sequences, the Dirichlet sequences in the form $\sum_{(n_1,n_2...n_I)\in\mathbf{N}_+^I} \frac{n_1^{\mu_1} n_2^{\mu_2} ... n_I^{\mu_I}}{p(n_1,n_2,..n_I)^s}$ have been largely studied, where $(\mu_1, \mu_2, ...\mu_I) \in \mathbf{N}_+^I$ and p is a I-variable function. R.H. Mellin [34] firstly proved in 1900 that the functions above have a meromorphic continuation to the whole complex plane when $\mu_i = 0$ for all indexes i, then K. Mahler [32] generalized the result to the case that μ_i are arbitrary positive integers when the polynomial satisfies the elliptic condition in 1927. In 1987, P. Sargos [41] proved that the condition "$\lim |p(n_1, n_2...n_I)| \to \infty$ when $|(n_1, n_2...n_I)| \to \infty$ and p is non-degenerate" is sufficient for these Dirichlet sequences to have a meromorphic continuation. In 1997, D. Essouabri [20] generalized the condition to the case, $\frac{\partial_{\underline{\mu}} p(\underline{x})}{p(\underline{x})} = \mathcal{O}(1), \forall \underline{\mu} \in \mathbf{Z}_+^n$. In Chapter 2 the author finds a sufficient condition for the sequences $\sum_{\underline{x}\in\mathbf{N}^n} \frac{a_{\underline{x}} \prod_{i=1}^n x_i^{\mu_i}}{P(\underline{x})^s}$ to have a meromorphic continuation over $\mathbf{C}$. The main result is announced as follows:

Theorem 1.3. *Let p be an elliptic polynomial of n variables and $(a_{\underline{x}})_{\underline{x}>\underline{0}}$ be q-automatic, then for a given n-tuple $\underline{\mu}$, the function $\sum_{(\underline{x})\in\mathbf{N}_+^n} \frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}$ admits an abscissa of convergence σ such that it converges absolutely on the half-plane $\Re(s) > \sigma$ and has a meromorphic continuation on the whole complex plane. Furthermore, the poles of this function (if any) are located on a finite number of left semi-lattices.*

1.1.2 Automatic power series

Another application of the finiteness of the k-kernel is to build links between automatic power series and functional equations.

Definition Let $\sum_{n\geq 0} a_n x^n$ be a formal power series defined over a field $\mathbf{F}$. We say it is automatic if the sequence of its coefficients is automatic.

It is proved that all automatic power sequences over an arbitrary field satisfy a functional equation.

Theorem 1.4. *([4]) let $f \in F((x))$ be a k-automatic power series, then there exist polynomials $a_0(x), a_1(x), ..., a_m(x) \in F[x]$ with $a_0(x)a_m(x) \neq 0$ such that*

$$\sum_{i=0}^{m} a_i(x) f(x^{k^i}) = 0.$$

We may ask the reversal question: given a functional equation, what does the solution look like? This question has been widely studied in [12] [17] [11] [16]. In [12] G. Christol answered this question for power sequences defined over a finite field: he proved that there is an equivalence between being a solution of such a functional equation and being an automatic power series.

Definition Let $\mathbf{F}$ be any field. We say that a formal power series $f(x) \in \mathbf{F}((x))$ is algebraic if it is algebraic over $\mathbf{F}(x)$, that is, if there exist $d \in \mathbf{N}$ and polynomials $a_i(x) \in \mathbf{F}[x]$ for $i = 0, ..., d$, such that $a_d(x) \neq 0$ and

$$\sum_{i=0}^{d} a_i(x) f(x)^i = 0.$$

Theorem 1.5. *([12]) Let $\mathbf{F}_q$ be a finite field of characteristic p. A formal power series $f(x) = \sum_{n\geq 0} a_n x^n \in F((x))$ is algebraic over the rational function field $\mathbf{F}_q(x)$ if and only if the sequence of its coefficients is q-automatic.*

However, things are going to be complicated for an arbitrary field, namely for $\mathbf{R}$ or $\mathbf{C}$, even though the simplest equation does not lead to general conclusions:

$$f(x) = p(x) f(x^k),$$

where p is a polynomial. It is known that coefficients of solutions of above functional equation satisfy a more complicated structure, but it is still unknown if there are some restrictions on p to make the solutions to be automatic. P. Dumas studied asymptotic properties of coefficients of solutions of above equation [17], S. Checcoli and J. Roques found all polynomials of low degree such that the power series f defined by the above functional equations are automatic [11].

Definition Let $(a(n))_{n\geq 0}$ be a sequence defined over a field $\mathbf{F}$, it is called k-regular if the $\mathbf{F}$-module generated by

$$Ker((a_n)_{n\in\mathbf{N}}) = \left\{(a_{k^l n+i})_{n\in\mathbf{N}} | l\in\mathbf{N}, 0\leq i<k^l\right\}$$

is a finitely-generated $\mathbf{F}$-module. A formal power series $f(x)\in\mathbf{F}((x))$ is called regular if the sequence of its coefficients is regular.

Theorem 1.6. *([17]) Let k be an integer larger then 1 and $p(x)\in\mathbf{F}[x]$, then all power sequences $f(x)=\sum_{n\geq 0} a_n x^n \in F((x))$ satisfying equation*

$$f(x) = p(x) f(x^k)$$

are k-regular.

In Chapter 3 the author finds some properties of automatic power sequence $f(x)$ defined over $\mathbf{Q}$ satisfying the equation above. Main results are as follows:

Theorem 1.7. *For given integers $q\geq 2$ and $d\geq 0$, there exist finitely many polynomials of degree d defined over the field of rational numbers $\mathbf{Q}$, say $p_1(x), p_2(x), ...p_k(x)$, such that $\prod_{s=0}^{\infty} p_j(x^{q^s}) = \sum_{i=1}^{\infty} c_i x^i$ is a q-automatic power series for $j: 1\leq j\leq k$.*

Theorem 1.8. *If the power series $F(x)=\prod_{s=0}^{\infty}\frac{1}{f(x^{q^s})}=\sum_{i=0}^{\infty} c_i x^i$ is q-regular, then there exists a polynomial $Q(x)$ such that $f(x)|\frac{Q(x^q)}{Q(x)}$, Furthermore $F(x)$ can be written as*

$$F(x) = Q(x)\prod_{i=1}^{\infty} R(x^q),$$

where $R(x)=\frac{Q(x^q)}{Q(x)f(x)}$ is a polynomial.

1.2 Automatic sequence and language theory

In this section, we will give a definition of automatic sequence using the notion of automaton. To do so, let us give firstly a definition of a finite automaton:

Definition Let a (finite) alphabet $\sum$ to be a (finite) set of elements, we say a word a over $\sum$ if a is a string of elements in $\sum$, in particuler the empty word is a word over all $\sum$, let us denote by $\sum^*$ the set of words over $\sum$. We say that L is a language over $\sum$ if L is a subset of $\sum^*$.

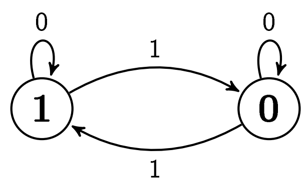
Definition A deterministic finite automaton, or DFA, $\mathcal{M}$ is defined to be a 5-tuple $\mathcal{M} = (\mathcal{Q}, \sum, \delta, q_0, \mathcal{F})$
where
$\mathcal{Q}$ is a finite set of states,
$\sum$ is the finite input alphabet,
$\delta : \mathcal{Q} \times \sum \to \mathcal{Q}$ is the transition function,
$q_0 \in \mathcal{Q}$ is the initial state, and
$\mathcal{F} \subseteq \mathcal{Q}$ is the set of accepting states.

A k-automatic sequence can be defined as an output of a k-deterministic finite automaton with output, which is denoted by k-DFAO. Comparing with a DFA, a DFAO is defined to be a 6-tuple $\mathcal{M} = (\mathcal{Q}, \sum_k, \delta, q_0, \Delta, \tau)$. Where $\mathcal{Q}, \sum, \delta, q_0$ are defined as they were in the previous definition, $\sum_k = \{0, 1, 2, ..., k-1\}$, Δ is the output alphabet, and $\tau : \mathcal{Q} \to \Delta$ is the output function. Such a machine $\mathcal{M}$ defines a function from $\sum^*$ to Δ, which we denote as $f\mathcal{M}(w)$, as follows: $f\mathcal{M}(\omega) = \tau(\delta(q_0, \omega))$.

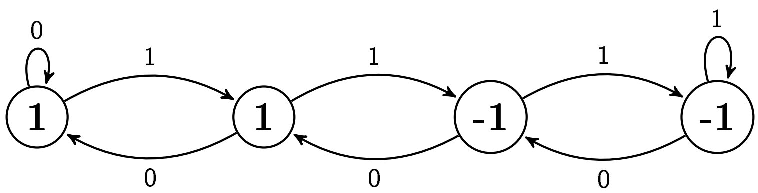
Definition We say the sequence $(a_n)_{n\in\mathbf{N}}$ over a finite alphabet Δ is k-automatic if there exists a k-DFAO $\mathcal{M} = (\mathcal{Q}, \sum_k, \delta, q_0, \Delta, \tau)$ such that $a_n = \tau(\delta(q_0, \omega))$ for all $n \geq 0$ and all w with $[w]_k = n$, where $[w]_k$ represents the number whose expansion in base k is w.

It is easy to see that there is a natural equivalence between this definition and the one presented in Section 1. Here we redefine the two examples in the previous section.

Example The DFAO of the Thue-Morse sequence is as follows:



The DFAO of the Rudin-Shapiro sequence is as follows:



Remarking that in language theory, there is another notion related to automaton, called regular languages.

Definition The set of regular languages over an alphabet $\sum$ is defined recursively as follows:
a) The empty language and the set of empty word are regular languages.
b) For each element $a \in \sum$, the language $\{a\}$ is a regular language.
c) If A and B are regular languages, then the union, the concatenation and the free monoid generated by one of them are regular languages.
d) No other languages over $\sum$ are regular.

Theorem 1.9. *(see,e.g.,[5]) A language is accepted by a DFA if and only if it is a regular language.*

Theorem 1.10. *(see,e.g.,[5]) Let $\mathcal{M} = (\mathcal{Q}, \sum_k, \delta, q_0, \Delta, \tau)$ be a k-DFAO,then for all $d \in \Delta$ the set $Id(\mathcal{M}) = \{w \in \sum^* : \tau(\delta(q_0, \omega)) = d\}$ is a regular language.*

As we can see, there is a strong relation between automatic sequences and regular languages, we may expect to use combinatorial properties of regular languages to prove some asymptotic properties of automatic sequences. Let us firstly recall some notion of densities used to study asymptotic properties of sequences:

Definition Let E be a subset of the integers. We say that the set E has an asymptotic density or a natural density d if the limit

$$d = \lim_{x \to \infty} \frac{1}{x} \sharp \{n \leq x; n \in E\}$$

exists.
We say that the set E has a logarithmic density δ if the limit

$$\delta = \lim_{x \to \infty} \frac{1}{\log x} \sum_{n \leq x, n \in E} \frac{1}{n}$$

exists.
We say that the set E has an analytic density δ' if the limit

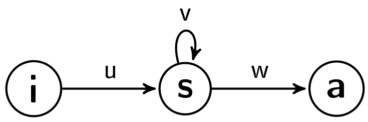
$$\delta' = \lim_{s \to 1+} (s-1) \sum_{n \geq 1, n \in E} \frac{1}{n^s}$$

exists. We call these limits respectively asymptotic density, logarithmic density and analytic density.

Theorem 1.11. *([2]) The existence of asymptotic density implies the existence of logarithmic density, and their values coincide;*
The existence of logarithmic density is equivalent to the existence of analytic density, and their values coincide.

Theorem 1.12. *(see,e.g.,[5]) Let $L \subseteq \sum^*$ be a regular language. Then there exists a constant $n \geq 1$ such that for all words $z \in L$ with length larger then n, there exists a decomposition $z = uvw$, where $u, v, w \in \sum^*$ and v is non-empty such that $uv^iw \in L$ for all $i \geq 0$. Furthermore, the constant n can be taken to be the number of states in the minimal DFA for L.*

This theorem is called pumping lemma, what it is saying is that if a word is long enough, then there must be some loops on the chain from the initial state to some acceptable state, so that all words obtained by repeating finitely many times one of the loops are also acceptable.



If we apply the same idea to study the asymptotic property of an automatic sequence then we have the following property

Theorem 1.13. *([42]) Let $(a_n)_{n\in\mathbf{N}}$be a k-automatic sequence and q be the number of states of a direct automaton generating $(a_n)_{n\in\mathbf{N}}$ then for any $m, y \in \mathbf{N}$. We have equality between the sets $\{a_n | mk^{q!} \leq n < (m+1)k^{q!}\} = \{a_n | mk^{yq!} \leq n < (m+1)k^{yq!}\}$.*

As a consequence,

Corollary 1.3. *Let $(a_n)_{n\in\mathbf{N}}$ be an automatic sequence, then there exists a integer k such that for all couple of integers r, s such that $r > k, s > 0$, if there is some a_n not in the set $\{a_n | s \leq n < s + r\}$, then either the element a_n does not have an asymptotic density or its density is 0.*

We remark that Theorem 1.13 is not announced as this form in [42], but it is an intermediate result of the proof of Theorem 1 in [42], for which the hypothesis of complete multiplicativity is not necessary. For more details, see Lemma 3, Theorem 1 in [42] and Proposition 3.1 in Chapter 3.

1.3 Substitution and complexity

The definition of Thue-Morse sequence presented in Section 1 can be reformulated by using a monoid endomorphism. Let ϕ be a monoid endomorphism of the set $\{0, 1\}^*$ such that $\phi(0) = 01$ and $\phi(1) = 10$, then

$$\phi^0(1) = 1$$
$$\phi^1(1) = 10$$
$$\phi^2(1) = 1001$$
$$\phi^3(1) = 10010110$$
$$\phi^4(1) = 1001011001101001...$$

So that the Thue-Morse sequence can be defined as $\lim_{n\to\infty} \phi^n(1)$ which exists. Let us denote this limit by $\phi^\infty(1)$.

Definition Let $\sum$ be an alphabet and k be a positive integer, A k-substitution is an endomorphism of $\sum^*$ such that the image of each letter in $\sum$ is of length k.

So the morphism ϕ defined as above is a 2-substitution.

For the Rudin-Shapiro sequence, there is a 2-substitution over an alphabet of 4 elements $\{A, B, C, D\}$ such that

$$s(A) = AB$$
$$s(B) = AC$$
$$s(C) = DB$$
$$s(D) = DC$$

So $s^\infty(A) = ABACABDBABACDCAC...$, then the Rudin-Shapiro sequence is obtained by replacing A, B by 1 and C, D by -1.

Theorem 1.14. *(Cobham [19]) Let k be a positive integer and $(a(n))_{n\geq 0}$ be a sequence. The following conditions are equivalent:*
a) The sequence $(a(n))_{n\geq 0}$ is k-automatic.
b) The k-kernel of $(a(n))_{n\geq 0}$ is finite.
c) The sequence $(a(n))_{n\geq 0}$ is the pointwise image of a fixed point of a k-substitution.

Now let us show how this kind of sequences are classified between “ordered” and totally disordered. To do so, we have to measure the “complexity" of a sequence. Here we introduce some different measures of the complexity of an infinite sequence.

Definition Let u be an infinite sequence of symbols from an alphabet $\sum$. We define the subword complexity $p_u(n)$ of u to be the number of different words of length n in u.

This definition is the one most used in complexity analysis, however some other complexities can be defined such as: Let u be an infinite sequence of symbols from an alphabet $\sum$.

Definition We define the square complexity $pd_u(n)$ of u to be the number of different words of type aa of length n in u, where $a \in \sum^*$.

Definition We define the cube complexity $pt_u(n)$ of u to be the number of different words of type aaa of length n in u, where $a \in \sum^*$.

Definition We define the palindromic complexity $pp_u(n)$ of u to be the number of different words a of length n in u, such that a equals to its reverse, i.e., $a = x_0x_1x_2...x_{k-1}x_kx_{k-1}...x_2x_1x_0$ or $a = x_0x_1x_2...x_{k-1}x_kx_kx_{k-1}...x_2x_1x_0$.

Definition We define the palindromic length complexity $pl_u(n)$ to be the minimal number of decomposition in palindromic words of the prefix of u of size n, i.e.,

$$pl_u(n) = \min\{k | u_0 u_1 u_2 ... u_n = p_1 p_2 ... p_k; p_i \text{ are palindromic words}\}.$$

The subword complexity varies from a constant to exponential:

Example The subword complexity of a constant sequence is constant;
The subword complexity of an ultimately periodic sequence is bounded;
The subword complexity of the Barbier infinite word

$$B = 1234567891011121314151617181920 21...$$

which is the infinite concatenation of the decimal expansions of the positive integers written in increasing order, is exponential: $p_B(n) = 10^n$ for all $n \geq 0$.

Theorem 1.15. *(see,e.g.,[5]) Almost all sequences w over a finite alphabet $\sum$ satisfy $p_w(n) = |\sum|^n$ for all $n \geq 0$.*

Theorem 1.16. *(Cobham [13]) If w is an automatic sequence, then $p_w(n) = \mathcal{O}(n)$ for all $n \geq 0$.*

The last two theorems show the distinction between automatic sequences and random sequences. Here let us consider the Thue-Morse sequences as example:

Theorem 1.17. *([31],[10],[36],[39],[9]) The Thue-Morse sequences $(a(n))_{n\geq 0}$ satisfies*
a) $p_a(n) \leq 4n$ for all $n \geq 0$,
b) all squares in $(a(n))_{n\geq 0}$ are of the form $\phi^k(00), \phi^k(11), \phi^k(010010), or\ \phi^k(101101)$ for some $k \geq 0$, as a consequence, $pd_a(n) \leq 4n$ for all $n \geq 0$,
c) $(a(n))_{n\geq 0}$ is an overlap-free sequence, i.e., it does not contain any words of type ababa with $a \in \sum, b \in \sum^*$, consequently $pt_a(n) = 0$ for all $n \geq 0$,
d) $pp_a(n)$ is 0,2 or 4.

In Chapter 5, the author studies the palindromic length complexity of Thue-Morse sequence and find all sequences which have the same palindromic length complexity as Thue-Morse's. The main result is as follows:

Theorem 1.18. *Let us define a class of infinite sequences $\mathcal{C}$ as follows:*
Let $\sum$ be an alphabet which contains at least two letters and let $a \in \sum$.
Let F be the set of bijections over $\sum$.
Let $(f_n)_{n\in\mathbf{N}}$ be a sequence over F and $(w_n)_{n\in\mathbf{N}}$ be a sequence of finite words over $\sum$ which are defined recursively as:

$$f_i \in F \text{ such that } f_i(w_i) \neq w_i \ \forall i \geq 0,$$

and

$$\begin{cases} w_0 = a \\ w_n = w_{n-1} f_{n-1}(w_{n-1}) f_{n-1}(w_{n-1}) w_{n-1}, \forall n > 0. \end{cases}$$

Then all sequences in $\mathcal{C}$ share the same palindromic length complexity.

1.4 Completely multiplicative functions

Definition A function $f : \mathbf{N} \to \mathbf{G}$, where G is a group, is called completely multiplicative if $f(nm) = f(n)f(m)$. It is called multiplicative if $f(nm) = f(n)f(m)$ $\forall(n, m)$ such that $(n, m) = 1$.

The motivation is to study the completely multiplicative automatic functions, and answer the two conjectures proposed respectively in [28] and [1].

Conjecture 1.1. *For any multiplicative q-automatic function $f : \mathbf{N} \to \mathbf{C}$ there exists an eventually periodic function $g : \mathbf{N} \to \mathbf{C}$ such that $f(p) = g(p)$ for all primes p.*

Definition Let $f : \mathbf{N} \to \mathbf{G}$ be completely multiplicative, we call f as a mock character if there are only finitely many primes p such that $f(p) = 0$.
Let f and g be two completely multiplicative functions and y be a positive integer, we define a distance $\mathbb{D}(\,.\, , \,.\, ; y)$ between these functions as $\mathbb{D}^2(f, g; y) = \sum_{p \leq y, p \in \mathbf{P}} \frac{1 - \Re(f(p)\overline{g(p)})}{y}$.

Conjecture 1.2. *For any mock character k, there exists a Dirichlet character χ such that $\mathbb{D}(k, \chi, y)$ is bounded. Conversely, if $k : \mathbf{Z} \to \mathbf{U}$ is completely multiplicative and a bounded distance from some Dirichlet character, then k must be a mock character. where $\mathbf{U}$ is the complex unit disc.*

This topic has been studied in [42] [1] [27] [30]. In article [42], the author proved that non-vanishing completely multiplicative automatic sequences are almost periodic (defined in [42]). In article [1], the authors gave a formal expression to all sequences in the non-vanishing case and also some examples in the vanishing case. In article [27], Y. Hu studied completely multiplicative sequences in general and extended the results under a weaker condition. Here we prove that completely multiplicative automatic sequences are rare and have the same expression as the one in [1]. To do so, let us begin by reviewing some asymptotic properties of (completely) multiplicative functions.

Let G be a group and let $f : \mathbf{N} \to G$ be a function. It was proved by I. Z. Ruzsa [38] that, if G is an Abelian group and f is a multiplicative function, then all elements in G have a nature density.

Theorem 1.19. *([38]) Let G be an Abelian group and f a multiplicative function over G, let*

$$E = \left\{ g | g \in G, \sum_{f(p)=g, p \in \mathbf{P}} \frac{1}{p} = \infty \right\}$$

where $\mathbf{P}$ is the set of primes and G_1 the subgroup of G generated by E. If

$$|G_1| < \infty, \sum_{f(p) \notin G_1} \frac{1}{p} < \infty$$

then $d(g) > 0$ for all $g \in Imf$ and the value depends only on the coset hG_1 in which g lies, otherwise $d(g) = 0$.

This theorem can be extended to the case $G \cup \{0\}$ with restriction that $\sum_{f(p)=0} \frac{1}{p} < \infty$ and f is completely multiplicative. A similar topic has also been studied by P. Erdös and A. Wintner. In [19] Erdös conjectured that if f evaluates over $\{1, -1\}$, then the mean value $M(f) = \lim_{N \to \infty} \frac{\sum_{i=1}^{N} f(i)}{N}$ exists. H. Delange [15] proved this for all real-evaluated multiplicative functions, under the condition that $\sum_{p \in \mathbf{P}} \frac{1-f(p)}{p} < \infty$ and $|f(n)| \leq 1$. In which case $M(f) \neq 0$ unless $f(2^k) = -1$ for every $k \in N$. After that E. Wirsing [43] proved that if the previous series diverges, then $M(f) = 0$. And these results were extended to complex valued multiplicative functions by G. Halász [26].

To study the subword complexity of f, it may be useful to check also the density of words of arbitrary length. It is conjectured by Elliott [18] that if the multiplicative functions $f_1, f_2, ...f_l$ defined over $\mathbf{U}$ are aperiodic, then for all distinct integers $n_1, n_2, ...n_l$ we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{m=1}^{N} f_0(m) f_1(m + n_1) ... f_l(m + n_l) = 0.$$

What we mean by aperiodic sequences are those sequences f satisfying $M(f(ax + b)) = 0$ for all couple of positive integers a, b. As a consequence, the subword complexity of aperiodic function over a finite alphabet is exponential. This conjecture was proved by N. Frantzikinakis [21] with a stronger hypothesis:

Theorem 1.20. *([21]) Let $p_1, ..., p_l$ be linearly independent polynomials over $\mathbf{Z}$ and $f_1, f_2, ...f_l$ be multiplicative functions over $\mathbf{U}$, at least one of which is strongly aperiodic. Then there exists a sequence of intervals $(M_k)_{k \in N}$ with $M_k \to \infty$ such that*

$$UD - \lim_{n \to \infty} \frac{1}{M_k} \sum_{m=1}^{M_k} f_0(m) f_1(m + p_1(n)) ... f_l(m + p_l(n)) = 0.$$

We write $UD - \lim_{n \to \infty} a(n) = c$, if for every $\epsilon > 0$ the set $\mathbf{n} \in \mathbf{N}^r : |a(\mathbf{n}) - c| \geq \epsilon$ has uniform density 0. A multiplicative function f is called strongly aperiodic if
$\min_{|t| \leq N} \mathbb{D}(f.\chi, n^{it}; N) \to \infty$ as $n \to \infty$ for every Dirichlet character χ.

With the same assumption of strongly aperiodic and dynamical analysis, the same author proves that the non-zero mean value in logarithm density implies the upper linear growth of subword complexity.

Theorem 1.21. *([22]) If the multiplicative function $f : \mathbf{N} \to \mathbf{U}$ has finite range, is strongly aperiodic, and does not converge to zero in logarithmic density, then $\lim_{n\to\infty} \frac{p(n)}{n} = \infty$, where $p(n)$ is the subword complexity of f of length n.*

Another result from dynamical analysis was obtained by M. Lemańczyk and C. Müllner [30] in which they consider an automatic sequence defined by primitive uniform substitution.

Theorem 1.22. *([30]) Let δ be a primitive substitution of constant length k. Then the numerical automatic sequence $(a_n)_{n\in\mathbf{N}}$ generated by infinite substitutions of an initial letter x is orthogonal to any bounded, aperiodic, multiplicative function $f : \mathbf{N} \to \mathbf{C}$ i.e., if $(a_n)_{n\in\mathbf{N}} = \lim_{m\to\infty} \delta^m(x)$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i\leq n} a_i f(i) = 0.$$

As we can see, most of completely multiplicative functions are aperiodic. So their subword complexities are larger than linear, while the subword complexities of automatic sequences are bounded by linear functions. As a consequence, the intersection of these two classes of sequences should be rare. The main result introduced in Chapter 3 is a classification of all completely multiplicative automatic sequences.

Theorem 1.23. *Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS; then, it can be written in the following form:*
-either there is at most one prime p such that $a_p \neq 0$ and $a_q = 0$ for all other primes q
-or $a_n = \epsilon^{v_p(n)} X(\frac{n}{p^{v_p(n)}})$, where $(X(n))_{n\in\mathbf{N}}$ is a Dirichlet character.

Chapter 2

Generalized Dirichlet series of n variables associated with automatic sequences

2.1 Introduction

The propose of this article is to give a sufficient condition for the meromorphic continuation of Dirichlet series of form $\sum_{\underline{x}\in\mathbf{N}_+^n}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{P(\underline{x})^s}$, where $(a_{\underline{x}})_{\underline{x}\in\mathbf{N}_+^n}$ is a q-automatic sequence of n parameters, $\mu_i\in\mathbf{Z}_+$ and $P:\mathbf{N}^n\to\mathbf{R}$ a polynomial, such that P does not have zeros in $\mathbf{Q}_+^n$. Some specific cases for $n=1$ are studied in this article as examples to show the possibility to have an holomorphic continuity on the whole complex plane. Some equivalences between infinite products are also built as consequences of these results. The Dirichlet sequences of the form $\sum_{n=1}^{\infty}\frac{a_n}{n^s}$ have been studied in [2], and our work is a natural generalization of the results in the above article by using the same method of calculations.

2.2 Notation, definitions and basic properties of automatic sequences

Here we define some notation used in this article. We let $\underline{x}$ denote an n-tuple $(x_1,x_2...x_n)$. We say $\underline{x}\geq\underline{y}$ (resp. $\underline{x}>\underline{y}$) if and only if $\underline{x}-\underline{y}\in\mathbf{R}_+^n$(resp. $\underline{x}-\underline{y}\in\mathbf{R}_+^n$), and we have an analogue definition for the symbol $\leq$ (resp. $<$). We let $\underline{x}^{\underline{\mu}}$ denote the n-tuple $(x_1^{\mu_1},x_2^{\mu_2}...x_n^{\mu_n})$. For a constant c, we let $\underline{c}$ denote the tuple $(c,c...c)$ and for two tuples $\underline{x}$ and $\underline{y}$, we let $<\underline{x},\underline{y}>$ denote the real number $\sum_{i=1}^n x_iy_i$. For an n-tuple $(x_1,x_2...x_n)$, we let $||\ ||_d$ denote the norm d and let $||\ ||$ or $||\ ||_2$ denote the norm 2.

Definition Let $q\geq 2$ be an integer. A sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ with values in the set $\mathcal{A}$ is called q-automatic if and only if its q-kernel $\mathcal{N}_q((a_{\underline{x}})_{\underline{x}\geq\underline{0}})$ is finite, where the q-kernel of the sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is the set of subsequences defined by

$$\mathcal{N}_q((a_{\underline{x}})_{\underline{x}\geq\underline{0}}) = \left\{(m_1, m_2...m_n) \longmapsto a_{(q^k m_1+l_1, q^k m_2+l_2,...,q^k m_n+l_n)}; k \geq 0, \underline{(0)} \leq \underline{l} \leq \underline{(q^k-1)}\right\}.$$

Remark A q-automatic sequence necessarily takes finitely many values. Hence we can assume that the set $\mathcal{A}$ is finite.

Because of the definition of q-automatic with n variables, there are some basic properties.

Theorem 2.1. *Let $q \geq 2$ be an integer and $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ be a sequence with values in $\mathcal{A}$. Then, the following properties are equivalent:*

(i) The sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is q-automatic

(ii) There exists an integer $t \geq 1$ and a set of t sequences $\mathcal{N}' = \left\{(a^1_{\underline{x}})_{\underline{x}\geq\underline{0}}, ..., (a^t_{\underline{x}})_{\underline{x}\geq\underline{0}}\right\})$ such that

- the sequence $(a^1_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is equal to the sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$

- the set $\mathcal{N}'$ is closed under the maps $(a_{\underline{x}})_{\underline{x}\geq\underline{0}} \longmapsto (a_{q\underline{x}+\underline{y}})_{\underline{x}\geq\underline{0}}$ for $\underline{0} \leq \underline{y} \leq \underline{q-1}$

(iii) There exist an integer $t \geq 1$ and a sequence $(A_{\underline{x}})_{\underline{x}\geq\underline{0}}$ with values in $\mathcal{A}^t$, that we denote as a column vector, as $(A_{1,1...1}, A_{2,1...1}, A_{1,2...1}...A_{1,1...2}, A_{2,2...1}...)^t$. There exist q^n matrices of size $t \times t$, say $M_{1,1...1}, M_{1,2...1}...M_{q,q..q}$, with the property that each row of each M_i has exactly one entry equal to 1, and the other $t-1$ entries equal to 0, such that:

- the first component of the vector $(A_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is the sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$

- for each $\underline{y}$ such that $\underline{0} \leq \underline{y} \leq \underline{q-1}$, the equality $A_{q\underline{x}+\underline{y}} = M_{\underline{y}} A_{\underline{x}}$ holds.

Proof. It is a natural consequence of the finiteness of the set $\mathcal{N}'$, see for example [40]. □

Proposition 2.1. *Let $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ be a q-automatic sequence and $(b_{\underline{x}})_{\underline{x}\geq\underline{0}}$ be a periodic sequence of period $\underline{c}$. Then the sequence $(a_{\underline{x}} \times b_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is also q-automatic and its q-kernel can be completed in such a way that all transition matrices of the maps $(a_{\underline{x}} \times b_{\underline{x}})_{\underline{x}\geq\underline{0}} \longmapsto (a_{q\underline{x}+\underline{y}} \times b_{q\underline{x}+\underline{y}})_{\underline{x}\geq\underline{0}}$ on the new set are independent on the choice of the values taken by the sequence $(b_{\underline{x}})_{\underline{x}\geq\underline{0}}$.*

Proof. As $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is a q-automatic sequence, we let $\mathcal{N}_a : \left\{(a^{(1)}_{\underline{x}})_{\underline{x}\geq\underline{0}}, (a^{(2)}_{\underline{x}})_{\underline{x}\geq\underline{0}}, ..., (a^{(l)}_{\underline{x}})_{\underline{x}\geq\underline{0}}\right\}$ denote its q-kernel. The sequence $(b_{\underline{x}})_{\underline{x}\geq\underline{0}}$ is a periodic sequence, thus it is also an q-automatic sequence, we let $\mathcal{N}_b : \left\{(b^{(1)}_{\underline{x}})_{\underline{x}\geq\underline{0}}, (b^{(2)}_{\underline{x}})_{\underline{x}\geq\underline{0}}, ..., (b^{(s)}_{\underline{x}})_{\underline{x}\geq\underline{0}}\right\}$ denote the q-kernel of $(b_{\underline{x}})_{\underline{x}\geq\underline{0}}$. As both of the q-kernel are finite, we can conclude that the set of Cartesian product of these two above sets is finite:

$$\mathcal{N}_{ab} : \left\{(a^{(i)}_{\underline{x}} \times b^{(j)}_{\underline{x}})_{\underline{x}\geq\underline{0}} | 0 \leq i \leq l, 0 \leq j \leq s\right\}$$

which is the q-kernel of the sequence $(a_{\underline{x}} \times b_{\underline{x}})_{\underline{x}\geq\underline{0}}$.

To the completion, we remark that there is a onto map from $\mathcal{N}_b'$ to $\mathcal{N}_b$, where $\mathcal{N}_b'$ is the q-kernel of the periodic sequence $(I_{\underline{x}})_{\underline{x}\geq\underline{0}}$ defined by

$$I_{(m_1,m_2,\ldots,m_n)} = (y_1, y_2, \ldots, y_n) \text{ where } y_i \equiv m_i \mod c, 1 \leq i \leq n$$

and the map is defined as

$$\mathcal{N}_b' \to \mathcal{N}_b : (I_{\underline{x}})_{\underline{x}\geq\underline{0}} \to (b_{I_{\underline{x}}})_{\underline{x}\geq\underline{0}}.$$

So it is enough to work on the finite set $\mathcal{N}_a \times \mathcal{N}_b'$. □

Let us consider the Dirichlet series $f(s) = \sum_{\underline{x}\in\mathbf{N}^n/\underline{0}} \frac{a_{\underline{x}}}{p(\underline{x})^s}$, Where $a_{\underline{x}}$ is q-automatic, a necessary condition of the convergence of such series is that $|p(\underline{x})| \to \infty$ when $||\underline{x}|| \to \infty$, here we want to find a sufficient condition.

An achievable assumption for $f(s)$ to be meromorphic is that the polynomial p is elliptic, which means that, if the degree of p is d then the homogeneous polynomial $p_d(\underline{x})$ of $p(\underline{x})$ satisfies the condition

$$p_d(\underline{x}) > 0,\ \forall \underline{x} \in [0, \infty[^n \setminus \{(0, 0, \ldots 0)\}.$$

Before announcing the main theorem, we would like to study some properties of elliptic polynomials:

Lemma 2.1. *Let $(r_1, r_2, .., r_n)$ be a vector on $\mathbf{Z}_+^n$ such that $\sum_{i=1}^n r_i < d$ and $\underline{x} \in \mathbf{R}_+^n$ then:*

$$\frac{x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n}}{\sum_{i=1}^n x_i^d} = \mathcal{O}(||x||^{\sum_{i=1}^n r_i - d}),\ \textit{when } ||x|| \to \infty.$$

Proof. It is enough to see the following inequality:

$$\sum_{i=1}^n x_i^d = \sum_{i=1}^n r_i \frac{1}{r_i} x_i^d = \sum_{i=1}^n \sum_{j=1}^{r_i} \frac{1}{r_i} x_i^d \geq (\prod_{i=1}^n \frac{1}{r_i^{r_i}} x_i^{dr_i})^{\frac{1}{\sum_{i=1}^n r_i}},$$

which is from the inequality of arithmetic and geometric means. With the assumption $\sum_{i=1}^n r_i < d$ and the equivalences between norms, we conclude $\frac{x_1^{r_1} x_2^{r_2} \ldots x_n^{r_n}}{\sum_{i=1}^n x_i^d} \leq \prod_{i=1}^n r_i^{r_i} ||x||_d^{\sum_{i=1}^n r_i - d} = \mathcal{O}(||x||^{\sum_{i=1}^n r_i - d})$, when $||x|| \to \infty$. □

Lemma 2.2. *If a polynomial p is elliptic of degree d and p_d is the homogeneous polynomial of degree d of p, then:*

(i) all coefficients of terms $x_1^d, x_2^d, \ldots, x_n^d$ are positive;

(ii) $p_d(\underline{x}) \to \infty$ when $||x|| \to \infty$;

(iii) there exists a positive number α such that $\forall \underline{x} \in [0, \infty[^n \setminus \{(0, 0, \ldots 0)\}$, $p_d'(\underline{x}) = p_d(\underline{x}) - \alpha \sum_{i=1}^n x_i^d > 0$;

(iv) $p(\underline{x}) \to \infty$ when $||x|| \to \infty$;

(v) there exists a positive number α such that $p'(\underline{x}) = p(\underline{x}) - \alpha \sum_{i=1}^{n} x_i^d \to \infty$ when $||x|| \to \infty$.

Proof. Assertion (i) is straightforward by evaluating the function at
$(1, 0, ..., 0), (0, 1, ..., 0), ..., (0, 0, ..., 1)$.
For (ii), let us consider the set

$$\{p_d(\underline{x}) | \forall \underline{x} \in [0, \infty[^n \setminus \{(0, 0, ...0)\}, ||\underline{x}||_d = 1\},$$

this set is closed because of closed map lemma and every element inside is larger then 0, so that such a set admits a non-zero infimum, let us denote it by ϵ. Then for an arbitrary $\underline{x}$, we have

$$p_d(\underline{x}) = ||\underline{x}||_d^d p_d(\frac{\underline{x}}{||\underline{x}||_d}) \geq ||\underline{x}||_d^d \epsilon.$$

For (iii), Setting

$$p'_d(\underline{x}) = p_d(\underline{x}) - \frac{\epsilon}{2} \sum_{i=1}^{n} x_i^d,$$

it is easy to check $p'_d(\underline{x}) \geq ||x||_d^d \epsilon - \frac{\epsilon}{2} \sum_{i=1}^{n} x_i^d > 0, \forall \underline{x} \in [0, \infty[^n \setminus \{(0, 0, ...0)\}$ and $||\underline{x}||_d = 1$, then it follows that

$$p'_d(\underline{x}) = ||\underline{x}||_d^d \frac{p'_d(\underline{x})}{||\underline{x}||_d^d} = ||\underline{x}||_d^d p'_d(\frac{\underline{x}}{||\underline{x}||_d}) > 0,$$

for all $\underline{x}$ in the set $[0, \infty[^n \setminus \{(0, 0, ...0)\}$.

For (iv) it is enough to point out that each monomial of degree smaller then d can be bounded above by a term of the form $\alpha \sum_{i=1}^{n} x_i^d$ because of Lemma 1, and conclude by (iii).

(v) is a direct consequence of (iii) and (iv).

□

Lemma 2.3. *Let P be an elliptic polynomial of degree d and let p be a polynomial with a degree smaller than d, then there exists un integer C such that for all n-tuples $\underline{x} \in \mathbf{N}_+^n$ with $< \underline{x}, \underline{1} >> C$,*

$$\frac{|p(\underline{x})|}{P(\underline{x})} < \frac{1}{P(\underline{x})^{\frac{1}{2d}}}$$

Proof. Let us consider the polynomial q defined by $q = P^{2d-1} - p^{2d}$. We can check that q is of degree $2d^2 - d$ and its homogeneous polynomial is uniquely defined by the one of P, which is from the fact that the degree of p^{2d} is at most $2d^2 - 2d$. So that q is an elliptic polynomial, thus there exists un integer C such that $P^{2d-1}(\underline{x}) - p^{2d}(\underline{x}) > 0$ if $< \underline{x}, \underline{1} >> C$. As a result, for all $\underline{x}$ with $< \underline{x}, \underline{1} >> C$,

$$(\frac{p(\underline{x})}{P(\underline{x})})^{2d} = \frac{1}{P(\underline{x})} \frac{p(\underline{x})^{2n}}{P(\underline{x})^{2n-1}} < \frac{1}{P(\underline{x})}.$$

□

2.3 Proof of the meromorphic continuation

In this section we prove the main result.

Theorem 2.2. *Let p be an elliptic polynomial of n variables and $(a_{\underline{x}})_{\underline{x}>\underline{0}}$ be q-automatic, then for a given n-tuple $\underline{\mu}$, the function $\sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}$ admits an abscissa of convergence σ such that it converges absolutely on the half plane $\Re(s)>\sigma$ and has a meromorphic continuation on whole complex plane. furthermore, the poles of this function (if any) are located on a finite number of left semi-lattices.*

This result will be obtained by proving several lemmas successively:

Proposition 2.2. *Let $a_{\underline{x}}$ be a q-automatic sequence, and $p(\underline{x})=\sum_{\underline{\alpha}} m_{\underline{\alpha}}x^{\underline{\alpha}}$ be a n-variable homogeneous elliptic polynomial of degree d, let $\underline{\mu}\in\mathbf{N}_+^n$ be a multi-index, for any $\underline{\beta}$ such that $\underline{0}\le\underline{\beta}\le\underline{q}$, define $p_{\underline{\beta}}(\underline{x})=q^{-n}(p(q\underline{x}+\underline{\beta})-p(q\underline{x}))$, then for any $k\in\mathbf{N}$, the function $f_{k,\underline{\beta},\underline{\mu}}: s\longrightarrow\sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{a_{\underline{x}}p_{\underline{\beta}}(\underline{x})^k\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^{s+k}}$ admits an abscissa of convergence $\sigma_{k,\underline{\beta},\underline{\mu}}$ such that $f_{k,\underline{\beta},\underline{\mu}}$ converges absolutely to an holomorphic function on the right half-plane $\Re(s)>\sigma_{k,\underline{\beta},\underline{\mu}}$.*

Proof. We firstly prove that $f_{0,\underline{0},\underline{0}}(s)$ converge when $\Re(s)>n$.

$$\begin{aligned}
|f_{0,\underline{0},\underline{0}}(s)| &= |\sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{a_{\underline{x}}}{p(\underline{x})^s}| \le \sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{|a_{\underline{x}}|}{p(\underline{x})^{\Re(s)}} \le \sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{|a_{\underline{x}}|}{(\alpha<\underline{x},\underline{1}>)^{\Re(s)}}\ (*)\\
&\le \frac{\max(|a_{\underline{x}}|)}{(\alpha)^{\Re(s)}}\Big(\sum_{<\underline{x},\underline{1}><n}\frac{1}{<\underline{x},\underline{1}>^{\Re(s)}}+\sum_{<\underline{x},\underline{1}>\ge n}\frac{1}{<\underline{x},\underline{1}>^{\Re(s)}}\Big)\\
&\le \frac{\max(|a_{\underline{x}}|)}{(\alpha)^{\Re(s)}}\Big(\sum_{<\underline{x},\underline{1}><n}\frac{1}{<\underline{x},\underline{1}>^{\Re(s)}}+\sum_{m\ge n}\frac{\binom{m+n-1}{n-1}}{m^{\Re(s)}}\Big) \qquad (2.1)\\
&\le \frac{\max(|a_{\underline{x}}|)}{(\alpha)^{\Re(s)}}\Big(\sum_{<\underline{x},\underline{1}><n}\frac{1}{<\underline{x},\underline{1}>^{\Re(s)}}+\sum_{m\ge n}\frac{m^{n-1}}{m^{\Re(s)}}\Big)\\
&\le \frac{\max(|a_{\underline{x}}|)}{(\alpha)^{\Re(s)}}\Big(\sum_{<\underline{x},\underline{1}><n}\frac{1}{<\underline{x},\underline{1}>^{\Re(s)}}+\sum_{m\ge n}\frac{1}{m^{\Re(s)+1-n}}\Big).
\end{aligned}$$

Inequality $(*)$ is obtained by Lemma 2.2 (iii) and taking α as it was in the lemma; and the sum $\sum_{m\ge n}\frac{1}{m^{\Re(s)+1-n}}$ exists and is bounded when $\Re(s)>n$.

For any $\underline{\beta}$ such that $\underline{0}\le\underline{\beta}\le\underline{q}$, we remark that $\prod_{i=0}^n(x_i+\beta_i)^{k_i}=\sum_{\underline{l}\le\underline{k}}C_{\underline{l}}\prod_{i=0}^n(x_i)^{l_i}$ with $C_{\underline{l}}<q^n$, which shows that all monomials of the polynomial $p_{\underline{\beta}}$ have a degree not larger than $d-1$. Lemma 2.3 leads $|\frac{p_{\underline{\beta}}(\underline{x})}{p(\underline{x})}|<\frac{1}{p(\underline{x})}^{\frac{1}{2d}}$ for all $\underline{x}$ satisfying $<\underline{x},\underline{1}>>C_1$, with C_1 defined in Lemma 2.3. While Lemma 2.1 and Lemma 2.2 (v) yield that there exists an integer l such that $|\frac{\prod_{i=1}^n x_i^{\mu_i}}{p^l(\underline{x})}|\to 0$ when $|\underline{x}|\to\infty$. As a result, there exists C_2 such that for all $<\underline{x},\underline{1}>\geq C_2, |\frac{\prod_{i=1}^n x_i^{\mu_i}}{p^l(\underline{x})}|<1$. Taking $C=\max(C_1,C_2)$,

$$\begin{aligned}\sum_{(\underline{x})\in\mathbf{N}_+^n}|\frac{a_{\underline{x}}p_{\underline{\beta}}^k(\underline{x})\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^{s+k}}| &\leq \sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{|p^{\Re(s)}(\underline{x})|}|\frac{p_{\underline{\beta}}(\underline{x})}{p(\underline{x})}|^k \\ &\leq \sum_{<\underline{x},\underline{1}><C}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{|p^{\Re(s)}(\underline{x})|}|\frac{p_{\underline{\beta}}(\underline{x})}{p(\underline{x})}|^k+\sum_{<\underline{x},\underline{1}>\geq C}\frac{a_{\underline{x}}}{p^{\Re(s)-l+\frac{k}{2d}}(\underline{x})}.\end{aligned} \tag{2.2}$$

With k a constant in $\mathbf{N}^+$, the above function converges to a holomorphic function on the half plane $\Re(s)>n+l-\frac{k}{2d}$. Furthermore, for all $b>k$, $\sum_{<\underline{x},\underline{1}>\geq x_1}|\frac{a_{\underline{x}}p_{\underline{\beta}}^b(\underline{x})}{p^{s-l+b}(\underline{x})}|$ is bounded on this half plane. □

Proposition 2.3. *With the same notation as above, if p is an homogeneous polynomial, then the function* $F: s\longrightarrow\sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}$ *admits a meromorphic continuation on the whole complex plane.*

Proof. In this proof, we consider the q-automatic sequence $(a_{\underline{x}})_{\underline{x}\geq\underline{0}}$ as itself multiplied by a constant sequence $(b_{\underline{x}})_{\underline{x}\geq\underline{0}}=1$, which is a q-periodic sequence. Because of Proposition 2.1, the q-kernel of this sequence admits a completion, we can define a sequence of vectors $(A_{\underline{x}})_{\underline{x}\geq\underline{0}}$ and the matrices of transition on this completion as in Theorem 2.1.

For any $\underline{\mu}\in\mathbf{N}_+^n$, there exists some $l\in\mathbf{Z}$ such that $<\underline{\mu},\underline{1}><ld$ and a constant $N_0\in\mathbf{N}$ such that $C<N_0nq$, where C is defined as in the previous lemma.

$$\begin{aligned}F_{\underline{\mu}}(s) &= \sum_{(\underline{x})\in\mathbf{N}_+^n}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} = \sum_{(\underline{x})<(\underline{N_0q})}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}+\sum_{(\underline{y})<(\underline{q})}\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}}\frac{A_{q\underline{z}+\underline{y}}\prod_{i=1}^n(qz_i+y_i)^{\mu_i}}{p^s(q\underline{z}+\underline{y})} \\ &= \sum_{(\underline{x})<(\underline{N_0q})}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}+\sum_{(\underline{y})<(\underline{q})}\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}}\frac{A_{q\underline{z}+\underline{y}}\prod_{i=1}^n(qz_i)^{\mu_i}}{p^s(q\underline{z}+\underline{y})} \\ &+\sum_{(\underline{\psi})<(\underline{\mu})}\sum_{(\underline{y})<(\underline{q})}\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}}\frac{A_{q\underline{z}+\underline{y}}C_{\underline{\psi},\underline{y}}\prod_{i=1}^n(qz_i)^{\psi_i}}{p^s(q\underline{z}+\underline{y})},\end{aligned} \tag{2.3}$$

where $C_{\underline{\psi},\underline{y}}$ is uniquely defined by $\underline{y}$ for given $\underline{\psi}$. So the sequence $(C_{\underline{\psi}}(\underline{x}))_{\underline{x}>\underline{0}}$ defined by

$$C_{\underline{\psi}}(\underline{x})=C_{\underline{\psi},\underline{y}}\text{ with }x_i\equiv y_i\mod q, 1\leq i\leq n$$

is periodic as a function of $\underline{x}$, we let $Res_{\underline{\mu}}(s)$ denote the term

$$Res_{\underline{\mu}}(s)=\sum_{(\underline{\psi})<(\underline{\mu})}\sum_{(\underline{y})<(\underline{q})}\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}}\frac{A_{q\underline{z}+\underline{y}}C_{\underline{\psi},\underline{y}}\prod_{i=1}^n(qz_i)^{\psi_i}}{p^s(q\underline{z}+\underline{y})}.$$

Remarking that all sequences $(b_{\underline{\psi}}(\underline{x}))_{\underline{x}>\underline{0}}$ defined by $b_{\underline{\psi}}(\underline{x})=A_{q\underline{z}+\underline{y}}C_{\underline{\psi},\underline{y}}$ if $\underline{x}=q\underline{z}+\underline{y}$ are in the form of a product of a specific q-automatic sequence by a $\underline{q}$-periodic one, because of Proposition 2.1, such sequences admit a unique completion the same one as $(A_{\underline{x}})_{\underline{x}\in\mathbf{N}_+^n}$ has, and the transition matrices on this completion do not depend on the choice of the $\underline{q}$-periodic sequences $(C_{\underline{\psi}}(\underline{x}))_{\underline{x}>\underline{0}}$.

Using the transition matrices, we have:

$$
\begin{aligned}
F_{\underline{\mu}}(s) &= \sum_{(\underline{x})\in\mathbf{N}^n/(\underline{0})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} = \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + \sum_{(\underline{y})<(\underline{q})} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{qz}+\underline{y}}\prod_{i=1}^n (qz_i+y_i)^{\mu_i}}{p^s(q\underline{z}+\underline{y})} \\
&= \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + \sum_{(\underline{y})<(\underline{q})} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{qz}+\underline{y}}\prod_{i=1}^n (qz_i)^{\mu_i}}{p^s(q\underline{z}+\underline{y})} + Res_\mu(s) \\
&= \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + \sum_{(\underline{y})<(\underline{q})} M_{\underline{y}} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{qz}}\prod_{i=1}^n (qz_i)^{\mu_i}}{p^s(q\underline{z})} \frac{1}{(1+\frac{p_y(\underline{z})}{p(\underline{z})})^s} + Res_\mu(s) \\
&= \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + \sum_{(\underline{y})<(\underline{q})} M_{\underline{y}} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{z}}\prod_{i=1}^n (qz_i)^{\mu_i}}{p^s(q\underline{z})} \sum_{k\geq 0} \binom{s+k-1}{k} (\frac{-p_y(\underline{z})}{p(\underline{z})})^k + Res_\mu(s) \\
&= \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + q^{<\underline{\zeta},\underline{1}>-ns} \sum_{(\underline{y})<(\underline{q})} M_{\underline{y}} \sum_{k\geq 0} \binom{s+k-1}{k} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{z}}\prod_{i=1}^n (z_i)^{\mu_i}(-p_y(\underline{z}))^k}{(p(\underline{z}))^{s+k}} \\
&\quad + Res_\mu(s).
\end{aligned}
\tag{2.4}
$$

The above equation gives:

$$
\begin{aligned}
&(Id - q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}})F_{\underline{\mu}}(s) = q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}} \sum_{k\geq 1} \binom{s+k-1}{k} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{z}}\prod_{i=1}^n (z_i)^{\mu_i}(-p_y(\underline{z}))^k}{(p(\underline{z}))^{s+k}} \\
&+ \sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + Res_\mu(s).
\end{aligned}
\tag{2.5}
$$

By multiplying by $com^t(Id - q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}})$ on both side, we have:

$$
\begin{aligned}
&det(Id - q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}})F_{\underline{\mu}}(s) = com^t(Id - q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}})(\sum_{(\underline{x})<(\underline{N_0 q})} \frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s} + Res_\mu(s) \\
&+ q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}} \sum_{k\geq 1} \binom{s+k-1}{k} \sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{z}}\prod_{i=1}^n (z_i)^{\mu_i}(-p_y(\underline{z}))^k}{(p(\underline{z}))^{s+k}}).
\end{aligned}
\tag{2.6}
$$

Because of Proposition 2.2, the infinite sum $Res_\mu(s)$ converges absolutely when $\Re(s) > l+n-\frac{1}{2d}$, and the infinite sum $\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}} \frac{A_{\underline{z}}\prod_{i=1}^n (z_i)^{\mu_i}(-p_y(\underline{z}))^k}{(p(\underline{z}))^{s+k}}$ is also convergent and bounded when $\Re(s) > l+n-\frac{k}{2d}$. Equation 2.5 shows that all terms on the right-hand side present a meromorphic continuity for $\Re(s) > l+n-\frac{1}{2d}$, so that $F_{\underline{\mu}}(s)$ has a meromorphic continuation on the half-plane $\Re(s) > l+n-\frac{1}{2d}$.

To guarantee that this argument works recursively over all rational numbers of type $l+n-\frac{k}{2d}$, we have to check that the meromorphic continuity of the $Res_\mu(s)$ can also be extended in this way. Using once more the above argument over $Res_{\mu}(s)$, we can deduce that this infinite sum can be extended as a meromorphic function on the half-plane $\Re(s) > l + n - \frac{2}{2d}$, however, with a new term "$Res_{\mu}(s)$", let us call it $Res'_{\mu}(s)$. Once more, we have to do the same thing for $Res'_{\mu}(s)$. But remarking that after each operation, the degree of the monomial at the numerator decreases strictly, so after finitely many times of such operation, the term Res vanishes. This fact guarantees that the iteration can be done successively to prove the meromorphic continuation of $F_{\underline{\mu}}(s)$ on the whole complex plane.

Furthermore, the poles of such a function can only be located at the zeros of the function $s \longrightarrow det(Id - q^{<\underline{\mu},\underline{1}>-ns} \sum_{(\underline{y})<(\underline{q})} M_{\underline{y}})$ for an arbitrary $\underline{\mu} \in \mathbf{N}_+^n$, so we conclude that all poles of function $F(s)$ are located in the set

$$s = \frac{1}{n}(\frac{\log \lambda}{\log q} + \frac{2ik\pi}{\log q} - l),$$

with λ any eigenvalue of the matrix $\sum_{(\underline{y})<(\underline{q})} M_{\underline{y}}, k \in \mathbf{Z}, l \in \mathbf{Z}$ and log is defined as complex logarithm.

□

Proof of Theorem 2.2. Let us write the polynomial $p(\underline{x})$ in the form $p(\underline{x}) = p_d(\underline{x}) + Res(\underline{x})$ where $p_d(\underline{x})$ is the homogeneous polynomial with maximum degree of $p(\underline{x})$, say d. By Lemma 1, $\frac{Res(\underline{x})}{p_d(\underline{x})} = \mathcal{O}(|x|^{-1/d})$. So for a given number $m \in \mathbf{R}$, there exist $x_0 \in \mathbf{N}_+$ and a positive integer k_0 such that for all $|\underline{x}| > x_0$ and all $k > k_0$, $|\frac{Res^k(\underline{x}) \prod_{i=1}^n x_i^{\mu_i}}{p_d^{k+m}(\underline{x})}| < (\frac{1}{2^k})$. For any given half-plane $\{s|\mathcal{R}_e(s) > m, m \in \mathbf{R}\}$, take an integer $s_0 > \max\{k_0, |m|\}$, we can compute that

$$\begin{aligned}
&|\sum_{|\underline{x}|>x_0} \frac{A_{\underline{x}} \prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s} \sum_{k=s_0+1}^{\infty} \binom{-s}{k} \frac{Res^k(\underline{x})}{p_d^k(\underline{x})}| = |\sum_{|\underline{x}|>x_0} \frac{A_{\underline{x}}}{(p_d(\underline{x}))^{s-m}} \sum_{k=s_0+1}^{\infty} \binom{-s}{k} \frac{Res^k(\underline{x}) \prod_{i=1}^n x_i^{\mu_i}}{p_d^{k+m}(\underline{x})}| \\
&< \sum_{|\underline{x}|>x_0} \frac{A_{\underline{x}}}{(p_d(\underline{x}))^{\Re(s-m)}} |\sum_{k=s_0+1}^{\infty} \binom{-\Re(s)}{k}| \frac{Res^k(\underline{x}) \prod_{i=1}^n x_i^{\mu_i}}{p_d^{k+m}(\underline{x})} \\
&< (\sum_{|\underline{x}|>x_0} \frac{A_{\underline{x}}}{(p_d(\underline{x}))^{\Re(s-m)}}) \mathcal{O}_{s_0}(\frac{1+|s|^{s_0+1}}{2^{s_0}}).
\end{aligned}$$

The above fact shows that the function $\phi(s) = \sum_{|\underline{x}|>x_0} \frac{A_{\underline{x}} \prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s} \sum_{k=s_0+1}^{\infty} \binom{-s+k}{k} \frac{Res^k(\underline{x})}{p_d^k(\underline{x})}$ admits a holomorphic continuation over the half-plane $\{s|\Re(s) > m, m \in \mathbf{R}\}$. Now let us consider the equivalence as below:

$$\sum_{(\underline{x})\in\mathbf{N}_+^n/(\underline{0})}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(x)^s}=\sum_{|\underline{x}|\leq x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(x)^s}+\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s}\frac{1}{(1+\frac{Res(\underline{x})}{p_d(\underline{x})})^s}$$
$$=\sum_{|\underline{x}|\leq x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(x)^s}+\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s}\sum_{k=0}^{s_0}\binom{-s+k}{k}\frac{Res^k(\underline{x})}{p_d^k(\underline{x})}+\phi(s). \quad (2.7)$$

For each $k\in\mathbf{N}$,

$$\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s}\frac{Res^k(\underline{x})}{p_d^k(\underline{x})}=\sum_{\underline{j}\leq k\underline{\mu}}\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\mathcal{C}_{\underline{i}}}{(p_d(\underline{x}))^{s+k}}\prod_{i=1}^n x_i^{j_i}, \quad (2.8)$$

where $\mathcal{C}_{\underline{i}}$ are constants depending on k, and $s\longrightarrow\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\mathcal{C}_{\underline{i}}}{(p_d(\underline{x}))^{s+k}}\prod_{i=1}^n x_i^{j_i}$ are meromorphic functions because of the previous lemma. As there are finitely many meromorphic function in (2.8), we can conclude that for every $k>0$, $s\longrightarrow\sum_{|\underline{x}|>x_0}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{(p_d(\underline{x}))^s}Res^k(\underline{x})$ is meromorphic. This fact implies that for an arbitrary $s_0\in\mathbf{R}$ the function $f(s)$ is a finite sum of meromorphic functions on the half-plane $\Re(s)>s_0$, so $f(s)$ itself is meromorphic on this half plane. As a result, the function $s\longrightarrow\sum_{(\underline{x})\in\mathbf{N}^n/(\underline{0})}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}$ is meromorphic on the whole complex plane.

□

Proposition 2.4. *Let $f(s)=\sum_{(\underline{x})\in\mathbf{N}^n/(\underline{0})}\frac{a_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}$ be the function defined as in Theorem 2.2. Let s_0 be its first pole on the axis of real numbers counting from plus infinity to minus infinity. Then the function $H(s)$ has a simple pole at this point.*

Proof. We recall a classical result on matrices (see [35]): Let B be a matrix of size $t\times t$ over any commutative field, $p_B(X)$ be its characteristic polynomial, and $\pi_B(X)$ be its monic minimal polynomial. Let $\Delta(X)$ be the monic gcd of the entries of (the transpose of) the comatrix of the matrix $(B-XI)$, then:

$$p_B(X)=(-1)^t\pi_B(X)\Delta(X)$$

We let B denote the matrix $(nq)^{-1}\sum_{(\underline{y})<(\underline{q})}M_{\underline{y}}$ and by T its size. By dividing by $\Delta(q^{n(s-1)+<\underline{\mu},\underline{1}>})$ both sides of Formula 2.6, we get:

$$\pi_B(q^{n(s-1)+<\underline{\mu},\underline{1}>})H(s)=\frac{com^t(q^{n(s-1)+<\underline{\mu},\underline{1}>}Id-\sum_{(\underline{y})<(\underline{q})}M_{\underline{y}})}{\Delta(q^{n(s-1)+<\underline{\mu},\underline{1}>})}(\sum_{(\underline{x})<(\underline{N_0q})}\frac{A_{\underline{x}}\prod_{i=1}^n x_i^{\mu_i}}{p(\underline{x})^s}+Res_\mu(s)$$
$$+q^{<\underline{\mu},\underline{1}>-ns}\sum_{(\underline{y})<(\underline{q})}M_{\underline{y}}\sum_{k\geq1}\binom{-s+k}{k}\sum_{(\underline{z})\in\mathbf{N}^n/\{\underline{t}<\underline{N_0}\}}\frac{A_{\underline{z}}\prod_{i=1}^n(z_i)^{\mu_i}(-p_y(\underline{z}))^k}{(p(\underline{z}))^{s+k}}). \quad (2.9)$$

The right-hand side of the above function is holomorphic when $\Re(s) > s_0$. As s_0 is the first pole of $H(s)$ on the real axis counting from plus infinity, it is a zero of the function $\pi_B(q^{n(s-1)+<\underline{\mu},\underline{1}>})$ associated with the eigenvalue 1 of the matrix B. On the other hand, as B is a stochastic matrix, $\pi_B(x)$ has a simple root at 1, so the function $\pi_B(q^{n(s-1)+<\underline{\mu},\underline{1}>})$ has a simple root at s_0 which concludes the proposition.

□

2.4 Review and remarks

The critical point of the above proof is the development of the term $(1-\frac{p_d(\underline{x}+\underline{\mu})-p_d(\underline{x})}{p_d(\underline{x})})^{-s}$ as an infinite sum, which works only if $|\frac{p_d(\underline{x}+\underline{\mu})-p_d(\underline{x})}{p_d(\underline{x})}|<1$. As we have to use this fact successively to deal with the term $Res(s)$ in Proposition 2.3, what we need actually is that $|\frac{\partial_{\underline{\mu}}(p_d)}{p_d}|$ is bounded by 1 for all $\underline{\mu}$ such that $<\underline{\mu},\underline{1}>\leq d$ when $||\underline{x}||$ is large. The assumption of ellipticity of the polynomial is a particular case of the above propriety. So we may expect to achieve the same result under the assumption

$$|\frac{\partial_{\underline{\mu}}(p_d)}{p_d}(\underline{x})|<1, \text{for all } <\underline{\mu},\underline{1}>\leq d, x_i\geq 1.$$

We may compare this assumption with that in [20], saying

$$\frac{\partial_{\underline{\mu}}(p_d)}{p_d}(\underline{x})=\mathcal{O}(1), x_i\geq 1,$$

which is the weakest assumption known to have a meromorphic continuation of Dirichlet series $\sum_{\underline{x}\in\mathbf{N}_+^n}\frac{1}{P(\underline{x})^s}$.

2.5 Infinite products

Let $P(x)=\sum_{i=0}^d a_i x^i$ be a polynomial which does not have zeros on $\mathbb{Q}$ and $\tilde{P}(x)$ be the polynomial defined by $\tilde{P}(x)=\sum_{i=0}^{d-1}-\frac{a_i}{a_d}x^{n-i}$, by definition, we have $P(x)=a_d x^d - a_d x^d\tilde{P}(\frac{1}{x})$. Let us define $c_i=\frac{a_i}{a_d}$ for all $i=0,1,...,d-1$.

In this section we consider two Dirichlet series generated by 1-index automatic sequences:

$$f(s)=\sum_{n=0}^{\infty}\frac{(\zeta)^{S_q(n)}}{(P(n+1))^s}$$

$$g(s)=\sum_{n=1}^{\infty}\frac{(\zeta)^{S_q(n)}}{(P(n))^s},$$

where q and r are two integers satisfying $2\leq r\leq q$ and r divides q, ζ is a r-th root of unity, such that $\zeta\neq 1$, $S_q(n)$ is the sum of digits of n in the q-ary expansion satisfying

$S_q(0) = 0$ and $S_q(qn + a) = S_q(n) + a$ for $0 \le a \le q - 1$.

Let us define:

$$\phi(s) = \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^s}$$

$$\psi(s) = \sum_{n=1}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n)^s}.$$

It is proved in [14] that ϕ and ψ have holomorphic continuations to the whole complex plane, and $\psi(s)(q^s - 1) = \phi(s)(\zeta q^s - 1)$ for all $s \in \mathbb{C}$.

Proposition 2.5. *f and g also have holomorphic continuations to the whole complex plane if $c_1, c_2, \ldots$ satisfy $max|c_i| < \frac{1}{d}$.*

Proof. We firstly remark that the hypothesis of $\max|c_i| < \frac{1}{d}$ implies the fact $|\tilde{P}(\frac{1}{n+1})| < 1$ for any $n \in \mathbf{N}_+$. Indeed, $|\tilde{P}(\frac{1}{n+1})| = |\sum_{i=0}^{d-1} -\frac{a_i}{a_d}(n+1)^{i-d}| \le \sum_{i=0}^{d-1} |\frac{a_i}{a_d}| < 1$.

$$\begin{aligned}
f(s) &= \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(P(n+1))^s} \\
&= a_d^{-s} \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^{ds}(1-\tilde{P}(\frac{1}{n+1}))^s} \\
&= a_d^{-s} \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^{ds}} \sum_{k=0}^{\infty} \binom{s+k-1}{k} \tilde{P}(\frac{1}{n+1}))^k \\
&= a_d^{-s} \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^{ds}} \sum_{k=0}^{\infty} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l}(n+1)^{-l} \qquad (2.10) \\
&= a_d^{-s} \sum_{k=0}^{\infty} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l} \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^{ds+l}} \\
&= a_d^{-s} \sum_{k=0}^{\infty} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l}\phi(ds+l) \\
&= a_d^{-s}\phi(ds) + a_d^{s} \sum_{k=1}^{\infty} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l}\phi(ds+l),
\end{aligned}$$

where $m_{k,l} = \sum_{M_{k,l} \in \mathcal{P}(\{c_i|1 \le i \le n-1\})} \prod_{c_i \in M_{k,l}} c_i$ and $M_{k,l}$ are sets of k elements included in $\{c_i|1 \le i \le n-1\}$ and the sum of indices of its elements equals l. The hypothesis $m = \max|ci| < \frac{1}{d}$ shows that $|\sum_{l=k}^{dk-k} m_{k,l}| \le (md)^k < 1$, so the right-hand side of (2.10) converges uniformly over the half plane $\{s|\Re(s) > 0\}$ because $\phi(s)$ is bounded for large $|s|$. By the same argument as in Theorem 2.2, we prove successively the holomorphic continuation of f on the whole complex plane. □

It is easy to check $f(0) = 0$, and because of the uniform convergence of the right-hand side of (2.10) over the half-plane $\{s|\Re(s) > 0\}$, dividing by s and letting s tend to 0 we have:

$$\begin{aligned} f'(0) &= d\phi'(0) + a_d^{-s} \sum_{k=1}^{\infty} \lim_{s\to 0} \frac{1}{s} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l}\phi(ds+l) \\ &= -d\log q/(\zeta-1) + \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\phi(l) \\ &= -d\log q/(\zeta-1) + \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l} \sum_{n=0}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n+1)^l} \\ &= -d\log q/(\zeta-1) + \sum_{n=0}^{\infty} (\zeta)^{S_q(n)} \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l} \frac{1}{(n+1)^l} \\ &= -d\log q/(\zeta-1) + \sum_{n=0}^{\infty} (\zeta)^{S_q(n)} \sum_{k=1}^{\infty} k^{-1} \tilde{P}^k(\frac{1}{n+1}) \\ &= -d\log q/(\zeta-1) + \sum_{n=0}^{\infty} (\zeta)^{S_q(n)} \log(1-\tilde{P}(\frac{1}{n+1})), \end{aligned} \tag{2.11}$$

on the other hand, one has for all s, $\psi(s)(q^s-1) = \phi(s)(\zeta q^s - 1)$.

$$\begin{aligned} f'(0) &= d\phi'(0) + a_d^{-s} \sum_{k=1}^{\infty} \lim_{s\to 0} \frac{1}{s} \binom{s+k-1}{k} \sum_{l=k}^{dk} m_{k,l}\phi(ds+l) \\ &= -d\log q/(\zeta-1) + \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\psi(l)(q^l-1)/(\zeta q^l-1) \\ &= -d\log q/(\zeta-1) + \zeta^{-1} \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\psi(l) + (\zeta^{-1}-1) \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\psi(l)/(\zeta q^l-1). \end{aligned} \tag{2.12}$$

By the same method as above, we can deduce by calculating $g'(0)$

$$\sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\psi(l) = \sum_{n=1}^{\infty} (\zeta)^{S_q(n)} \log(1-\tilde{P}(\frac{1}{n}))$$

and

$$\sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l}\psi(l)/(\zeta q^l - 1) = \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l} \sum_{n=1}^{\infty} \frac{(\zeta)^{S_q(n)}}{(n)^l} \sum_{r=1}^{\infty} (\zeta q^l)^{-r}$$
$$= \sum_{r=1}^{\infty} (\zeta)^{-r} \sum_{n=1}^{\infty} (\zeta)^{S_q(n)} \sum_{k=1}^{\infty} k^{-1} \sum_{l=k}^{dk} m_{k,l} \frac{1}{(nq^r)^l}$$
$$= \sum_{r=1}^{\infty} (\zeta)^{-r} \sum_{n=1}^{\infty} (\zeta)^{S_q(n)} \log(1 - \tilde{P}(\frac{1}{nq^r})).$$

As a consequence,

$$\sum_{n=0}^{\infty} (\zeta)^{S_q(n)} \log(1-\tilde{P}(\frac{1}{n+1})) = \zeta^{-1} \sum_{n=1}^{\infty} (\zeta)^{S_q(n)} \log(1-\tilde{P}(\frac{1}{n})) + (\zeta^{-1}-1) \sum_{r=1}^{\infty} \sum_{n=1}^{\infty} (\zeta)^{S_q(n)-r} \log(1-\tilde{P}(\frac{1}{nq^r})). \tag{2.13}$$

Proposition 2.6. *We have the equality*

$$\prod_{n=0}^{\infty} (1 - \tilde{P}(\frac{1}{n+1}))^{\zeta^{S_{q} n}} \times \prod_{n=1}^{\infty} (1 - \tilde{P}(\frac{1}{n}))^{-\zeta^{S_{q} n - 1}} \times (\prod_{r=1}^{\infty} \prod_{n=1}^{\infty} (1 - \tilde{P}(\frac{1}{nq^r}))^{\zeta^{S_{q} n - r}})^{1-\zeta^{-1}} = 1$$

Moreover, if we suppose, for any j

$$x_j(m) = \begin{cases} \frac{r-1}{r} & \text{if } s_q(m) = j \text{ (mod r)} \\ -\frac{1}{r} & \text{if } s_q(m) \neq j \text{ (mod r)} \end{cases}$$

We clearly have

$$\sum_{j \mod r} x_j(m) = 0.(*)$$

Furthermore,

$$\sum_{j \mod r} x_j(m)\zeta^j = \zeta^{S_q(m)}.(**)$$

Using $(*)$, Formula (2.13) can be reformulated as

$$\sum_{j \mod r} \zeta^j \sum_{n=0}^{\infty} x_j(n)(\log(1 - \tilde{P}(\frac{1}{n+1}) + \sum_{r=1}^{\infty} \zeta^{-r} \log(1 - \tilde{P}(\frac{1}{nq^r}))) =$$
$$\sum_{j \mod r} \zeta^{j-1} \sum_{n=1}^{\infty} x_j(n)(\log(1 - \tilde{P}(\frac{1}{n}) + \sum_{r=1}^{\infty} \zeta^{-r} \log(1 - \tilde{P}(\frac{1}{nq^r}))). \tag{2.14}$$

Let now η be a primitive root of unity, we can apply relation (2.14) successively to $\zeta = \eta^a$ for $a = 1, 2, ..., r-1$.
Because of $(**)$, we also have

$$\begin{aligned}&\sum_{j \mod r}\sum_{n=0}^{\infty} x_j(n)(\log(1-\tilde{P}(\frac{1}{n+1}))+\sum_{r=1}^{\infty}\zeta^{-r}\log(1-\tilde{P}(\frac{1}{nq^r}))) = \\ &\sum_{j \mod r}\sum_{n=1}^{\infty} x_j(n)(\log(1-\tilde{P}(\frac{1}{n}))+\sum_{r=1}^{\infty}\zeta^{-r}\log(1-\tilde{P}(\frac{1}{nq^r}))) = 0.\end{aligned} \tag{2.15}$$

Define the matrices: Mat_1 to be $Mat_1 = (\eta^{ij})$ and Mat_2 to be $Mat_2 = (\eta^{ij-i})$, $i = 0, 1, ..., r-1; j = 0, 1, ..., r-1$, define λ and β by

$$\lambda(j) = \sum_{n=0}^{\infty} x_j(n)(\log(1-\tilde{P}(\frac{1}{n+1}))+\sum_{r=1}^{\infty}\zeta^{-r}\log(1-\tilde{P}(\frac{1}{nq^r})))$$

$$\beta(j) = \sum_{n=1}^{\infty} x_j(n)(\log(1-\tilde{P}(\frac{1}{n}))+\sum_{r=1}^{\infty}\zeta^{-r}\log(1-\tilde{P}(\frac{1}{nq^r}))).$$

Let

$$A = \begin{pmatrix} \lambda(0) \\ \lambda(1) \\ ... \\ \lambda(r-1) \end{pmatrix}$$

$$B = \begin{pmatrix} \beta(0) \\ \beta(1) \\ ... \\ \beta(r-1) \end{pmatrix}.$$

Then we have

$$Mat_1 A = Mat_2 B.$$

On the other hand, A is invertible and $Mat_1 = Mat_2 \times Mat_3$ with

$$Mat_3 = \begin{pmatrix} 0 & 0 & ... & 0 & 1 \\ 1 & 0 & ... & 0 & 0 \\ 0 & 1 & ... & 0 & 0 \\ 0 & 0 & ... & 1 & 0 \end{pmatrix}.$$

So we have

$$A = Mat_3 \times B.$$

Proposition 2.7. *We have the equality $\lambda(i) = \beta(i-1)$ for $i = 1, 2, ..., r-1$ and $\lambda(0) = \beta(r-1)$, which leads to, for $i = 1, 2, ..., r-1$,*

$$\prod_{n=0}^{\infty}((1-\tilde{P}(\frac{1}{n+1})\times\prod_{r=1}^{\infty}(1-\tilde{P}(\frac{1}{nq^r})))^{x_j(n)} = \prod_{n=1}^{\infty}((1-\tilde{P}(\frac{1}{n})\times\prod_{r=1}^{\infty}(1-\tilde{P}(\frac{1}{nq^r})))^{x_{j-1}(n)},$$

and for $i = 0$,

$$\prod_{n=0}^{\infty}((1-\tilde{P}(\frac{1}{n+1}) \times \prod_{r=1}^{\infty}(1-\tilde{P}(\frac{1}{nq^r})))^{x_0(n)} = \prod_{n=1}^{\infty}((1-\tilde{P}(\frac{1}{n}) \times \prod_{r=1}^{\infty}(1-\tilde{P}(\frac{1}{nq^r})))^{x_{r-1}(n)}.$$

2.6 Acknowledgement

We would like to thank D. Essouabri for his helpful remarks and discussions on the meromorphic continuation of general Dirichlet series.

Chapter 3

On completely multiplicative automatic sequences

3.1 Introduction

In this article, we describe the decomposition of completely multiplicative automatic sequences, which will be referred to as CMAS. In article [42], the author proves that a non-vanishing CMAS is almost periodic (defined in [42]). In article [1], the authors give a formal expression of all non-vanishing CMAS and also some examples in the vanishing case (named mock characters). In article [27], the author studies completely multiplicative sequences, which will be referred to as CMS, taking values in a general field that have finitely many prime numbers such that $a_p \neq 1$; she proves that such CMS have complexity $p_a(n) = O(n^k)$, where $k = \# \{p | p \in \mathbf{P}, a_p \neq 1, 0\}$. In this article, we prove that all completely multiplicative sequences $(a_n)_{n\in\mathbf{N}}$ defined on $\mathbf{C}$, vanishing or not, can be written in the form $a_n = b_n \chi_n$, where $(b_n)_{n\in\mathbf{N}}$ is an almost constant sequence, and $(\chi_n)_{n\in\mathbf{N}}$ is a Dirichlet character.

Let us consider a CMAS $(a_n)_{n\in\mathbf{N}}$ defined on $\mathbf{C}$. We first prove that all CMAS are mock characters (defined in [1]) with an exceptional case. Second, we study the CMAS satisfying the condition C:

$$\sum_{p | a_p \neq 1, p \in \mathbf{P}} \frac{1}{p} < \infty,$$

where $\mathbf{P}$ is the set of prime numbers. We prove that in this case, there is at most one prime p such that $a_p \neq 1$ or 0. In the third part, we prove that all CMAS are either Dirichlet-like sequences or strongly aperiodic sequences. Finally, we conclude by proving that a strongly aperiodic sequence cannot be automatic.

3.2 Definitions, notation and basic propositions

Let us recall the definition of automatic sequences and complete multiplicativity:

Definition Let $(a_n)_{n\in\mathbf{N}}$ be an infinite sequence and $k \geq 2$ be an integer; we say that this sequence is k-automatic if there is a finite set of sequences containing $(a_n)_{n\in\mathbf{N}}$ and closed under the maps

$$a_n \to a_{kn+i}, i = 0, 1, ...k - 1.$$

There is another definition of a k-automatic sequence $(a_n)_{n\in\mathbf{N}}$ via an automaton. An automaton is an oriented graph with one state distinguished as the initial state, and, for each state, there are exactly k edges pointing from this state to other states; these edges are labeled as $0, 1, ..., k-1$. There is an output function f, which maps the set of states to a set U. For an arbitrary $n \in \mathbf{N}$, the n-th element of the automatic sequence can be computed as follows: writing the k-ary expansion of n, start from the initial state and move from one state to another by taking the edge read in the k-ary expansion one by one until stopping on some state. The value of a_n is the evaluation of f on the stopping state. If we read the expansion from right to left, then we call this automaton a reverse automaton of the sequence; otherwise, it is called a direct automaton.

In this article, all automata considered are direct automata.

Definition We define a subword[1] of a sequence as a finite length string of the sequence. We let $\overline{w}_l$ denote a subword of length l.

Definition Let $(a_n)_{n\in\mathbf{N}}$ be an infinite sequence. We say that this sequence is completely multiplicative if, for any $p, q \in \mathbf{N}$, we have $a_p a_q = a_{pq}$.

It is easy to see that a CMAS can only take finite many values, either 0 or a k-th root of unity (see, for example, Lemma 1 [42]).

Definition Let $(a_n)_{n\in\mathbf{N}}$ be a CMS. We say that a_p is a prime factor of $(a_n)_{n\in\mathbf{N}}$ if p is a prime number and $a_p \neq 1$. Moreover, we say that a_p is a non-trivial factor if $a_p \neq 0$, and we say that a_p is a 0-factor if $a_p = 0$. We say that a sequence $(a_n)_{n\in\mathbf{N}}$ is generated by $a_{p_1}, a_{p_2}, ...$ if and only if $a_{p_1}, a_{p_2}, ...$ are the only prime factors of the sequence.

[1] what we call a *subword* here is also called a *factor* in the literature; however, we use *factor* with a different meaning.

Definition We say that a sequence is an almost-0 sequence if there is only one non-trivial factor a_p and $a_q = 0$ for all primes $q \neq p$.

Proposition 3.1. *Let $(a_n)_{n\in\mathbf{N}}$ be a k-CMAS and q be the number of states of a direct automaton generating $(a_n)_{n\in\mathbf{N}}$; then, for any $m, y \in \mathbf{N}$, we have equality between the sets* $\{a_n | mk^{q!} \leq n < (m+1)k^{q!}\} = \{a_n | mk^{yq!} \leq n < (m+1)k^{yq!}\}$.

Proof. In article [42] (Lemma 3 and Theorem 1), the author proves that, in an automaton, every state that can be reached from a specific state, say, s, with $q!$ steps, can be reached with $yq!$ steps for every $y \geq 1$; conversely, if a state can be reached with $yq!$ steps for some $y \geq 1$, then it can already be reached with $q!$ steps. This proves the proposition.

□

Let us consider a CMS $(a_n)_{n\in\mathbf{N}}$ taking values in a finite Abelian group G. We define

$$E = \left\{ g | g \in G, \sum_{a_p = g, p\in\mathbf{P}} \frac{1}{p} = \infty \right\}$$

and G_1 as the subgroup of G generated by E.

Definition Let $(a_n)_{n\in\mathbf{N}}$ be a sequence, we say that the sequence $(a_n)_{n\in\mathbf{N}}$ has a mean value if and only if $\lim_{N\to\infty} \frac{\sum_{n=0}^{N} a_n}{N+1}$ exists.

Proposition 3.2. *Let $(a_n)_{n\in\mathbf{N}}$ be a CMS taking values in a finite Abelian group G; then all elements $g \in G$ have a non-zero natural density. Furthermore, this density depends only on the coset rG_1 on which the element g lies. The statement is still true in the case that G is a semi-group generated by a finite group and 0 under the condition that there are finitely many primes p such that $a_p = 0$.*

Proof. When G is an Abelien group, the proposition is proved in Theorem 3.10, [38], and when G is a semi-group, Theorem 7.3, [38] shows that all elements in G have a natural density. To conclude the proof, it is sufficient to consider the following fact: let f_0 be a CMS such that there exists a prime p with $f_0(p) = 0$, and let f_1 be another CMS such that

$$f_1(q) = \begin{cases} f_0(q) \text{ if } q \in \mathcal{P}, q \neq p \\ 1 \text{ otherwise,} \end{cases}$$

If $d_0(g)$ and $d_1(g)$ denote the natural density of g in the sequence $(f_0(n))_{n\in\mathbf{n}}$ and $(f_1(n))_{n\in\mathbf{N}}$, respectively, then we have the equality

$$d_1(g) = d_0(g) + \frac{1}{p} d_0(g) + \frac{1}{p^2} d_0(g) ... = \frac{p}{p-1} d_0(g).$$

Doing this iteratively until a non-vanishing sequence, we can conclude the proof by the first part of the proposition. □

3.3 Finiteness of the numbers of 0-factors

In this section, we will prove that a CMAS is either a mock character, which means that it has only finitely many 0-factors, or an almost-0 sequence, that is, $a_m = 0$ for all m that are not a power of p, and $a_{p^k} = \delta^k$ for some δ, where δ is a root of unity or 0 and p is a prime number.

Proposition 3.3. *Let $(a_n)_{n\in\mathbf{N}}$ be a p-CMAS; then, it is either a mock character or an 0-almost sequence.*

Proof. If $(a_n)_{n\in\mathbf{N}}$ is not a mock character, then it contains infinitely many 0-factors. Here, we prove that, in this case, if there is some $a_m \neq 0$, then m must be a power of p, and p must be a prime number. Let us suppose that there are q states of the automaton generating the sequence. As there are infinitely many 0-factors, it is easy to find a subword of length $p^{2q!}$ such that all its elements are 0:
This is equivalent to finding some $m \in \mathbf{N}$ and $p^{2q!}$ 0-factors, say, $a_{p_1}, a_{p_2}, ..., a_{p_{p^{2q!}}}$, such that

$$\begin{cases} m \equiv 0 \pmod{p_1} \\ m+1 \equiv 0 \pmod{p_2} \\ m+2 \equiv 0 \pmod{p_3} \\ \ldots \\ m+p^{2q!}-1 \equiv 0 \pmod{p_{p^{2q!}}} \end{cases}$$

If m is a solution of the above system, then the subword $\overline{a_m a_{m+1} \ldots a_{m+p^{2q!}-1}}$ is constant to 0. Therefore, there exists an m' such that $m \le m'p^{q!} < (m'+1)p^{q!} \le m + p^{2q!}$. Because of Proposition 3.1, for any $y \in \mathbf{N}$, $a_k = 0$ for all k such that $m'p^{yq!} \le k < (m'+1)p^{yq!}$. Taking an arbitrary prime r, if r and p are not multiplicatively dependent, then $a_r = 0$ because there exists a power of r satisfying $m'p^{yq!} \le r^t < (m'+1)p^{yq!}$. This inequality holds because we can find some integers t and y such that

$$\log_p m' \le t \log_p r - yq! < \log_p (m'+1).$$

The above argument shows that if $(a_n)_{n\in\mathbf{N}}$ is not a sequence such that $a_m = 0$ for all $m > 1$, then p must be a power of a prime number p'. Otherwise, as p is not multiplicatively dependent from any other prime number, $a_m = 0$ for all $m > 1$. Furthermore, the sequence $(a_n)_{n\in\mathbf{N}}$ can have at most one non-zero prime factor, and if it exists, it should be $a_{p'}$. Using automaticity, we can replace p' with p. □

3.4 CMAS satisfying condition C

From this section, we consider only the CMAS with finitely many 0-factors.

In this section, we prove that all CMAS satisfying C can have at most one non-trivial factor, and we do this in several steps.

Proposition 3.4. *Let $(a_n)_{n\in\mathbf{N}}$ be a non-vanishing CMS taking values in the set $G = \{\zeta^r | r \in \mathbf{N}\}$, where ζ is a non-trivial k-th root of unity, having u prime factors $a_{p_1}, a_{p_2}, ...a_{p_u}$; then, there exist $g \in G$ (where $a_{p_1} = g$) and a subword $\overline{w}_u$ appearing periodically in the sequence $(a_n)_{n\in\mathbf{N}}$ such that all its letters are different from g. Furthermore, the period does not have any prime factor other than $p_1, p_2, ..., p_u$.*

Proof. We prove the statement by induction. For $u = 1$, the above statement is trivial. It is easy to check that the sequence $(a_{np_1^{k+1}+p_1^k})_{n\in\mathbf{N}}$ is a constant sequence of 1, the period is p_1^{k+1}, and $g = a_{p_1}$.

Supposing that the statement is true for some $u = n_0$, let us consider the case $u = n_0+1$. We first consider the sequence $(a'_n)_{n\in\mathbf{N}}$ defined by $a'_n = a_{\frac{n}{p_{n_0+1}^{v_{p_{n_0+1}}(n)}}}$, a sequence having n_0 prime factors, where $v_p(n)$ denotes the largest integer r such that $p^r|n$. Using the hypothesis of induction, we obtain a subword $\overline{w}_{n_0}$ satisfying the statement. Let us suppose that the first letter of this subword appears in the sequence $(a'_{m_{n_0}n+l_{n_0}})_{n\in\mathbf{N}}$. We can extract from this sequence a sequence of the form $(a'_{m_{n'_0}n+l_{n_0}})_{n\in\mathbf{N}}$ such that $m_{n'_0} = m_{n_0}\prod_{j=1}^{n_0} p_j^{d_j}$ for some $d_j \in \mathbf{N}^+$ and $v_{p_j}(m_{n'_0}n + l_{n_0} + n_0) = v_{p_j}(l_{n_0} + n_0)$ for all $j \leq n_0$. In this case, the sequence $(a'_{m_{n'_0}n+l_{n_0}+n_0})_{n\in\mathbf{N}}$ is a constant sequence, say, all letters equal C.

Here, we consider two residue classes $N_1(n)$ and $N_2(n)$, separately satisfying the following conditions:

$$m_{n'_0}N_1(n) \equiv -l_{n_0} - n_0 \mod p_{n_0+1}$$
$$m_{n'_0}N_1(n) \not\equiv -l_{n_0} - n_0 \mod p_{n_0+1}^2$$

and

$$m_{n'_0}N_2(n) \equiv -l_{n_0} - n_0 \mod p_{n_0+1}^2$$
$$m_{n'_0}N_2(n) \not\equiv -l_{n_0} - n_0 \mod p_{n_0+1}^3$$

In these two cases, we have $a_{m_{n'_0}N_1(n)+l_{n_0}+n_0} = Ca_{p_{n_0+1}}$ and $a_{m_{n'_0}N_2(n)+l_{n_0}+n_0} = Ca_{p_{n_0+1}}^2$ for all $n \in \mathbf{N}$. Because $a_{p_{n_0+1}} \neq 1$, there is at least one element of $Ca_{p_{n_0+1}}, Ca_{p_{n_0+1}}^2$ not equal to g. If $N_i(n)$ is the associated residue class, then $N_i(n) = p_{n_0+1}^{i+1}n + t$ for all integers n with $t \in \mathbf{N}$, $i = 1$ or 2.

Now, let us choose $m_{n_0+1} = m_{n'_0} p_{n_0+1}^{i+1}$ and $l_{n_0+1} = l_{n_0} + t m_{n'_0}$ so that the sequence $(a'_{m_{n_0+1}n+l_{n_0+1}})_{n\in\mathbf{N}}$ is a subsequence of $(a'_{m_{n_0}n+l_{n_0}})_{n\in\mathbf{N}}$; thus, the subword of length n $\overline{a'_{m_{n_0+1}n+l_{n_0+1}} a'_{m_{n_0+1}n+l_{n_0+1}+1} \dots a'_{m_{n_0+1}n+l_{n_0+1}+n_0-1}}$ is constant, and none of its letters equal g because of the hypothesis of induction. Furthermore, $a_{m_{n_0+1}n+l_{n_0+1}+n_0} = a_{m'_{n_0}N_i(n)+l_{n_0}+n_0}$ is constant and different from g because of the choice of residue class. The properties saying that the prime number p_{n_0+1} is larger than n_0+1 and $p_{n_0+1} | m_{n'_0} N_i(n) + l_{n_0} + n_0$ by construction imply that, for all j such that $0 \leq j \leq n_0 - 1$, $p_{n_0+1} \nmid m_{n_0+1}n + l_{n_0+1} + j$. Therefore, we conclude that, for all $n, j \in \mathbf{N}$ such that $0 \leq j \leq n_0 - 1$, $v_{p_{n_0+1}}(m_{n_0+1}n + l_{n_0+1} + j) = 0$. This means that the subword $\overline{a_{m_{n_0+1}n+l_{n_0+1}} a_{m_{n_0+1}n+l_{n_0+1}+1} \dots a_{m_{n_0+1}n+l_{n_0+1}+n_0}}$ is a subword of length $n_0 + 1$ independent of n and that none of its letters equals g; moreover, m_{n_0+1} does not have any prime factor other than $p_1, p_2 \dots p_{n_0}$. □

Proposition 3.5. *Let $(a_n)_{n\in\mathbf{N}}$ be a non-vanishing CMS defined on a finite set G satisfying condition $\mathcal{C}$, and let $(a'_n)_{n\in\mathbf{N}}$ be another CMS generated by the first r prime factors of $(a_n)_{n\in\mathbf{N}}$, say, $a_{p_1}, a_{p_2}, \dots, a_{p_r}$. If there is a subword $\overline{w}_r$ appearing periodically in $(a'_n)_{n\in\mathbf{N}}$ and if the period does not have any prime factors other than $p_1, p_2, \dots, p_r$, then this subword appears at least once in $(a_n)_{n\in\mathbf{N}}$.*

Proof. Let us denote by $p_1, p_2 \dots$ the sequence of prime numbers such that $a_{p_i} \neq 1$. Supposing that the first letter of the subword $\overline{w}_r$ belongs to the sequence $(a'_{m_r n + l_r})_{n\in\mathbf{N}}$ for some $m_r \in \mathbf{N}, l_r \in \mathbf{N}$, by hypothesis, m_r does not have any prime factors other than $p_1, p_2, \dots, p_r$. Thus, the total number of such subwords in the sequence $(a_n)_{n\in\mathbf{N}}$ can be bounded by the inequality:

$$\begin{aligned} &\#\{a_k | k \leq n, \overline{a_k, a_{k+1}, \dots, a_{k+r-1}} = \overline{w}_r\} \\ &\geq \#\left\{a_k | k \leq n, k = m_r k' + l_r, k' \in \mathbf{N}; p_i \nmid k + j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, i > r\right\} \end{aligned} \tag{3.1}$$

Let us consider the sequence defined by $N(t) = \prod_{j=1}^{t} p_{r+j}$; we have

$$\begin{aligned} &\#\left\{a_k | k \leq N(t) m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k + j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} \\ &= \prod_{j=1}^{t} (p_{r+j} - r) \end{aligned} \tag{3.2}$$

This equality holds because of the Chinese reminder theorem and the fact that $p_{r+j} \nmid m_r$ and $p_{r+j} > r$ for all $j \geq 1$.

Therefore, we have

$$
\begin{aligned}
&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, i > r\right\} \\
>&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} \\
&- \#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \mid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, i > r+t\right\} \\
>&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} \\
&- \sum_{i > r+t} \#\left\{a_k | k \leq N(i)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \mid k+j, \forall j \text{ with } 0 \leq j \leq r-1\right\} \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r \sum_{i > r+t, p_i < N(t)+r} [\frac{N(t)}{p_i}] \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r \sum_{i > r+t, p_i < N(t)+r} \frac{N(t)}{p_i} - r\pi(N(t)+r).
\end{aligned}
\tag{3.3}
$$

where $[a]$ represents the smallest integer larger than a and π is the prime counting function. However,

$$
\prod_{j=1}^{t}(p_{r+j} - r) = \prod_{j=1}^{t} \frac{p_{r+j} - r}{p_{r+j}} N(t) \geq \prod_{j=1}^{\infty} \frac{p_{r+j} - r}{p_{r+j}} N(t). \tag{3.4}
$$

The last formula can be approximated as $\prod_{j=1}^{\infty} \frac{p_{r+j}-r}{p_{r+j}} = \exp(\sum_{j=1}^{\infty} \log(\frac{p_{r+j}-r}{p_{r+j}})) = \exp(-\Theta(\sum_{j=1}^{\infty} \frac{r}{p_{r+j}}))$, and the last equality holds because $\log(1-x) \sim x$ when x is small. Because of C, the above quantity does not diverge to 0; we conclude that, if t is sufficiently large, there exists a c with $0 < c < 1$ such that $\prod_{j=1}^{t}(p_{r+j} - r) > cN(t)$.

On the other hand, we remark that for all $i > r+t$, $p_i^t > \prod_{j=1}^{t} p_{r+j} = N(t)$; thus, $p_i > N(t)^{\frac{1}{t}}$

$$
\sum_{i > r+t, p_i < N(t)+r} \frac{N(t)}{p_i} < N(t) \sum_{N(t)^{\frac{1}{t}} < p < N(t)+r} \frac{1}{p}. \tag{3.5}
$$

The term $N(t)^{\frac{1}{t}}$ can be bounded by

$$
N(t)^{\frac{1}{t}} = (\prod_{j=1}^{t} p_{r+j})^{\frac{1}{t}} \geq \frac{t}{\sum_{j=1}^{t} \frac{1}{p_{r+j}}} > \frac{t}{\sum_{j=1}^{t} \frac{1}{q_j}}. \tag{3.6}
$$

where q_j is the j-th prime number in $\mathbf{N}$. For any $x \in \mathbf{N}$, $\#\{p_i | p_i \leq x\} \sim \frac{x}{\log(x)}$ and $\sum_{p_i \leq x} \frac{1}{p_i} \sim \log\log(x)$; thus, $N(t)^{\frac{1}{t}}$ tends to infinity when t tends to infinity. Because of C, we can conclude that there exists some $t_0 \in \mathbf{N}$ such that, for all $t > t_0$, $\sum_{N(t)^{\frac{1}{t}} < p < N(t)+r} \frac{1}{p} < \frac{1}{2r}c$.

To conclude, for all $t > t_0$,

$$\begin{aligned}
&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k' + l_r, k' \in \mathbf{N}; k + j \nmid p_i, \forall (i,j) \text{ with } 0 \leq j \leq r-1, \forall i > r\right\} \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r\sum_{k>r+t}\frac{N(t)}{p_k} - r\pi(N(t)+r) \\
>&cN(t) - \frac{1}{2}cN(t) - r\pi(N(t)+r).
\end{aligned} \tag{3.7}$$

When t tends to infinity, the set $\#\left\{a_k | k \leq n, \overline{a_k, a_{k+1}, ..., a_{k+r-1}} = \overline{w}_r\right\}$ is not empty.

□

Proposition 3.6. *Let $(a_n)_{n\in\mathbf{N}}$ be a p-CMAS, vanishing or not, satisfying condition C. Then, there exists at most one prime number k such that $a_k \neq 1$ or 0.*

Proof. Suppose that the sequence $(a_n)_{n\in\mathbf{N}}$ has infinitely many prime factors not equal to 0 or 1. Let us consider first the sequence $(a'_n)_{n\in\mathbf{N}}$ defined as follows:

$$a'_n = a_{\frac{n}{\prod_{p_i\in\mathbf{Z}} p_i^{v_{p_i}(n)}}},$$

where $\mathbf{Z} = \{p | p \in \mathbf{P}, a_p = 0\}$; because of Proposition 3.3, this set is finite.

Using Propositions 3.4 and 3.5, there exists a subword of length $p^{2q!}$, say, $\overline{v}_{p^{2q!}}$, appearing in $(a'_n)_{n\in\mathbf{N}}$ such that none of its letters equal $g = a'_{p_1} = a_{p_1}$, where q is the number of states of the automaton generating $(a_n)_{n\in\mathbf{N}}$. Then, by construction, there is a subword of the same length, say, $\overline{w}_{p^{2q!}}$, appearing at the same position on the sequence $(a_n)_{n\in\mathbf{N}}$ such that none of its letters equal g. Extracting a subword $\overline{w'}_{p^{q!}}$ contained in $\overline{w}_{p^{2q!}}$ of the form $\overline{a_{up^{q!}}a_{up^{q!}+2}...a_{(u+1)p^{q!}-1}}$ for some $u \in \mathbf{N}$ and using Proposition 3.1, we have, for every y such that $y \geq 1$ and every m such that $0 \leq m \leq p^{yq!} - 1$, $a_{up^{yq!}+m} \neq g$. In particular,

$$\lim_{y\to\infty}\frac{1}{p^{yq!}}\#\left\{a_s = g | up^{yq!} \leq s < (u+1)p^{yq!} - 1\right\} = 0.$$

which contradicts the fact that g has a non-zero natural density proved by Proposition 3.2.

Therefore, we have proven that the sequence $(a_n)_{n\in\mathbf{N}}$ must have finitely many prime factors. However, Corollary 2 of [27] proves that, in this case, the sequence $(a_n)_{n\in\mathbf{N}}$ can have at most one prime k such that $a_k \neq 1$ or 0.

□

3.5 Classification of CMAS

In this section, we will prove that a CMAS is either strongly aperiodic or a Dirichlet-like sequence.

Definition A sequence $(a_n)_{n\in\mathbf{N}}$ is said to be aperiodic if and only if, for any pair of integers (s, r), we have

$$\lim_{N\to\infty} \frac{\sum_{i=0}^{N} a_{si+r}}{N} = 0.$$

Definition Let $\mathcal{M}$ be the set of completely multiplicative functions. Let $\mathbf{D} : \mathcal{M} \times \mathcal{M} \times \mathbf{N} \to [0, \infty]$ be given by

$$\mathbf{D}(f, g, N)^2 = \sum_{p\in\mathbf{P}\cap[N]} \frac{1 - Re(f(p)\overline{g(p)})}{p}$$

and $M : \mathcal{M} \times \mathbf{N} \to [0, \infty)$ be given by

$$M(f, \mathbf{N}) = \min_{|t|\leq N} \mathbf{D}(f, n^{it}, N)^2$$

A sequence $(a_n)_{n\in\mathbf{N}}$ is said to be strongly aperiodic if and only if $M(f\chi, N) \to \infty$ as $N \to \infty$ for every Dirichlet character χ.

Definition A sequence $(a_n)_{n\in\mathbf{N}}$ is said to be (trivial) Dirichlet-like if and only if there exists a (trivial) Dirichlet character $X(n)_{n\in\mathbf{N}}$ such that there exists at most one prime number p satisfying $a_p \neq X(p)$.

Proposition 3.7. *Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS; then, either there exists a Dirichlet character $(X(n))_{n\in\mathbf{N}}$ such that the sequence $(a_n X(n))_{n\in\mathbf{N}}$ is a trivial Dirichlet-like character or it is strongly aperiodic.*

Proof. First, it is easy to check that there is an integer r such that a_p is a r-th root of unity for all but finitely many primes p (see Lemma 1 [42]). If $(a_n)_{n\in\mathbf{N}}$ is not strongly aperiodic, then because of Proposition 6.1 in [21], there exists a Dirichlet character $(X(n))_{n\in\mathbf{N}}$ such that

$$\lim_{N\to\infty} \mathbf{D}(a, X, N) < \infty \ (*).$$

However, the sequence $(a_n\overline{X(n)})_{n\in\mathbf{N}}$ is also CMAS and satisfies condition $\mathcal{C}$; the last fact is from $(*)$. Because of Proposition 3.6, $(a_n\overline{X(n)})_{n\in\mathbf{N}}$ is a trivial Dirichlet-like character. □

Proposition 3.8. *Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS and $X_t(n)_{n\in\mathbf{N}}$ be a Dirichlet character (mod t). If the sequence $(a_n X_t(n))_{n\in\mathbf{N}}$ is the trivial Dirichlet-like character (mod t), then $(a_n)_{n\in\mathbf{N}}$ is either a Dirichlet character (mod t) or a Dirichlet-like character $a_n = \epsilon^{v_p(n)} X(\frac{n}{p^{v_p(n)}})$, where p is a prime divisor of t and ϵ is a root of unity.*

Proof. Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS satisfying the above hypothesis; then, all possibilities for such $(a_n)_{n\in\mathbf{N}}$ are the sequences of the form

$$a_n = \prod_{i=1}^{m} \epsilon_i^{v_{p_i}(n)} X\left(\frac{n}{\prod_{i=1}^{m} p_i^{v_{p_i}(n)}}\right),$$

for each n, where ϵ_i are all non-zero complex numbers and p_i are all prime factors of t.

Let us consider the Dirichlet sequence $f(s)$ associated with the sequence $(a_n)_{n\in\mathbf{N}}$, which can be written as

$$f(s) = L(s, X_t)\prod_{i=1}^{m}\frac{1-\frac{1}{p_i^s}}{1-\frac{a_{p_i}}{p_i^s}}.$$

Therefore, all the poles of $f(s)$ can be found on

$$s = \frac{\log a_{p_i} + 2in\pi}{\log p_i},$$

for all i such that $1 \leq i \leq m$ and $n \in \mathbf{Z}$.

However, if $(a_n)_{n\in\mathbf{N}}$ is a k-automatic sequence for some integer k, then the poles should be located at points

$$s = \frac{\log\lambda}{\log k} + \frac{2im\pi}{\log k} - l + 1,$$

where λ is any eigenvalue of a certain matrix defined from the sequence $(\chi_n)_{n\in\mathbf{N}}$, and $m \in \mathbf{Z}, l \in \mathbf{N}$, and log is a branch of the complex logarithm [2]. By comparing the two sets of possible locations of poles for the same function, we can see that there is at most one $a_{p_i} \neq 0$. □

3.6 Conclusion

In this section, we conclude this article by proving that strongly aperiodic CMAS does not exist. To do so, we recall the definition of the block complexity of sequences.

Definition Let $(a_n)_{n\in\mathbf{N}}$ be a sequence. The block complexity of $(a_n)_{n\in\mathbf{N}}$ is a sequence, which will be denoted by $(p(k))_{k\in\mathbf{N}}$, such that $p(k)$ is the number of subwords of length k that occur (as consecutive values) in $(a_n)_{n\in\mathbf{N}}$

Proposition 3.9. *If $(a_n)_{n\in\mathbf{N}}$ is a CMAS, then it is not strongly aperiodic.*

Proof. From Theorem 2 in ([22]) and the remark following this theorem, the block complexity of the sequence $(a_n)_{n\in\mathbf{N}}$ should satisfy the property that $\lim_{n\to\infty}\frac{p(n)}{n} = \infty$, which contradicts the fact that the block complexity of an automatic sequence is bounded by a linear function [13]. Therefore, the non-existence of strongly aperiodic CMAS is proved. □

Theorem 3.1. *Let* $(a_n)_{n\in\mathbf{N}}$ *be a CMAS; then, it can be written in the following form:*
-either there is at most one prime p *such that* $a_p \neq 0$ *and* $a_q = 0$ *for all other primes* q
-or $a_n = \epsilon^{v_p(n)} X(\frac{n}{p^{v_p(n)}})$, *where* $(X(n))_{n\in\mathbf{N}}$ *is a Dirichlet character.*

3.7 Acknowledgement

We found some results in the recent literature on similar topics that have applications to the classification of CMAS. In [30], the authors proved that all continuous observables in a substitutional dynamical system (X_θ, S) are orthogonal to any bounded, aperiodic, multiplicative function, where θ represents a primitive uniform substitution and S is the shift operator. As an application, all multiplicative and automatic sequences produced by primitive automata are Weyl rationally almost periodic. We remark that a sequence $(b_n)_{n\in\mathbf{N}}$ is called Weyl rationally almost periodic if it can be approximated by periodic sequences in same alphabet in the pseudo-metric

$$d_W(a,b) = \limsup_{N\to\infty} \sup_{l\geq 1} \frac{1}{N} |\{l \leq n < l+N : a(n) \neq b(n)\}|.$$

This result could probably be generated in the non-primitive case.

In [29], the authors considered general multiplicative functions with the condition $\liminf_{N\to\infty} |b_{n+1} - b_n| > 0$. They proved that if $(b_n)_{n\in\mathbf{N}}$ is a completely multiplicative sequence, then most primes, at a fixed power, give the same values as a Dirichlet character.

Chapter 4

Automatic sequences defined by Theta functions and some infinite products

4.1 Introduction

Let $p(x) \in C(x)$ be a rational function satisfying the condition $p(0) = 1$ and q an integer larger than 1, in this article we will consider the expansion in power series of the infinite product

$$f(x) = \prod_{s=0}^{\infty} p(x^{q^s}) = \sum_{i=0}^{\infty} c_i x^i,$$

and study when the sequence $(c_i)_{i\in\mathbf{N}}$ is q-automatic. This topic has been studied by many authors, such as [17], [16] and [11] , using analytical approach, here we want to review this topic by a basic algebraic approach.

The main result is that for given integers $q \geq 2$ and $d \geq 0$, there exist finitely many polynomials of degree d defined over the field of rational numbers $\mathbf{Q}$, such that $f(x) = \prod_{s=0}^{\infty} p(x^{q^s}) = \sum_{i=0}^{\infty} c_i x^i$ is a q-automatic power series.

4.2 Definitions and generality

Definition Let $(a_n)_{n\in\mathbf{N}}$ be a sequence, we say it is q-automatic if the set

$$Ker((a_i)_{i\in\mathbf{N}}) = \left\{(a_{q^l n+b})_{n\in\mathbf{N}} | l \in \mathbf{N}, 0 \leq b < q^l\right\}$$

is finite. This set will be called the q-kernel of $(a_n)_{n\in\mathbf{N}}$.

For every couple of integers (l, b) satisfying $l \in \mathbf{N}, 0 \leq b < q^l$, let us define a relation $R_{l,b}$ over the sequence space: we say $R_{l,b}((a_n)_{n\in\mathbf{N}}, (b_n)_{n\in\mathbf{N}})$ if and only if

$$\forall n \in \mathbf{Z}, b_n = a_{q^l n+b}.$$

Definition Let $\sum_{i=0}^{\infty} a_i x^i$ be a power series, we say it is q-automatic if the sequence of coefficients $(a_n)_{n\in\mathbf{N}}$ is q-automatic.

Similarly we define operators $O_{l,b}$ over the space of power series:

$$O_{l,b}(\sum_{n=0}^{\infty} a_n x^n) = \sum_{n=0}^{\infty} a_{q^l n+b} x^n.$$

Now let us consider a detailed version of a well-known theorem, see, for example, [4].

Proposition 4.1. *let $f \in F((x))$ be a k-automatic power series, then there exist polynomials $a_0(x), a_1(x), ..., a_m(x) \in F[x]$ with $a_0(x)a_m(x) \neq 0$ such that*

$$\sum_{i=0}^{m} a_i(x) f(x^{k^i}) = 0.$$

Furthermore, the coefficients of $a_0(x), a_1(x), ..., a_x(t)$ depend only on $R_{l,b}$ relations over the q-kernel of the sequence of the coefficients of f.

Proof. Let B denote the k-kernel of the sequence of coefficients of f, and N denote the cardinal of B. We can then associate each element in B with a power series by

$$(a_n)_{n\in\mathbf{N}} \to \sum_{n=0}^{\infty} a_n x^n.$$

Let B' denote the image of B by the previous map. For each power series in B', we have

$$\sum_{i=0}^{\infty} a_i x^i = \sum_{i=0}^{k-1} x^i (\sum_{j=0}^{\infty} a_{kj+i} x^{kj}).$$

Remarking that if the sequence $(a_n)_{n\in\mathbf{N}}$ is in B, then $(a_{kn+j})_{n\in\mathbf{N}}$ is also in B, for $j = 0, 1, ..., k-1$. If we write

$$\sum_{i=0}^{\infty} a_i x^i = \sum_{(b_n)_{n\in\mathbf{N}} \in B'} c_b \sum_{i=0}^{\infty} b_i x^{ki},$$

Then

$$c_b = \begin{cases} x^i, \text{ if } R_{1,i}((a_n)_{n\in\mathbf{N}}, (b_n)_{n\in\mathbf{N}}) \\ 0, \text{ otherwise.} \end{cases}$$

Particularly, we can do the same thing for $f(x), f(x^k), ..., f(x^{k^N})$:

$$\begin{cases} f(x) = \sum_{(b_n)_{n\in\mathbf{N}}\in B} c_b^1 \sum_{i=0}^{\infty} b_i x^{k^{N+1}i}, \\ f(x^k) = \sum_{(b_n)_{n\in\mathbf{N}}\in B} c_b^2 \sum_{i=0}^{\infty} b_i x^{k^{N+1}i}, \\ ... \\ f(x^{k^N}) = \sum_{(b_n)_{n\in\mathbf{N}}\in B} c_b^N \sum_{i=0}^{\infty} b_i x^{k^{N+1}i}; \end{cases}$$

with c_b^j defined only by $R_{l,b}$ relations. But as the cardinal of B' is N, the linear forms at the right-hand side of above equalities are linearly dependent. As a result, if we neglect the linear dependence between elements in B', we can have a linear dependence between $f(x), f(x^k), ..., f(x^{k^N})$ such that the coefficients depend only on c_b^j. So these coefficients depend only on $R_{l,b}$ relations. □

Here we make this proposition precise by some examples:

Example Let us consider a periodic sequence

$$a, b, a, b, a, b, a, b...$$

which is 2-automatic.

Now let us write down the associated power sequence $F(x) = a + bx + ax^2 + bx^3 + ...$ and two other sequences $A(x) = a + ax + ax^2 + ax^3...$, $B(x) = b + bx + bx^2 + bx^3...$ with constant coefficients.

So

$$F(x) = A(x^2) + xB(x^2)$$

$$A(x) = (1 + x)A(x^2)$$

$$B(x) = (1 + x)B(x^2)$$

so we have the following dependence:

$$F(x) = (1 + x^2)(1 + x^4)A(x^8) + x(1 + x^2)(1 + x^4)B(x^8)$$

$$F(x^2) = (1 + x^4)A(x^8) + x^2(1 + x^4)B(x^8)$$

$$F(x^4) = A(x^8) + x^4B(x^8)$$

$F(x)$ satisfies the functional equation

$$(x^8-x^6+x^4-x^2)((1+x^2)F(x^2)-F(x)) = (x^4-x^3+x^2-x)(1+x^4)((1+x^4)F(x^4)-F(x^2))$$

This functional equation does not depend on the values of a and b.

Example Let us consider the Thue-Morse sequence

$$a, b, b, a, b, a, a, b, b, a, a, b, a, b, b, a...$$

which is 2-automatic.

Now let us write down the associated power sequence $F(x) = a + bx + bx^2 + ax^3 + ...$ and another sequence $G(x) = b + ax + ax^2 + bx^3...$, by changing a to b and b to a:

So

$$F(x) = F(x^2) + xG(x^2)$$

and

$$G(x) = G(x^2) + xF(x^2)$$

so we have the following dependence:

$$G(x^2) = G(x^4) + x^2F(x^4)$$

$$x^2G(x^4) = F(x^2) - F(x^4)$$

$$x^2G(x^2) = x^2G(x^4) + x^4F(x^4)$$

$F(x)$ satisfies the functional equation

$$(x^4 - 1)F(x^4) + (1 + x)F(x^2) - xF(x) = 0$$

This functional equation does not depend on the values of a and b.

Proposition 4.2. *For a given functional equation $F : \sum_{s=0}^{m} a_s(t)f(t^{k^s}) = 0$, there exist finitely many polynomials $p_1, p_2, ..., p_r$ with $p_i(0) = 1, \forall i \in [0, r]$, such that the associated theta functions $G_r(x) = \prod_{s=0}^{\infty} p_r(x^{q^s})$ satisfying equation F.*

Proof. If $p(x)$ is a such polynomial satisfying $p(0) = 1$. Let us denote by $G(x)$ the associated power series. By hypothesis, it satisfies the functional equation F:

$$\sum_{s=0}^{m} a_s(x)G(x^{q^s}) = 0.$$

On the other hand, the power series G satisfies another functional equation:

$$G(x) = p(x)G(x^q).$$

Plugging the second equation into the first one, we get

$$\sum_{s=0}^{m} a_s(x) \prod_{r=s}^{m} p(x^{q^{m-r}}) = 0.$$

An observation is that all terms in the sum contain a factor $p(x^{q^{m-1}})$ except the last one. So we have

$$p(x^{q^{m-1}})|a_m(x)$$

with $p(0) = 1$, so there are finitely many choices for $p(x)$. □

Proposition 4.3. *For a fixed number k, there are finitely many polynomials $p_1, p_2, ..., p_r$ such that the theta functions $G_j(x) = \prod_{s=0}^{\infty} p_j(x^{q^s})$ are q-automatic and the sizes of their q-kernels are bounded by k.*

Proof. Fixing the size of the q-kernel, we fix the number of possibilities of $R_{l,b}$ relations, so the possible functional equations, and we conclude by Proposition 4.2. □

4.3 Infinite product of polynomials

Let $p = \sum_{i=1}^{n} a_i x^i$ be a polynomial with coefficients in $\mathbf{C}$ and q be an integer larger than 1. It is known that the coefficients of the power series

$$f(x) = \prod_{s=0}^{\infty} p(x^{q^s})$$

form a q-regular sequence [17], here we want to study when this sequence is q-automatic.

Firstly, let us suppose that the degree of p, noted $deg(p)$, satisfies $q^{k-1} < deg(p) \leq q^k$ for some $k \in \mathbf{N}$ and write

$$f(x) = \prod_{s=0}^{\infty} p(x^{q^s}) = \sum_{i=1}^{\infty} c_i x^i.$$

Then the coefficients c_i satisfy a recurrence relation:

$$c_{nq+r} = \sum_{\substack{0 \leq j \leq q^k) \\ j \equiv r \pmod q}} a_j c_{n+\frac{r-j}{q}} \tag{4.1}$$

for all r such that $0 \leq r \leq q-1$ and $c_n = 0$ for all negative indices.

Lemma 4.1. *The sequences $(c_{qn+i-j})_{n \in \mathbf{N}}$, for all i and j such that $0 \leq i \leq q-1$ and $0 \leq j \leq 2q^k$, can be represented as linear combinations of sequences $\{(c_{n-i})_{n \in \mathbf{N}} | 0 \leq i < 2q^k\}$.*

Proof. Because of the previous equality, we have

$$c_{nq+i-j} = \sum_{\substack{0 \leq s \leq q^k \\ s \equiv i-j \pmod q}} a_s c_{n+\frac{i-j-s}{q}}$$

for all n, i, j defined as above. Now let us check that all sequences appearing on the right-hand side of these equalities are in the set defined in the statement. It is enough to calculate the shifting indices and we have the bounds as follows,

$$-2q^k < -3q^{k-1} \leq \frac{i-j-s}{q} \leq 0$$

which proves the statement. □

Example Let us consider the case where $p(x) = 1 + x + x^2 + x^3 + x^4$ and $q = 2$, the sequence of coefficients of the power series $F(x) = \prod_{s=0}^{\infty} p(x^{q^s})$ is denoted by $(c_n)_{n\in\mathbf{N}}$, so we have

$$p(x) = (1 + x + x^2 + x^3 + x^4)F(x^2)$$

from which we can deduce

$$c_{2n} = c_n + c_{n-1} + c_{n-2},$$

$$c_{2n+1} = c_n + c_{n-1}.$$

Using the above lemma, we get

$$\begin{pmatrix} c_{2n} \\ c_{2n-1} \\ c_{2n-2} \\ c_{2n-3} \\ c_{2n-4} \\ c_{2n-5} \\ c_{2n-6} \\ c_{2n-7} \\ c_{2n-8} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \\ c_{n-4} \\ c_{n-5} \\ c_{n-6} \\ c_{n-7} \\ c_{n-8} \end{pmatrix}$$

and

$$\begin{pmatrix} c_{2n+1} \\ c_{2n} \\ c_{2n-1} \\ c_{2n-2} \\ c_{2n-3} \\ c_{2n-4} \\ c_{2n-5} \\ c_{2n-6} \\ c_{2n-7} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \\ c_{n-4} \\ c_{n-5} \\ c_{n-6} \\ c_{n-7} \\ c_{n-8} \end{pmatrix}$$

Because of the previous fact, we can introduce some transition matrices: for all integers r

such that $0 \leq r \leq q-1$ let us define Γ_r as a square matrix of size $2q^k+1$ satisfying

$$\Gamma_r \begin{pmatrix} c_n \\ c_{n-1} \\ \dots \\ c_{n-2q^k} \end{pmatrix} = \begin{pmatrix} c_{qn+r} \\ c_{qn+r-1} \\ \dots \\ c_{qn+r-2q^k} \end{pmatrix}$$

for all $n \in \mathbf{N}$.

Let us denote by G the semi-group generated by all Γ_r and multiplication.

Proposition 4.4. *$a \in \{c_n | n \in \mathbf{N}\}$ if and only if there exists a matrix $g \in G$ such that a is the first element in the first row of the matrix g, in other words, $a = g(1,1)$. Furthermore, $(c_n)_{n\in\mathbf{N}}$ is automatic if and only if G is a finite semi-group.*

Proof. The first part of this proposition is trivial, for any $r \in \mathbf{N}$, let us consider its q-ary expansion $r = \overline{s_{k_1}s_{k_1-1}...s_0}$. Using Lemma 4.1, we have

$$\begin{pmatrix} c_r \\ c_{r-1} \\ \dots \\ c_{r-2q^k} \end{pmatrix} = \Gamma_{s_{k_1}}\Gamma_{s_{k_1-1}}..\Gamma_{s_0} \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix},$$

which proves the first part of the statement.

For the second part, let us define maps γ_r for all integers r by $\gamma_r(n) = q(q(...q(q(n)+s_0)...)+s_{k_1-1})+s_{k_1}$ for all $n \in \mathbf{N}$ if $r = \overline{s_{k_1}s_{k_1-1}...s_0}$. Then there is an equality for all r:

$$\begin{pmatrix} c_{\gamma_r(0)} & c_{\gamma_r(1)} & \cdots & c_{\gamma_r(2q^k)} \\ c_{\gamma_r(0)-1} & c_{\gamma_r(1)-1} & \cdots & c_{\gamma_r(2q^k)-2q^k} \\ \dots & & & \\ c_{\gamma_r(0)-2q^k} & c_{\gamma_r(1)-2q^k} & \cdots & c_{\gamma_r(2q^k)-2q^k} \end{pmatrix} = \Gamma_{s_{k_1}}\Gamma_{s_{k_1-1}}..\Gamma_{s_0} \begin{pmatrix} a_0 & a_1 & \dots & a_{2q^k} \\ 0 & a_0 & \dots & a_{2q^k-1} \\ \dots & & & \\ 0 & 0 & \dots & a_0 \end{pmatrix}.$$

But the the last matrix in the above equality is constant and invertible, so each element of a matrix $g \in G$ is a finite linear composition of elements in the sequence $(c_n)_{n\in\mathbf{N}}$, so the finiteness of elements in $(c_n)_{n\in\mathbf{N}}$ is equivalent to the finiteness of elements in G. And using the fact that $(c_n)_{n\in\mathbf{N}}$ is an automatic sequence, we conclude the statement. $\square$

Proposition 4.5. *For given integers $q \geq 2$ and $d \geq 0$, there exist finitely many polynomials of degree d defined over the field of rational numbers $\mathbf{Q}$, such that $\prod_{s=0}^{\infty} p(x^{q^s}) = \sum_{i=1}^{\infty} c_i x^i$ is a q-automatic power series.*

Proof. Suppose that the sequence $(c_n)_{n\in\mathbf{N}}$ generated by $\prod_{s=0}^{\infty} p(x^{q^s}) = \sum_{i=1}^{\infty} c_i x^i$ is automatic. Let us consider a sequence of matrices $(\Gamma_n)_{n\in\mathbf{N}}$, such that Γ_i are defined as above for $i = 0, 1, .., q-1$ and $\Gamma_{qi+j} = \Gamma_i \Gamma_j$ for all $i \geq 1$ and $j = 0, 1, ..., q-1$.

It is easy to see that this matrix sequence is automatic because G is finite. And also the automata of this matrix sequence is the same as the one of $(c_n)_{n\in\mathbf{N}}$, because c_n is exactly the element at the position $(1,1)$ of the matrix Γ_n. To conclude the statement, we have to prove two things: firstly the number of automata generating the sequences $(\Gamma_n)_{n\in\mathbf{N}}$ is finite, secondly, the output functions for each automaton are also finite.

For the first point, it is enough to show that $|G|$ is bounded by a function depending only on d and q, which is proved by Theorem 1.3 of [33]. It says that given naturals n and k, there exist, up to semi-group isomorphism, only a finite number of finite sub-semi-groups of $M_n(F)$ generated by at most k elements.

For the second point, it is a consequence of Proposition 4.3. □

Proposition 4.6. *Let f be a polynomial satisfying the hypothesis in Proposition 4.5, then all its coefficients belong to* $\mathbf{Z}$.

Proof. Let us denote by d the degree of f and write down all coefficients of f in the form $a_i = \frac{p_i}{q_i}$ such that $(p_i, q_i) = 1$, and similarly for all coefficient of F, let us write down $c_i = \frac{r_i}{t_i}$ with $(r_i, t_i) = 1$. If there are some coefficients of f which are rational numbers but not integers, then there exist a prime p and two integers d_1 and d_2 satisfying :

$$d_1 = \max\left\{t | t \in \mathbf{N}, \exists q_i, p^t | q_i\right\}$$

and

$$d_2 = \max\left\{t | t \in \mathbf{N}, \exists t_i, p^t | t_i\right\}$$

with $d_1 > 0, d_2 > 0$. In fact, because of the hypothesis, there exists $a_i = \frac{p_i}{q_i}$ with $q_i \neq 1$. So there exists a prime p such that $p | q_i$, thus $d_1 \neq 0$. Let us suppose $a_j = \frac{p_j}{q_j}$ with the smallest index such that $p^{d_1} | q_j$. Now let us check

$$c_j = a_j + \sum_{qk+s=j, k>0} a_k c_s.$$

If $c_j = \frac{r_j}{t_j}$ with $p | t_j$ then $d_2 \geq 1$; otherwise, there are some a_k, c_j such that $p^{d_1} | q_k t_j$, but with the assumption of smallest index, $p^{d_1} \nmid q_k$, so $p | t_j$ thus $d_2 \geq 1$.

Let l_1 be the smallest index such that $p^{d_1} | q_{l_1}$ and similarly let l_2 be the smallest index such that $p^{d_2} | s_{l_2}$. Now let us consider the coefficient $c_{l_2 q + l_1}$, which can be calculated as

$$c_{l_2 q + l_1} = \sum_{0 \leq i \leq d, qj+i = l_2 q + l_1} a_i c_j.$$

Let us consider the sum at the right-hand side, for any couple of (a_i, c_j), if $i < t_1$, then $p^{d_1} \nmid q_i$, the maximality of d_2 leads to $p^{d_1+d_2} \nmid q_i t_j$; similarly, if $i > t_1$, then $j < t_2$ thus

$p^{d_2} \nmid t_j$, so that $p^{d_1+d_2} \nmid q_i t_j$; but if $i = t_1$, then $j = t_2$, so $p^{d_1}|q_i$ and $p^{d_2}|t_i$. As a result, $p^{d_1+d_2}|c_{t_2 q+t_1}$, contradicts the maximality of d_2. □

4.4 Rational functions generated by infinite products

Here we consider the following question: for a given polynomial p and an integer q, when does $F(x) = \prod_{s=0}^{\infty} p(x^{q^s})$ equal a rational function. This question has already been studies in [16] when restricting the polynomial to the cyclotomic case, this section can be considered as a generalization of the previous work.

Proposition 4.7. *Let p be a polynomial taking coefficients over $\mathbf{C}$ and q be an integer larger than 1, then there is an equivalence between:*

(1) $\prod_{s=0}^{\infty} p(x^{q^s})$ is a rational function.

(2) there exists a polynomial $Q(x)$ such that $p(x) = \frac{Q(x^q)}{Q(x)}$ and all roots of $Q(x)$ are roots of unity, if δ is a root of $Q(x)$ then δ^{q^t} is a root of Q for all $t \in \mathbf{N}$.

Proof. (2) implies (1) is straightforward, let us check (1) implies (2).

Let $F(x) = \prod_{s=0}^{\infty} p(x^{q^s})$ be a rational function, say $F(x) = \frac{P(x)}{Q(x)}$, where $P(x)$ and $Q(x)$ are coprime, using the functional equation $F(x) = p(x)F(x^q)$, we get

$$\frac{P(x)Q(x^q)}{P(x^q)Q(x)} = p(x).$$

As $deg(p(x)) > 0$, so that $deg(Q(x)) > deg(P(x))$, and $P(x^q)|P(x)Q(x^q)$ if $deg(P(x)) > 0$, then $P(x^q)$ and $Q(x^q)$ should have at least one common root, which contradicts that $P(x)$ and $Q(x)$ are coprime, so we have

$$F(x) = \frac{1}{Q(x)}$$

and

$$p(x) = \frac{Q(x^q)}{Q(x)}$$

Now let us study the roots of $Q(x)$, let us suppose $0 \leq |r_1| \leq |r_2| \leq ... \leq |r_m|$ where r_i are the roots of $Q(x)$ and $|r_i|$ is the modulus of r_i. Firstly $|r_m|$ can not be too large, if $|r_m| > 1$ then each root of $Q(x^q)$ should have a modulus strictly smaller than $|r_m|$, on the other hand $Q(x)|Q(x^q)$, which is impossible. For the same reason, $|r_1|$ can not be a real number between 0 and 1. So $|r_i|$ are either 0 or 1, but if $x|Q(x)$, the infinite product of $p(x)$ will not converge, so $|r_i| = 1$ for all roots of $Q(x)$. Using once more $Q(x)|Q(x^q)$, if δ is a root of $Q(x)$ then it is a root of $Q(x^q)$ which implies δ^q is a root of $Q(x)$, we can do it recursively and we obtain δ^{q^t} is a root of Q for all $t \in \mathbf{N}$, as a corollary, δ can only be a root of unity. So we prove (2) using (1).

□

4.5 Infinite product of inverse of polynomials

In this section, we consider the power sequence defined as follows:

$$F(x) = \prod_{s=0}^{\infty} \frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i,$$

where q is an integer larger than 1 and $p = \sum_{i=0}^{n} b_i x^i$ is a polynomial such that $p(0) = 1$ defined as before.
Such a sequence satisfies the functional equation

$$F(x) = \frac{1}{p(x)} F(x^q).$$

If we write $\frac{1}{p(x)} = \sum_{i=0}^{\infty} a_i x^i$, then

$$c_{qn+i} = \sum_{j=0}^{n} a_{qj+i} c_{n-j},$$

for all $n \in \mathbf{N}$ and i such that $0 \leq i \leq q-1$.

Proposition 4.8. *If the coefficients of the power series $F(x) = \prod_{s=0}^{\infty} \frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i$ take finitely many values in $\mathbf{C}$, then the roots of f are all of modulus 1.*

Proof. Firstly, let us prove that the moduli of all roots of p are not smaller than 1. Otherwise, let us chose one of those which have smallest modulus, say α, because of the above definition, we can conclude that

$$p(\alpha^k) \neq 0$$

for all k larger than 1.

Let us consider the equality,

$$\prod_{s=0}^{\infty} \frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i,$$

the right-hand side converges when x tends to α while the left-hand side diverges, in fact $\prod_{s=1}^{\infty} \frac{1}{p(\alpha^{q^s})}$ converges to a non-zero value because

$$\log(\prod_{s=1}^{\infty} \frac{1}{p(\alpha^{q^s})}) = -\sum_{s=1}^{\infty} \log(p(\alpha^{q^s}))$$

which converges, however, $\frac{1}{p(x^{q^s})}$ has a pole at $x = \alpha$.

Secondly, let us prove that the moduli of all roots of p are not larger than 1. Otherwise, let us chose one of them, say β, and an integer t such that $|\beta|^{q^t} > |a|/|b| + 1$, where $|a|$ is the largest modulus of the sequence $(c_i)_{i\in\mathbf{N}}$ and $|b|$ is the smallest non-zero modulus of this sequence. Now consider the following series

$$\frac{1}{1-\beta}\prod_{s=t}^{\infty}\frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} d_i x^i.$$

It is easy to see that $\{d_i|i \in \mathbf{N}\}$ is finite, because such a series can be obtained by multiplying a polynomial to $F(x)$, but on the other hand, we have the inequality,

$$|d_{q^t i}| = |\sum_{j=0}^{i} \beta^{q^t j} c_{q^t(i-j)}| \geq -|a|\sum_{j=0}^{i-1}|\beta^{q^t j}| + |b||\beta^{q^t i}| > 0$$

which diverges. This contradicts the fact that $\{d_i|i \in \mathbf{N}\}$ is finite. In conclusion, the roots of f are all of modulus 1.

□

Proposition 4.9. *If the power series $F(x) = \prod_{s=0}^{\infty}\frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i$ is a q-regular sequence, then the roots of p are all roots of unity, furthermore, the order of each root is multiple of q.*

Proof. If $F(x) = \prod_{s=0}^{\infty}\frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i$ is a q-regular sequence, then $F'(x) = \sum_{i=1}^{\infty} c_i i x^{i-1}$ is also q-regular. On the other hand, we know $\frac{1}{F(x)} = \prod_{s=0}^{\infty} p(x^{q^s})$ is q-regular, so

$$\frac{F'(x)}{F(x)} = (\log F(x))'$$

is q-regular. In the same way we have $(\log F(x^q))')$ is q-regular so that

$$(\log F(x))' - (\log F(x^q))' = \frac{p'(x)}{p(x)}$$

is q-regular, then we conclude by Theorem 3.3 [4] that all roots are roots of unity.

To prove the second part, we use a method introduced in [8]. We firstly define some notation. Let us denote by $A_{t,i}$ the operator of power series:

$$A_{t,i}(\sum_{j=0}^{\infty} a_j x^j) = \sum_{j=0}^{\infty} a_{q^t j+i} x^{q^t j+i}$$

for all i such that $0 \leq i \leq q^t - 1$.

If there exists a root of p which's order is not a multiple of q, say α, then for all formal power series f, let us define $ord(f(x))$ to be the order of pole of f at point α. It is easy to check that there exists a $t \in \mathbf{N}$ such that for all $f \in F[[x]]$, $ord(f(x)) = ord(f(x^{q^t}))$ so there are some i such that $ord(f(x)) \leq ord(A_{t,i}(f(x)))$.

Now let us define a sequence of power series $(s_i)_{i\in\mathbf{N}}$ and a sequence of integer $(I_i)_{i\in\mathbf{N}}$ such that $s_0 = 1$, $0 \leq I_i \leq q^t - 1, \forall i$ and $ord(A_{I_i}(\frac{s_i}{p(x)})) \geq ord(\frac{s_i}{p(x)})$ and we define $s_{i+1} = A_{I_i}(\frac{s_i}{p(x)})$, so we can easily check

$$A_{I_i}(s_i F(x)) = A_{I_i}(\frac{s_i}{p(x)})F(x) = s_{i+1}F(x),$$

and by induction

$$A_{I_i}A_{I_{i-1}}...A_{I_0}(F(x)) = s_{i+1}F(x).$$

However,

$$ord(s_i) < ord(s_{i+1}),$$

the sequence s_i are linearly independent, so $F(x)$ can not be a regular sequence. □

Theorem 4.1. *If the power series $F(x) = \prod_{s=0}^{\infty} \frac{1}{p(x^{q^s})} = \sum_{i=0}^{\infty} c_i x^i$ is a q-regular sequence, then there exists a polynomial $Q(x)$ such that $p(x)|\frac{Q(x^q)}{Q(x)}$, so $F(x)$ can be written as*

$$F(x) = Q(x)\prod_{i=1}^{\infty} R(x^q),$$

where $R(x) = \frac{Q(x^q)}{Q(x)F(x)}$, which is a polynomial.

4.6 Applications

In this section we will consider some examples of automatic power series of type

$$F(x) = \prod_{s=0}^{\infty} p(x^{l^s}) = \sum_{i=1}^{\infty} c_i x^i,$$

where p is a polynomial of degree d with coefficients in $\mathbf{Q}$ and $l \geq 2$. It has been proved by Proposition 4.5 that the number of such polynomials p is fixed once given the degree d of the polynomial and l. But when l and d are both large, it will be difficult to compute the semi-group of matrix discussed in Section 4.2. Here we show a method applied on a particular example to generate the couples (p, l) such that $\prod_{s=0}^{\infty} p(x^{l^s}) = \sum_{i=1}^{\infty} c_i x^i$ is an automatic power sequence.

Let us consider firstly the power series $F_1(x)$ defined by $p_1(x) = 1 + x - x^3 - x^4$ and $l = 2$, it is easy to check that

$$F_1(x) = \prod_{s=0}^{\infty} p_1(x^{2^s}) = \prod_{s=0}^{\infty}(1 + x^{2^s})\prod_{s=0}^{\infty}(1 - (x^3)^{2^s}).$$

And it is well known that $\prod_{s=0}^{\infty}(1+x^{2^s}) = \frac{1}{1-x} = \sum_{i=1}^{\infty} x^i$ and $\prod_{s=0}^{\infty}(1-x^{2^s}) = \sum_{i=1}^{\infty} b_n x^i$, where $(b_n)_{n\in\mathbf{N}}$ is the Thue-Morse sequence beginning with $1, -1$. So the coefficient of term x^n in $F_1(x)$, say $f_1(n)$, can be calculated by

$$f_1(n) = \sum_{3i\leq n} b_i.$$

The sequence $(f_1(n))_{n\in\mathbf{N}}$ is bounded because of the fact that $b_{2n+1} + b_{2n} = 0$, so $F_1(x)$ is a 2-automatic power sequence. Moreover the transition matrices Γ_1 and Γ_0 can be defined by

$$\begin{pmatrix} c_{2n} \\ c_{2n-1} \\ c_{2n-2} \\ c_{2n-3} \end{pmatrix} = \Gamma_1 \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \end{pmatrix}$$

$$\begin{pmatrix} c_{2n+1} \\ c_{2n} \\ c_{2n-1} \\ c_{2n-2} \end{pmatrix} = \Gamma_0 \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} c_n \\ c_{n-1} \\ c_{n-2} \\ c_{n-3} \end{pmatrix}.$$

Remarking that

$$\Gamma_0^2 = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}, \Gamma_1\Gamma_0 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & -2 & 1 \end{pmatrix},$$

$$\Gamma_0\Gamma_1 = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \Gamma_1^2 = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix},$$

let us consider the the power series $F_2(x)$ defined by $p_2(x) = 1 + x + x^2 - x^4 - x^5 - 2x^6 - x^7 - x^8 + x^{10} + x^{11} + x^{12} = (x^2 + x + 1)(x^6 - 1)(x^4 - 1)$ and $l = 4$, the transition matrices of this polynomial are

$$\alpha_0 = \begin{pmatrix} 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 \end{pmatrix} \alpha_1 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{pmatrix}$$

$$\alpha_2 = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & -2 & 1 & 0 \end{pmatrix} \alpha_3 = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix}$$

If we define a sequence of matrices $(\alpha_n)_{n\in\mathbf{N}}$ by $\alpha_{4n+i} = \alpha_n\alpha_i, 0 \le i \le 3$, then the n-th coefficient of $F_2(x)$ is $f_2(n) = \alpha_n(1,1)$. However the matrices α_i for $i = 0,1,2,3$ are all of form $\begin{pmatrix} A_i & 0 \\ B_i & C_i \end{pmatrix}$ with A_i of size 4×4, B_i of size 4×1, C_i of size 1×1 and 0 the 0-matrix of size 1×4, so $\alpha_n(1,1)$ can be calculated only by the multiplications between A_i. Remarking that this four matrices are nothing else then $\Gamma_0^2, \Gamma_1\Gamma_0, \Gamma_0\Gamma_1, \Gamma_0^2$, we conclude that the sequence $(f_2(n))_{n\in\mathbf{N}}$ is bounded so 4-automatic.

By the same method, the power series $F_3(x)$ defined by $p_3(x) = 1 + x + x^2 - x^4 - x^5 + x^7 + x^8 - x^{10} - x^{11} - x^{12} = (x^2+x+1)(x^6+1)(1-x^4)$ and $l = 4$ is also 4-automatic. In fact, its transition matrices are

$$\beta_0 = \begin{pmatrix} 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 \end{pmatrix} \beta_1 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{pmatrix}$$

$$\beta_2 = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix} \beta_3 = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix}$$

and once more they are of form $\begin{pmatrix} A_i & 0 \\ B_i & C_i \end{pmatrix}$ with $A_0 = -\Gamma_0\Gamma_0\Gamma_1\Gamma_1, A_1 = -\Gamma_1\Gamma_0\Gamma_0\Gamma_0, A_2 = \Gamma_0\Gamma_1\Gamma_0\Gamma_0, A_3 = \Gamma_1\Gamma_1\Gamma_0\Gamma_0$.

Furthermore, as

$$\prod_{s=0}^{\infty}((x^2)^{4^s}+1)(x^{4^s}+1) = \prod_{s=0}^{\infty}\frac{(x^4)^{4^s}-1}{x^{4^s}-1} = \frac{1}{1-x}$$

we have

$$(1-x)F_2(x) = \prod_{s=0}^{\infty}((x^2)^{4^s}+x^{4^s}+1)((x^6)^{4^s}-1)\frac{(x^4)^{4^s}-1}{((x^2)^{4^s}+1)(x^{4^s}+1)} = \prod_{s=0}^{\infty}(x^9)^{4^s}-(x^6)^{4^s}-(x^3)^{4^s}+1,$$

$$(1-x)F_3(x) = \prod_{s=0}^{\infty}((x^2)^{4^s}+x^{4^s}+1)((x^6)^{4^s}+1)\frac{(x^4)^{4^s}-1}{((x^2)^{4^s}+1)(x^{4^s}+1)} = \prod_{s=0}^{\infty}-(x^9)^{4^s}+(x^6)^{4^s}-(x^3)^{4^s}+1.$$

Proposition 4.10. *The power series*

$$F_2(x) = \prod_{s=0}^{\infty}((x^2)^{4^s}+x^{4^s}+1)((x^6)^{4^s}-1)((x^4)^{4^s}-1)$$

and

$$F_3(x) = \prod_{s=0}^{\infty}((x^2)^{4^s}+x^{4^s}+1)((x^6)^{4^s}+1)(-(x^4)^{4^s}+1)$$

are 4*-automatic.*

Chapter 5

Palindromic length complexity and a generalization of Thue-Morse sequences

5.1 Introduction

The notion of palindromic length of a finite word as well as an infinite word was first introduced by Frid, Puzynina and Zamboni [25]. They conjectured that if the palindromic length of an infinite word is bounded, then this sequence is eventually periodic. This conjecture is widely studied in [25][23][6], and the palindromic length of some specific sequences are studied as well: Frid [23] showed that Sturmian words have an unbounded palindromic length PL_u and Ambrož and Pelantová [6] showed that PL_u grows arbitrarily slowly. [7] studied palindromic lengths of fixed points of a specific class of morphisms and gave upper bounds for the Fibonacci word and the Thue-Morse word. In this article, we give a formal expression of the palindromic length of Thue-Morse sequence and find all sequences which have the same palindromic length as Thue-Morse's. After writing a first version of this paper, we found that some results in the same direction were obtained by Frid [24] for Thue-Morse sequences. However, we will indicate how our results can be applied for a type of generalization of Thue-Morse sequences.

5.2 Definitions and notation

Let $(a_n)_{n\in\mathbf{N}}$ be a sequence and let us define a (finite) word, or a factor, of a sequence to be a (finite) string of the sequence. Let $w_a(x,y)$ denote the factor of the sequence $(a_n)_{n\in\mathbf{N}}$ beginning with a_x and of length y, in other words $w_a(x,y) = \overline{a_x a_{x+1} ... a_{x+y-1}}$.

Let $\widetilde{w}$ denote the reversal of w, that is to say, if $w = \overline{w_0 w_1 ... w_k}$ then $\widetilde{w} = \overline{w_k w_{k-1} ... w_0}$, we say a word w is palindromic if $w = \widetilde{w}$. Let Pal denote the set of all palindromic words.

We define the palindromic length of a word w, which will be denoted by $|w|_{pal}$, to be:

$$|w|_{pal} = \min \{k | w = p_1 p_2 ... p_k, p_i \in Pal,\ \forall i \in [1,k]\},$$

in this case we say $w = p_1 p_2 ... p_k, p_i$ is an optimal palindromic decomposition of w.

Let us define the palindromic length sequence $(pl_a(n))_{n\in\mathbf{N}}$ of the sequence $(a_n)_{n\in\mathbf{N}}$ to be

$$pl_a(n) = |w_a(0,n)|_{pal},$$

in other words, $pl_a(n)$ is the palindromic length of the word $\overline{a_0 a_1 ... a_{n-1}}$.

Now let us define a class of infinite sequences $\mathcal{C}$ which can be considered as a generalization of the Thue-Morse sequence:
Let $\sum$ be an alphabet which contains at least two letters and let $a \in \sum$.
Let F be the set of bijections over $\sum$.
Let $(f_n)_{n\in\mathbf{N}}$ be a sequence over F and $(w_n)_{n\in\mathbf{N}}$ be a sequence of finite words over $\sum$ which are defined recursively as:

$$f_i \in F \text{ such that } f_i(w_i) \neq w_i\ \forall i \geq 0,$$

and

$$\begin{cases} w_0 = a \\ w_n = w_{n-1} f_{n-1}(w_{n-1}) f_{n-1}(w_{n-1}) w_{n-1}, \forall n > 0. \end{cases}$$

Let $f(a)$ be the limit of the sequence $(w_n)_{n\in\mathbf{N}}$ which exists because of the definition.
The class $\mathcal{C}$ is the set of all infinite limits defined as above. It is easy to see that, if the size of $\sum$ is equal to 2, say $\sum = \{a,b\}$, then all sequences in $\mathcal{C}$ are Thue-Morse sequences, they may be written as

$$a, b, b, a, b, a, a, b, b, a, a, b, a, b, b, a...$$

and the one by changing a to b and b to a. Let $(pl(n))_{n\in\mathbf{N}}$ be the palindromic length sequence of Thue-Morse, the first elements of this sequence are

$$1, 2, 2, 1, 2, 3, 3, 2, 3, 4, 3, 2, 3, 3, 2, 1...$$

5.3 Palindromic length of sequences in $\mathcal{C}$

In this section we will study palindromic lengths of sequences in $\mathcal{C}$ and prove that they all have the same palindromic length, as the one of Thue-Morse.

Let $(a_n)_{n\in\mathbf{N}}$ be a sequence in $\mathcal{C}$, we will begin with some properties of palindromic factors of this sequence.

Lemma 5.1. *For any integer x, $w_a(4x, 4)$ is of type abba such that $a \neq b$.*
As a corollary, $a(2n+1) \neq a(2n)$.

Proof. This lemma is trivial because of the definition. □

Lemma 5.2. *Let $w_a(x, y)$ be a palindromic factor of a sequence in $\mathcal{C}$ such that y is odd, then y is either 1 or 3.*

Proof. If $w_a(x, y)$ is of size larger than 3, then it contains at least one palindromic word in the center of size 5, however a word of size 5 should be inside a word of type $xyyxxyyx$ or $xyyxf(x)f(y)f(y)f(x)$,where f is a bijection over the alphabet defined as above, but none of them contains a palindromic word of such a size. □

Lemma 5.3. *Let $w_a(x, y)$ be a palindromic word of $(a_n)_{n\in\mathbf{N}}$ such that y is even, then either there exist $z, r \in \mathbf{N}$ such that $w_a(x, y)$ is embedded into the center of palindromic word $w_a(4z, 4r)$, that is to say $4z < x, x + y < 4z + 4r$ and $x - 4z = 4z + 4r - x - y$, or $x \equiv 3 \mod 4$ and $y \equiv 2 \mod 4$.*

Proof. We first prove that $x+y/2-1$ is odd, otherwise $x+y/2-1 = 2t$ and $x+y/2 = 2t+1$ for some t, so that $a_{2t} = a_{2t+1}$ contradicts to Lemma 5.1. This fact implies that

$$\begin{cases} \text{if } x \equiv 0 \mod 4 \text{ then } x + y - 1 \equiv 3 \mod 4 \\ \text{if } x \equiv 1 \mod 4 \text{ then } x + y - 1 \equiv 2 \mod 4 \\ \text{if } x \equiv 2 \mod 4 \text{ then } x + y - 1 \equiv 1 \mod 4 \\ \text{if } x \equiv 3 \mod 4 \text{ then } x + y - 1 \equiv 0 \mod 4. \end{cases}.$$

For the last case, we have $x \equiv 3 \mod 4$ and $y \equiv 2 \mod 4$. Now let us check that, for other cases, the word $w_a(x, y)$ can be embedded into the center of a palindromic word of type $w_a(4z, 4r)$. Let $w_a(4z, 4r)$ be the shortest factor of above type including $w_a(x, y)$, here we prove that this factor is palindromic. It is easy to see that $w_a(x, y)$ is at the center of $w_a(4z, 4r)$ and the word $w_a(4(z+1), 4(r-1))$ is palindromic because of the palindromicity of $w_a(x, y)$; furthermore we have the fact that $w_a(4z, 4) = \widetilde{w_a}(4(z + r - 1), 4)$ when $x \not\equiv 3 \mod 4$, because these two words of length 4 are both palindromic and uniquely defined by respectively a prefix or a suffix of $w_a(x, y)$ of size smaller than 4 but larger than 1. In conclusion, the word $w_a(4z, 4r)$ is palindromic. □

Lemma 5.4. *Let $w_a(0,s)$ be a prefix of $(a_n)_{n\in\mathbf{N}}$ in $\mathcal{C}$, and let $w_a(0,s) = p_1p_2...p_r$ be an optimal palindromic decomposition such that for all $i : 1 \le i \le r$, p_i is either singleton or can be embedded into the center of palindromic word of type $w_a(4z, 4t)$, then there exists at least one optimal palindromic decomposition of $w_a(0,s)$ of following forms:*

$$\begin{cases} w_a(0,s) = q_1q_2...q_r;\, s \equiv 0 \mod 4;\\ w_a(0,s) = q_1q_2...q_{r-1}t_1;\, s \equiv 1 \mod 4;\\ w_a(0,s) = q_1q_2...q_{r-2}t_1t_2;\, s \equiv 2 \mod 4;\\ w_a(0,s) = q_1q_2...q_{r-2}t_1l_1;\, s \equiv 3 \mod 4;\\ w_a(0,s) = q_1q_2...q_{r-3}t_1t_2l_1;\, s \equiv 2 \mod 4; \end{cases}$$

where q_i are palindromes of length $4k_i$, t_i are singletons and l_1 are palindromes of length $2k$.

Proof. Let us consider a factor of $(a_n)_{n\in\mathbf{N}}$ of type $rq_1q_2..q_{2l}$ where $1 \le |r| \le 2$ beginning at some position $4x$ where q_i are palindromic words of even size and can be embedded into the center of palindromic word of type $w_a(4z, 4r)$. Here we prove that there exists an other palindromic decomposition of same length such that

$$rq_1q_2..q_{2l} = q'_1q'_2..q'_{2l}r,$$

where all q'_i are of size $4k_i$.

As q_1 is palindromic, because of Lemma 5.3, $rq_1\widetilde{r}$ is also palindromic, let us denote this word by q'_1, its size is multiple of 4. By excluding the case that $|r| = |q_2| = 2$, q_2 can be written as $\widetilde{r}q'_2r$, where q'_2 is either a palindromic word of size $4m$ or empty, so we have the equality $rq_1q_2 = q'_1q'_2r$ and the last r begins at some position $4x$. We do it recursively and we end up with the expression $rq_1q_2..q_{2l} = q'_1q'_2..q'_{2l}r$.

In such a way we can accumulate the singletons in the decomposition $w_a(0,s) = p_1p_2...p_r$ and push them to the end. An easy observation is that there are at most two singletons in an optimal decomposition, since once there are three singletons, they will meet each other by the above algorithm in a block $w_a(4k, 4l)$ hence two of them will create a palindromic word of length 2 which contradicts the optimality. The above process ends up with five possibilities:

$$\begin{cases} w_a(0,s) = q_1q_2...q_r; \\ w_a(0,s) = q_1q_2...q_{r-1}t_1; \\ w_a(0,s) = q_1q_2...q_{r-2}t_1t_2; \\ w_a(0,s) = q_1q_2...q_{r-2}t_1l_1; \\ w_a(0,s) = q_1q_2...q_{r-3}t_1t_2l_1; \end{cases}$$

where q_i are palindromes whose lengths are multiple of 4, t_i are singletons and l_i are palindromes whose lengths are multiple of 2.

The first case leads to $s \equiv 0 \mod 4$; the second one leads to $s \equiv 1 \mod 4$ and the third one leads to $s \equiv 2 \mod 4$; for the fourth one we can check that $|l_1| \not\equiv 0 \mod 4$ because of Lemma 5.3, so that $s \equiv 3 \mod 4$; the fifth case, $|l_1|$ must be a multiple of 4, so $s \equiv 2 \mod 4$.

□

Corollary 5.1. *Let $(pl(n))_{n\in\mathbf{N}}$ be the palindromic length of a sequence in $\mathcal{C}$ such that all its prefixes admit an optimal palindromic decomposition satisfying the constrains listed as in the previous lemma, then for all $k \geq 0$:*
$pl(4k+i) \geq pl(4k+3)+1$ for $i = 1,2$ and $pl(4k) \geq pl(4k+3)$.

Proof. For $i = 0$, $w_a(0, 4k+1)$ is of the form $q_1q_2...q_rt_1$. Using Lemma 5.1 we have $w_a(0, 4k+4) = q_1q_2...q_rq$ is a palindromic decomposition, not necessarily optimal, with $q = w_a(4k, 4)$, so $pl(4k+3) \leq r+1 = pl(4k)$

For $i = 1$, there are 2 cases: if $w_a(0, 4k+2)$ is of the form $q_1q_2...q_rt_1t_2$, then using Lemma 5.1 once more we have $w_a(0, 4k+4) = q_1q_2...q_rq$ is a palindromic decomposition, with $q = w_a(4k, 4) = t_1t_2t_2t_1$; if $w_a(0, 4k+2)$ is of the form $q_1q_2...q_rt_1t_2l_1$, using the hypothesis we have $w_a(0, 4k+4) = q_1q_2...q_rq$ is a palindromic decomposition, with $q = t_1t_2l_1t_2t_1$.

For $i = 2$, $w_a(0, 4k+3) = q_1q_2...q_rt_1l_1$, using the hypothesis we have $w_a(0, 4k+4) = q_1q_2...q_rq$ is a palindromic decomposition, with $q = t_1l_1t_1$.

So all inequalities as above are proved.

□

Lemma 5.5. *Let $w_a(0,k)$ be a prefix of $(a_n)_{n\in\mathbf{N}}$, then there is an optimal decomposition $w_a(0,k) = p_1p_2...p_s$ such that none of these palindromes is of length 3, furthermore, if p_i is of even size then it can be embedded into the center of palindromic word of type $w_a(4z, 4r)$.*

Proof. Let us suppose that k is the smallest number such that $w_a(0,k)$ does not satisfy one of the two constrains above, then either the last palindromic factor in all optimal compositions is of length 3, or it can not be embedded into the center of palindromic word of type $w_a(4z, 4r)$. If it is in the first case, then the last factor can be either

$\overline{a_{4t-1}a_{4t}a_{4t+1}}$ or $\overline{a_{4t-2}a_{4t-1}a_{4t}}$. If $k = 4t+2$, use the hypothesis that the last factors in all optimal decompositions of $w_a(0,k)$ are the same, which is $\overline{a_{4t-1}a_{4t}a_{4t+1}}$, we can decompose $w_a(0,4t+2)$ as $w_a(0,4t+2) = w_a(0,4t-1)p$ with $p = \overline{a_{4t-1}a_{4t}a_{4t+1}}$, so that the palindromic length is $pl(4t+1) = 1+pl(4t-2)$. However, if we decompose the same word as $w_a(0,4t+2) = w_a(0,4t)a_{4t}a_{4t+1}$, we have a length $pl(4t-1)+2$, so that

$$pl(4t+1) = 1 + pl(4t-2) < pl(4t-1) + 2.$$

Similarly for the case that $k = 4t+1$, by considering the decomposition $w_a(0,4t+1) = w_a(0,4t)a_{4t}$, we have

$$pl(4k) = 1 + pl(4k-3) < pl(4k-1) + 1,$$

Both inequalities contradict the previous corollary.

If the last factor can not be embedded into the center of a palindromic word of type $w_a(4z,4r)$, then because of Lemma 5.3 it can be found at some position $\overline{a_{4t-1}a_{4t}...a_{4l}}$, so the optimal decomposition is $w_a(0,4l) = w_a(0,4t-1)\overline{a_{4t-1}a_{4t}...a_{4l}}$. However, if we consider another composition $w_a(0,4l) = w_a(0,4t)\overline{a_{4t}a_{4t}...a_{4l-1}}a_{4l}$, we have

$$pl(4t-2) + 1 < pl(4t-1) + 2,$$

which contradicts the previous corollary.

□

Corollary 5.2. *Corollary 5.1 is valid for all sequences in $\mathcal{C}$.*

Lemma 5.6. *The palindromic length sequence $(pl(n))_{n\in\mathbf{N}}$ satisfies for $k \geq 0$:*

$$pl(4k+i) \leq pl(4k+3) + 2,$$

when $i = 0$ or 1; and

$$pl(4k+2) \leq pl(4k+3) + 1.$$

Proof. Let $w_a(0,4k+4) = p_1p_2...p_s$ be an optimal palindromic decomposition such that all p_j are of size $4r_j$ which exists because of Lemma 5.5.

If the size of p_s is larger than 4, then for $i = 1,2$ or 3, we can write $p_s = ab\widetilde{a}$ where a is the prefix of p_s of length $4-i$ so in this case

$$w_a(0,4k+i) = p_1p_2...p_{s-1}ab$$

and $|a|_{pal} = 2$ when $i = 3$ and $|a|_{pal} = 1$ otherwise.

If the size of p_s is 4, then for $i = 1,2$ or 3, we can write $p_s = ab$ where a is the prefix of p_s of length i so in this case

$$w_a(0, 4k+i) = p_1 p_2 ... p_{s-1} a$$

and $|a|_{pal} = 1$ when $i = 1$ and $|a|_{pal} = 2$ otherwise. In both cases the above inequalities hold. □

Lemma 5.7. *Let $(a_n)_{n\in\mathbf{N}}$ be a sequence in $\mathcal{C}$ defined over the alphabet $\sum$, let f be a bijection from $\sum^4$ to a new alphabet $\sum'$, then the sequence $(b_n)_{n\in\mathbf{N}}$ defined as*

$$b_n = f(\overline{a_{4n}a_{4n+1}a_{4n+2}a_{4n+3}})\ \forall n \in \mathbf{N}$$

is also in $\mathcal{C}$. As a consequence, $w_a(0, 4t) = p_1 p_2 ... p_k$ is an optimal palindromic decomposition of $w_a(0, 4t)$ if and only if $w_b(0, t) = f(p_1) f(p_2) ... f(p_k)$ is an optimal palindromic decomposition of $w_b(0, t)$ and the palindromic length sequence $(pl(n))_{n\in\mathbf{N}}$ satisfies for $k \geq 0$:

$$pl(4k+3) = pl(k)$$

Proof. The first part is easy to check by induction. For the second part, applying the algorithm introduced in Lemma 5.3 to $w_a(0, 4k+4)$, we get an optimal decomposition such that all palindromic words in the optimal decomposition are of size $4k_i$ and begin at some position $4r_i$. Applying f to $w_a(0, 4k+4)$ as well as each palindromic factor, we get a decomposition of a word of length $k+1$, which is a prefix of the sequence $(b_n)_{n\in\mathbf{N}}$, this decomposition is optimal because of the bijectivity of f. □

Corollary 5.3. *The palindromic length sequence $(pl(n))_{n\in\mathbf{N}}$ satisfies for $k \geq 0$:*
$pl(4k+3) = pl(k)$;
$pl(4k+2) = pl(4k+3) + 1$;
$pl(4k+1) = pl(4k+3) + 1$ or $pl(4k+3) + 2$;
$pl(4k) = pl(4k+3)$, $pl(4k+3) + 1$ or $pl(4k+3) + 2$.

Proposition 5.1. *The palindromic length sequence $(pl(n))_{n\in\mathbf{N}}$ satisfies for $k \geq 0$:*
$pl(4k+1) = pl(4k+3) + 1$ if $k \equiv 0 \mod 4$;
$pl(4k+1) = pl(4k+3) + 2$ if $k \equiv 2, 3 \mod 4$;
$pl(4k+1) - pl(4k+3) = pl(k) - pl(k+2)$ if $k \equiv 1 \mod 4$;
$pl(4k) = pl(k-1) + 1$.

Proof. If $k \equiv 0 \mod 4$, applying the bijection introduced in Lemma 5.7, the optimal decomposition of $w_a(0, 4k+4)$ is $w_a(0, 4k)\overline{a_{4k}a_{4k+1}a_{4k+2}a_{4k+3}}$, so that $w_a(0, 4k+2) = w_a(0, 4k)\overline{a_{4k}a_{4k+1}}$ is a decomposition of $w_a(0, 4k+2)$. As a result, $pl(4k+1) \leq pl(4k-1) + 2 = pl(4k+3) + 1$.

If $k \equiv 2, 3 \mod 4$, it is enough to prove that the last factor in any optimal palindromic decompositions of $w_a(0, 4k+4)$ is of length larger than 4. This is trivial by applying the bijection f to $w_a(0, 4k+4)$ and concluding by the classification in Lemma 5.4.

If $k \equiv 1 \mod 4$, applying the bijection introduced in Lemma 5.7 and Lemma 5.4, the optimal decomposition of $w_a(0, 4k+4)$ is either of type $p_1p_2...p_kt_1t_2$ or of type $p_1p_2...p_kt_1t_2l$, with p_i and l of length $16r_i$ and t_i of length 4. The first case implies $pl(4k+1)-pl(4k+3) = 1$ while the second case implies $pl(4k+1) - pl(4k+3) = 2$. However, if we apply f to $w_a(0, 4k+4)$ we get a word of length $k+1$ and $pl(k) - pl(k+2) = 1$ in the first case and $pl(k) - pl(k+2) = 2$ in the second case.

The last equality is a consequence of Lemma 5.4 and Lemma 5.7. □

Proposition 5.2. *All sequences in $\mathcal{C}$ share the same palindromic length sequence $(pl(n))_{n\in\mathbf{N}}$. Furthermore, this palindromic length sequence is 4-regular.*

Proof. The 4-kernel of $(pl(n))_{n\in\mathbf{N}}$ is generated by the elements in

$$\{(pl(n))_{n\in\mathbf{N}}, (pl(n-1))_{n\in\mathbf{N}}, (pl(n+1))_{n\in\mathbf{N}}, (b(n))_{n\in\mathbf{N}}, (1)_{n\in\mathbf{N}}\},$$

where the sequence $(b(n))_{n\in\mathbf{N}}$ is defined by $b(n) = pl(4n+1) - pl(4n+3)$. By definition, the sequence $(b(n))_{n\in\mathbf{N}}$ satisfies relations $b(4n) = 1$, $b(4n+1) = b(n)$ and $b(4n+2) = b(4n+3) = 2$. Les us define matrices Γ_i by

$$\begin{pmatrix} (pl(4n+i-1))_{n\in\mathbf{N}} \\ (pl(4n+i))_{n\in\mathbf{N}} \\ (pl(4n+i+1))_{n\in\mathbf{N}} \\ (b(4n+i))_{n\in\mathbf{N}} \\ (1)_{n\in\mathbf{N}} \end{pmatrix} = \Gamma_i \begin{pmatrix} (pl(n-1))_{n\in\mathbf{N}} \\ (pl(n))_{n\in\mathbf{N}} \\ (pl(n+1))_{n\in\mathbf{N}} \\ (b(n))_{n\in\mathbf{N}} \\ (1)_{n\in\mathbf{N}} \end{pmatrix}$$

for $i = 0, 1, 2, 3$. Using relations determined in Corollary 5.3 and Proposition 5.1, we have

$$\begin{aligned}
pl(4n-1) &= pl(4(n-1)+3) = pl(n-1) \\
pl(4n) &= pl(n-1)+1 \\
pl(4n+1) &= b(n) + pl(4n+3) = b(n) + pl(n) \\
pl(4n+2) &= pl(4n+3)+1 = pl(n)+1 \\
pl(4n+3) &= pl(n) \\
pl(4n+4) &= pl(n)+1
\end{aligned}$$

So that the Γ_i can be determined as follows:

$$\Gamma_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \Gamma_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\Gamma_2 = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \Gamma_3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

□

Proposition 5.3. *Let k be a positive integer, then for any integer $i = 0, 1, 2, ...15$, $pl(16k + i) \leq pl(16k - 1) + 3$. As a consequence,* $\limsup pl(n) = 3\frac{\ln(n)}{\ln(16)}$.

Proof. Using Proposition 5.1 and Corollary 5.3, we have: $pl(16k-1) = pl(16(k-1)+15) = pl(k-1)$, furthermore,

$$\begin{aligned}
&pl(16k) = pl(4k-1) + 1 = pl(k-1) + 1 \\
&pl(16k+1) = pl(16k+3) + 1 = pl(4k) + 1 = pl(k-1) + 2 \\
&pl(16k+2) = pl(16k+3) + 1 = pl(k-1) + 2 \\
&pl(16k+3) = pl(4k) = pl(k-1) + 1 \\
&pl(16k+4) = pl(16k+3) + 1 = pl(k-1) + 2 \\
&pl(16k+5) \leq pl(16k+4) + 1 \leq pl(k-1) + 3 \\
&pl(16k+6) = pl(16k+7) + 1 = pl(4k+1) + 1 \leq pl(4k) + 2 = pl(k-1) + 3 \\
&pl(16k+7) = pl(4k+1) \leq pl(k-1) + 2 \\
&pl(16k+8) = pl(4k+1) \leq pl(k-1) + 2 \\
&pl(16k+9) = pl(4k+2) = pl(k) + 1 \leq pl(k-1) + 1 \\
&pl(16k+10) = pl(16k+11) + 1 = pl(4k+2) + 1 = pl(k) + 2 \leq pl(k-1) + 3 \\
&pl(16k+11) = pl(4k+2) \leq pl(k-1) + 2 \\
&pl(16k+12) = pl(4k+2) + 1 \leq pl(k-1) + 3 \\
&pl(16k+13) = pl(16k+15) + 2 = pl(k) + 2 \leq pl(k-1) + 3 \\
&pl(16k+14) = pl(16k+15) + 1 = pl(k) + 1 \leq pl(k-1) + 2 \\
&pl(16k+15) = pl(k) \leq pl(k-1) + 1.
\end{aligned}$$

The first part of the statement is a straightforward consequence of above equations, furthermore, we have $\limsup pl(n) \leq 3\frac{\ln(n)}{\ln(16)}$.

On the other hand, if we define a sequence $(k(i))_{i\in\mathbf{N}}$ by $k(1) = 5,\ k(i+1) = k(i)\times 16+9$, then we can get easily $pl(k(i)) = 3i$. consequently $\limsup pl(n) = 3\frac{\ln(n)}{\ln(16)}$.

□

Remark Lemma 5.7 and Proposition 5.1 are critical in the proof because they show the importance of the hypothesis that $f_n(w_n) \neq w_n$. Because of this hypothesis, we can guarantee that the set $\mathcal{C}$ is closed under bijections (and their inverses) defined in Lemma 5.5, and do not have factors like $aaaa$ in the sequence. So that we can apply some inductive properties by saying that $w_a(0, 4k+4)$ and $w_b(0, k+1)$ share the "same" optimal palindromic decomposition, which is the key point to make Proposition 5.1 work.

Corollary 5.4. *$pl(n) + 1 \geq pl(n + 1)$;*
if there exists an integer n satisfying $pl(n) + 2 = pl(n + 1) + 1 = pl(n + 2)$, then $n \equiv 3 \mod 4$;
if $pl(4k) = pl(4k + 3)$ then $pl(4k + 1) = pl(4k + 2) = pl(4k + 3) + 1$;
if $pl(4k) = pl(4k + 1)$ then $pl(4k) = pl(4k + 1) = pl(4k + 3) + 2$.

Proof. The first statement is trivial because of a decomposition $w_a(0, n + 2) = w_a(0, n + 1)a_{n+1}$.

For the second statement, remarking the fact that $pl(4k + 3) = pl(4k + 2) - 1$, we have either $n \equiv 3 \mod 4$ or $n + 3 \equiv 3 \mod 4$, but if we are in the last case, then $pl(n) + 2 = pl(n + 2) = pl(n + 3) + 1$ so that $pl(n + 3) > pl(n)$ which contradicts Corollary 5.3.

For the last two statements, $pl(4k) = pl(4k+3)$ implies that the last palindromic factor in optimal decompositions of $w_a(0, 4k+4)$ is $\overline{a_{4k}a_{4k+1}a_{4k+2}a_{4k+3}}$ which proves $pl(4k+1) = pl(4k+2) = pl(4k+3)+1$. On the contrary, if $pl(4k) = pl(4k+1)$ then $pl(4k) \neq pl(4k+3)$, so that the last palindromic factor in optimal decompositions of $w_a(0, 4k + 4)$ is of length larger than 4, which leads to the fact $pl(4k) = pl(4k + 1) = pl(4k + 3) + 2$. □

5.4 All sequences sharing $(pl(n))_{n\in\mathbf{N}}$

In this section, we are going to prove that all sequences sharing the same palindromic length $(pl(n))_{n\in\mathbf{N}}$ defined in the previous section are exactly the sequences in $\mathcal{C}$.

Lemma 5.8. *Let $(b_n)_{n\in\mathbf{N}}$ be a sequence such that all words $w_b(4k, 4)$ are of form $xyyx$, then*
1) if $w_b(a, b)$ is a palindromic word and b is odd, then $b \leq 3$, furthermore, if $b = 3$, then

$a \equiv 3$ *or* $0 \mod 4$.
2) if $w_b(a,b)$ *is a palindromic word and* b *is even, then* $a+b/2-1$ *is odd.*

Proof. It is analogous to Lemma 5.2 and Lemma 5.3. □

Lemma 5.9. *Let* $(b_n)_{n\in\mathbf{N}}$ *be a sequence such that its palindromic length sequence coincides with* $(pl(n))_{n\in\mathbf{N}}$*, then all word* $w_b(4k,4)$ *are of form* $xyyx$ *with* $x\neq y$*.*

Proof. We prove the statement by induction:

Firstly the statement holds for $s=0$. Suppose that this statement is true for all $s\leq s_0$, we will prove it for $s=s_0+1$.

Let us consider a decomposition $w_b(0,4s_0+4)=p_1p_2...p_r$ such that $r=pl(4s_0+4)$, and let n denote the length of p_r.
Firstly n can not be too small: if $n<4$ then $pl(4s_0+3)=1+pl(4s_0+3-n)>pl(4s_0+3)$ which contradicts Corollary 5.2.
Secondly, if n is odd then it can not be too large: if $n=2n_0+1$ and $n_0>4$ then $w_b(4s_0+6-2n_0,2n_0-7)$ is a palindrome of odd size larger or equal to 3 and finishing at the position $4s_0-1$, which does not exist because of the Lemma 5.8.
Thirdly, if n is even and large enough: if $n=2n_0$ and $n_0\geq 4$, then, because of Lemma 5.8, n is a multiple of 4 and $w_b(4s_0-1,4)$ is the inverse of some word $xyyx$.
So there are 5 other cases to study: $n=4,5,6,7,9$.

When $n=4$, $w_b(4s_0,4)$ is either of type $xxxx$ or $xyyx$, and $pl(4s_0-1)+1=pl(4s_0+3)$, if $w_b(4s_0,4)$ is of type $xxxx$ then $w_b(0,4s_0+3)=w_b(0,4s_0)xxxx$ so $pl(4s_0+2)\leq pl(4s_0-1)+1=pl(4s_0+3)$, contradicting Corollary 5.2.

When $n=5$ or 6, $pl(4s_0+3)=pl(4s_0+3-n)+1>pl(4s_0-1)+1$, however, $pl(4s_0+3)\leq pl(4s_0-1)+1$, contradiction.

When $n=7$, $pl(4s_0+3)=pl(4s_0-4)+1=pl(4s_0-5)+2$. On the other hand, $pl(4s_0-1)\leq pl(4s_0-5)+1$ and $pl(4s_0+3)\leq pl(4s_0-1)+1$, so

$$pl(4s_0+3)=pl(4s_0-1)+1=pl(4s_0-5)+2. \ (*)$$

After Corollary 5.4, $4s_0-4\equiv 0 \mod 16$ and

$$pl(4s_0+7)=pl(4s_0+3) \text{ or } pl(4s_0+7)=pl(4s_0+3)-1 \ (**).$$

If we write $w_b(0,4s_0+4)=w_b(0,4s_0-4)abbaxabb$ let us consider the last palindromic factor of $w_b(0,4s_0+8)$:

1) The length can not be smaller than 4, otherwise $pl(4s_0+7) = pl(4s_0+i)+1 > pl(4s_0+7)$ with $3 < i < 6$, contradicting Corollary 5.3.
2) The length can not be 4, 5, 6, 7, otherwise $pl(4s_0+7) = pl(4s_0+i)+1$ with $-1 < i \leq 3$, but $pl(4s_0+i) \geq pl(4s_0+3)$, so that $pl(4s_0+7) > pl(4s_0+3)+1$, contradicting (**).
3) The length can not be 8, otherwise $w_b(0, 4s_0+8) = w_b(0, 4s_0-4)abbaxabbbbax$ and $pl(4s_0+7) = pl(4s_0-1)+1 = pl(4s_0+3)$. But on the other hand, $pl(4s_0+4) = pl(4s_0+3)+1$ and $pl(4s_0+6) = pl(4s_0+7)+1 = pl(4s_0+3)+1$ because of Proposition 5.1; $pl(4s_0+5) = pl(4s_0+3)+1$ because of the decomposition $w_b(0, 4s_0+6) = w_b(0, 4s_0+4)bb$ so that $pl(4s_0+4) = pl(4s_0+5) = pl(4s_0+6)$ which contradicts Corollary 5.4.
4) The length can not be 9, 10, otherwise $pl(4s_0+7) = pl(4s_0-i)+1$ with $i = 2, 3$, but $pl(4s_0-i) \geq pl(4s_0-1)+1$ so that $pl(4s_0+7) > pl(4s_0-1)+1 = pl(4s_0)+3$, which contradicts (**).
5) The length can not be 11, 12, 13, 14, 16, 17, because the last factor can not be palindromic.
6) The length can not be 15, otherwise, $w_b(0, 4s_0+8) = w_b(0, 4s_0-8)cddcabbaxabbacdd$, with $a \neq b, c \neq d$. Let us check a decomposition $w_b(0, 4s_0+5) = w_b(0, 4s_0-4)abbaxabba$, so that $pl(4s_0+4) \leq pl(4s_0-5)+1$, but $pl(4s_0+4) = pl(4s_0+3)+1$ which implies $pl(4s_0+3) \leq pl(4s_0-5)$, which contradicts (*).
7) The length can not be an odd number larger than 15, otherwise, there is a palindromic factor of odd size larger than 3 in $w_a(0, 4s_0)$ finishing at position $4s_0-1$, contradicting Lemma 5.8.
8) The length can not be an even number larger than 14, otherwise, because of Lemma 5.8, the length is a multiple of 4, which implies the factor $w_a(4s_0, 4)$ is the symmetrical of some words $w_a(4x, 4)$, by hypothesis, it is of type *abba* but not *xabb*.
In conclusion, the last palindromic factor of $w_a(0, 4s_0+4)$ can not be 7.

When $n = 9$, $pl(4s_0+3) = pl(4s_0-6)+1 = pl(4s_0-5)+2 \geq pl(4s_0-1)+1$. On the other hand, $pl(4s_0+3) \leq pl(4s_0-1)+1$, so $pl(4s_0+3) = pl(4s_0-1)+1$; another observation is that $pl(4s_0+2) \leq pl(4s_0-5)+1$ because $\overline{b_{4s_0-4}, b_{4s_0-3}, b_{4s_0-2}, b_{4s_0-1}, b_{4s_0}, b_{4s_0+1}, b_{4s_0+2}}$ is palindromic, but $pl(4s_0+2) = pl(4s_0+3)+1$ so $pl(4s_0+3)+2 \leq pl(4s_0-5)+2 = pl(4s_0+3)$, contradiction.

In conclusion, for all possible cases $w_b(4s_0, 4)$ is of type *xyyx*.

□

Proposition 5.4. *Let w be a finite word of length 4^k, such that its palindromic length sequence coincides with a prefix of $(pl(n))_{n\in\mathbf{N}}$, then w is a prefix of a sequence in $\mathcal{C}$.*

Proof. Let us prove it by induction. The statement is trivially true when $k = 0$. Now suppose the statement is true for $k = s_0$, let us consider the case $k = s_0+1$:

Remarking that Lemma 5.2, 5.3, 5.4, 5.5 work under the weaker condition on sequences announced in previous proposition, we can apply the same arguments to prove each prefix of w of length $4k$ admits an optimal palindromic decomposition of type $p_1p_2..p_r$ such that the length of all this factors are multiples of 4. Using Lemma 5.7 there is another alphabet Σ_1 and a bijection $f : \Sigma^4 \to \Sigma_1$ such that $f(w)$ is still a word which palindromic length sequence coincides with a prefix of $(pl(n))_{n\in\mathbf{N}}$, however the length of $f(w)$ is 4^{s_0}, using the hypothesis of induction, it is a prefix of a sequence in $\mathcal{C}$, so w is also a prefix of a sequence in $\mathcal{C}$, by applying the inverse of f.

□

Theorem 5.1. *All sequences such that their palindromic length sequences coincide with the one of Thue-Morse's are in $\mathcal{C}$.*

Bibliography

[1] J.-P. Allouche and L. Goldmakher. Mock characters and the Kronecker symbol. *Journal of Number Theory*, 192:356–372, 2018.

[2] J.-P. Allouche, M. Mendès France, and J. Peyrière. Automatic Dirichlet series. *Journal of Number Theory*, 81(2):359–373, 2000.

[3] J.-P. Allouche, S. Riasat, and J. Shallit. More infinite products: Thue–Morse and the gamma function. *The Ramanujan Journal*, 49(1):115–128, 2019.

[4] J.-P. Allouche and J. Shallit. The ring of k-regular sequences. *Theoretical Computer Science*, 98(2):163–197, 1992.

[5] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, New York, NY, USA, 2003.

[6] P. Ambrož and E. Pelantová. A note on palindromic length of Sturmian sequence. *European Journal of Combinatorics, submitted*, 2018. http://arxiv.org/abs/1808.08879.

[7] P. Ambrož, O. Kadlec, Z. Masáková, and E. Pelantová. Palindromic length of words and morphisms in class $\mathcal{P}$. *Theoretical Computer Science*, 780:74–83, 2019.

[8] P. G. Becker. k-Regular power series and Mahler-type functional equations. *Journal of Number Theory*, 49(3):269–286, 1994.

[9] A. Blondin-Massé, S. Brlek, and S. Labbé. Palindromic lacunas of the Thue-Morse word. *Proc GASCOM 2008*, pages 53–67, 01 2008.

[10] S. Brlek. Enumeration of factors in the Thue-Morse word. *Discrete Applied Mathematics*, 24(1):83–96, 1989.

[11] S. Checcoli and J. Roques. On some arithmetic properties of Mahler functions. *Israel Journal of Mathematics*, 228(2):801–833, 2018.

[12] G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bulletin de la Société Mathématique de France*, 108:401–419, 1980.

[13] A. Cobham. Uniform tag sequences. *Mathematical Systems Theory*, 6(1):164–192, 1972.

[14] H. Cohen and J. P. Allouche. Dirichlet series and curious infinite products. *Bulletin of the London Mathematical Society*, 17(6):531–538, 11 1985.

[15] H. Delange. Sur les fonctions arithmétiques multiplicatives. *Ann. Sci. École Norm. Sup.*, 78:273–304, 1961.

[16] W. Duke and H. N. Nguyen. Infinite products of cyclotomic polynomials. *Bulletin of the Australian Mathematical Society*, 91(3):400–411, 2015.

[17] P. Dumas. *Mahlerian recurrences, automatic sequences, asymptotic studies.* Thèse, Université Bordeaux I, Sept. 1993.

[18] P. D. T. A. Elliott. *Multiplicative Functions* $|g| \leq 1$ *and their Convolutions: An Overview*, pages 63–75. Birkhäuser Boston, Boston, MA, 1990.

[19] P. Erdős. Some unsolved problems. *Michigan Math. J.*, 4(3):291–300, 1957.

[20] D. Essouabri. Singularité de séries de Dirichlet associées à des polynômes de plusieurs variables et applications en théorie analytique des nombres. *Annales de l'Institut Fourier*, 47(2):429–483, 1997.

[21] N. Frantzikinakis. An Averaged Chowla and Elliott Conjecture Along Independent Polynomials. *International Mathematics Research Notices*, 2018(12):3721–3743, 2018.

[22] N. Frantzikinakis and B. Host. Furstenberg Systems of Bounded Multiplicative Functions and Applications. *International Mathematics Research Notices, to appear*, 2019. preprint, `http://arxiv.org/abs/1804.18556`.

[23] A. Frid. Sturmian numeration systems and decompositions to palindromes. *European Journal of Combinatorics*, 71:202–212, 2018.

[24] A. Frid. First Lower Bounds for Palindromic Length. In *Developments in Language Theory*, pages 234–243, Cham, 2019. Springer International Publishing.

[25] A. Frid, S. Puzynina, and L. Zamboni. On palindromic factorization of words. *Advances in Applied Mathematics*, 50(5):737–748, 2013.

[26] G. Halász. Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen. *Acta Mathematica Hungarica*, 19(3-4):365–403, 1968.

[27] Y. Hu. Subword complexity and non-automaticity of certain completely multiplicative functions. *Advances in Applied Mathematics*, 84:73–81, 2017.

[28] N. B. J. P. Bell and M. Coons. Transcendence of generating functions whose coefficients are multiplicative. *Transactions of the American Mathematical Society*, 364(2):933–959, 2012.

[29] O. Klurman and A. P. Mangerel. Rigidity theorems for multiplicative functions. *Mathematische Annalen*, 07 2017.

[30] M. Lemańczyk and C. Müllner. Automatic sequences are orthogonal to aperiodic multiplicative functions. 11 2018. preprint, `http://arxiv.org/abs/1811.00594`.

[31] A. Luca. On some combinatorial problems in free monoids. *Discrete Mathematics*, 38(2):207–225, 1982.

[32] K. Mahler. Über einen Satz von Mellin. *Mathematische Annalen*, 100:384–398, 1928.

[33] A. Mandel and I. Simon. On finite semigroups of matrices. *Theoretical Computer Science*, 5(2):101 – 111, 1977.

[34] R. H. Mellin. Eine Formel für den Logarithmus transcendenter Funktionen von endlichen Geschlecht. *Acta Soc. Scient. Fennicae*, 4:29, 1900.

[35] H. Minc. *Nonnegative matrices*. Wiley, 1988.

[36] J. J. Pansiot. The Morse sequence and iterated morphisms. *Information Processing Letters*, 12(2):68–70, 1981.

[37] D. Robbins. Solution to problem E2692. *The American Mathematical Monthly*, 86:394–395, 1979.

[38] I. Z. Ruzsa. General multiplicative functions. *Acta Arithmetica*, 32(4):313–347, 1977.

[39] J. S. S. Brown, N. Rampersad and T. Vasiga. Squares and overlaps in the Thue-Morse sequence and some variants. *RAIRO - Theoretical Informatics and Applications - Informatique Théorique et Applications*, 40(3):473–484, 2006.

[40] O. Salon. Suites automatiques à multi-indices. *Séminaire de Théorie des Nombres de Bordeaux*, pages 1–36A, 1986.

[41] P. Sargos. Prolongement méromorphe des séries de Dirichlet associées à des fractions rationnelles de plusieurs variables. *Annales de l'Institut Fourier*, 34(3):83–123, 1984.

[42] J.-C. Schlage-Puchta. Completely multiplicative automatic functions. *Integers*, 11:8, 2011.

[43] E. Wirsing. Das asymptotische Verhalten von Summen über multiplikative Funktionen. II. *Acta Mathematica Hungarica*, 18(3-4):411–467, 1967.

[44] D. R. Woods. Problem proposal E2692. *The American Mathematical Monthly*, 85:48, 1978.