On completely multiplicative automatic sequences

Shuo Li

On completely multiplicative automatic sequences

Shuo LI

CNRS, Institut de Mathématiques de Jussieu-PRG
Université Pierre et Marie Curie, Case 247
4 Place Jussieu
F-75252 Paris Cedex 05 (France)
shuo.li@imj-prg.fr

Abstract

In this article we prove that all completely multiplicative automatic 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.

1 Introduction

In this article we give a decomposition of completely multiplicative automatic sequences, which will be referred as CMAS. In article [SP11], the author proves that a non-vanishing CMAS is almost periodic (defined in [SP11]). In article [AG18], the authors give a formal expression to all non-vanishing CMAS and also some examples in the vanishing case (named as mock characters). In article [Hu17], the author studies completely multiplicative sequences, which will be referred as CMS, taking values on a general field, which 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 firstly prove that all CMAS are mock characters (defined in [AG18]) with an exceptional case. Secondly, 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 sequence or strongly aperiodic sequences. Lastly we conclude by proving that a strongly aperiodic sequence can not be automatic.

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 $k \geq 2$ an integer, we say 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 initial state, and for each state there are exactly k edges pointing from this state to others, 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, starting from the initial state and moving from one state to another by taking the edge read in the k-ary expansion one by one until stop 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 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 [SP11]).

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 respectively, 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.

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 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 [SP11](Lemma 3 and Theorem 1), the author proves that in an automaton, every state which can be reached from a specific state, say s, with $q!$ steps, can be reached with $yq!$ steps for every $y \geq 1$; and 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.

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

Let us consider $(a_n)_{n\in\mathbf{N}}$ a CMS 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 the subgroup of G generated by E.

Definition We say that an element ζ of a sequence $(a_n)_{n\in\mathbf{N}}$ has a natural density if and only if $\lim_{N\to\infty} \frac{\sharp\{n|a_n=\zeta, 0\le n\le N\}}{N+1}$ exists, and 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 2 *Let $(a_n)_{n\in\mathbf{N}}$ be a CMS taking values in a finite Abelian group G, then for all elements $g \in G$, the sequence $a^{-1}(g) = \{n : a_n = g\}$ has 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, [Ruz77], and when G is a semi-group, the Theorem 7.3, [Ruz77] shows that all elements in G have a natural density. To conclude the proof, it is enough to consider the following fact: let f_0 be a CMS such that there exists a prime p with $a_p = 0$, 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 $d_0(g), d_1(g)$ denote respectively the natural density of g in the sequence $(f_0(n))_{n\in\mathbf{n}}$ and $(f_1(n))_{n\in\mathbf{N}}$, 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 regressively until a non-vanishing sequence, we can conclude by the first part of the proposition.

3 Finiteness of the numbers of 0-factors

In this section we will prove that a CMAS is either a mock character, which means it has only finitely many 0-factors, or an almost-0-sequence, that is to say $a_m = 0$ for all m which is 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 *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 on 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:
It is equivalent to find some $m \in \mathbf{N}$ and $p^{2q!}$ 0-factors, say $a_{p_1}, a_{p_2}, ..., a_{p_{p^{2q!}}}$, such that

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

If m is a solution of the above system, then the subword $\overline{a_m a_{m+1} ... a_{m+p^{2q!}-1}}$ is constant to 0. So there exists a m' such that $m \leq m'p^{q!} < (m'+1)p^{q!} \leq m + p^{2q!}$. Because of Proposition 1, for any $y \in \mathbf{N}$, $a_k = 0$ for all k such that $m'p^{yq!} \leq k < (m'+1)p^{yq!}$. 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 $m'p^{yq!} \leq r^t < (m'+1)p^{yq!}$, this inequality holds because we can find some integer t and y such that:

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

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 with any prime numbers, $a_m = 0$ 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 the automaticity, we can replace p' with p.

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 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 other prime factor than $p_1, p_2, ..., p_u$.*

Proof We prove it 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 the statement is true for some $u = n_0$, let us consider the case $u = n_0 + 1$. We firstly consider the the sequence $(a'_n)_{n \in \mathbf{N}}$ defined as $a'_n = a_{\frac{n}{p_{n_0+1}^{v_{p_{n_0+1}}(n)}}}$, 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 get 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 $N_1(n)$, $N_2(n)$ satisfying separately the 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 respectively $a_{m_{n_0'} N_1(n)+l_{n_0}+n_0} = Ca_{p_{n_0+1}}$ and $a_{m_{n_0'} N_2(n)+l_{n_0}+n_0} = Ca^2_{p_{n_0+1}}$ for all $n \in \mathbf{N}$. As $a_{p_{n_0+1}} \neq 1$, there is at least one element of $Ca_{p_{n_0+1}}, Ca^2_{p_{n_0+1}}$ not equal to g. If $N_i(n)$ is the associated residue class, then $N_i(n) = p^{i+1}_{n_0+1} n + t$ for all integer n with $t \in \mathbf{N}$, $i = 1$ or 2.

Now let us choose $m_{n_0+1} = m_{n_0'} p^{i+1}_{n_0+1}$ and $l_{n_0+1} = l_{n_0} + tm_{n_0'}$, 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 $\overline{a'_{m_{n_0+1}n+l_{n_0+1}} a'_{m_{n_0+1}n+l_{n_0+1}+1} \ldots a'_{m_{n_0+1}n+l_{n_0+1}+n_0-1}}$ is constant and none of its letters equals 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 independent of n 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 the fact that for all j such that $0 \leq j \leq n_0 - 1$, $p_{n_0+1} \nmid m_{n_0+1} n + l_{n_0+1} + j$. So 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 $\overline{a_{m_{n_0+1}n+l_{n_0+1}} a_{m_{n_0+1}n+l_{n_0+1}+1} \ldots a_{m_{n_0+1}n+l_{n_0+1}+n_0}}$ is a subword of length $n_0 + 1$ independent of n and none of its letters equals g, what is more m_{n_0+1} does not have any prime factor other than $p_1, p_2 \ldots p_{n_0}$.

Proposition 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 $a_{p_1}, a_{p_2}, \ldots, a_{p_r}$. If there is a subword $\overline{w}_r$ appearing periodically in $(a'_n)_{n\in\mathbf{N}}$, and the period does not have any other prime factors than $p_1, p_2, \ldots, p_r$, then this subword appears at least once in $(a_n)_{n\in\mathbf{N}}$.*

Proof Let us denote by $p_1, p_2 \ldots$ 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 other prime factor than $p_1, p_2, \ldots, p_r$. So the total number of such subwords in the sequence $(a_n)_{n\in\mathbf{N}}$ can be bounded by the inequality:

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

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

$$\#\left\{a_k | k \leq N(t) m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k + j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} = \prod_{j=1}^{t} (p_{r+j} - r) \quad (2)$$

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

So we have

$$
\begin{aligned}
&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, i > r\right\} \\
>&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} \\
&- \#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \mid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, i > r+t\right\} \\
>&\#\left\{a_k | k \leq N(t)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \nmid k+j, \forall (i,j) \text{ with } 0 \leq j \leq r-1, r < i \leq r+t\right\} \\
&- \sum_{i > r+t} \#\left\{a_k | k \leq N(i)m_r + l_r, k = m_r k^{'} + l_r, k^{'} \in \mathbf{N}; p_i \mid k+j, \forall j \text{ with } 0 \leq j \leq r-1\right\} \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r \sum_{i > r+t, p_i < N(t)+r} [\frac{N(t)}{p_i}] \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r \sum_{i > r+t, p_i < N(t)+r} \frac{N(t)}{p_i} - r\pi(N(t)+r).
\end{aligned} \tag{3}
$$

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

$$
\prod_{j=1}^{t}(p_{r+j} - r) = \prod_{j=1}^{t} \frac{p_{r+j} - r}{p_{r+j}} N(t) \geq \prod_{j=1}^{\infty} \frac{p_{r+j} - r}{p_{r+j}} N(t). \tag{4}
$$

The last formula can be approximates as $\prod_{j=1}^{\infty} \frac{p_{r+j}-r}{p_{r+j}} = \exp(\sum_{j=1}^{\infty} \log(\frac{p_{r+j}-r}{p_{r+j}})) = \exp(-\Theta(\sum_{j=1}^{\infty} \frac{r}{p_{r+j}}))$, 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 large enough, 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 $i > r+t$, $p_i^t > \prod_{j=1}^{t} p_{r+j} = N(t)$, so $p_i > N(t)^{\frac{1}{t}}$

$$
\sum_{i > r+t, p_i < N(t)+r} \frac{N(t)}{p_i} < N(t) \sum_{N(t)^{\frac{1}{t}} < p < N(t)+r} \frac{1}{p}. \tag{5}
$$

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

$$
N(t)^{\frac{1}{t}} = (\prod_{j=1}^{t} p_{r+j})^{\frac{1}{t}} \geq \frac{t}{\sum_{j=1}^{t} \frac{1}{p_{r+j}}} > \frac{t}{\sum_{j=1}^{t} \frac{1}{q_j}}. \tag{6}
$$

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 $\sum_{p_i \leq x} \frac{1}{p_i} \sim \log\log(x)$, so $N(t)^{\frac{1}{t}}$ 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 $t > t_0$,

$$
\begin{aligned}
&\#\{a_k | k \leq N(t)m_r + l_r, k = m_r k' + l_r, k' \in \mathbf{N}; k+j \nmid p_i, \forall (i,j) \text{ with } 0 \leq j \leq r-1, \forall i > r\} \\
>&\prod_{j=1}^{t}(p_{r+j} - r) - r \sum_{k > r+t} \frac{N(t)}{p_k} - r\pi(N(t)+r) \\
>&cN(t) - \frac{1}{2}cN(t) - r\pi(N(t)+r).
\end{aligned} \tag{7}
$$

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

Proposition 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 Supposing that the sequence $(a_n)_{n\in\mathbf{N}}$ has infinitely many prime factors not equal to 0 or 1. Let us consider firstly 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, this set is finite.

Using the Proposition 4 and 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 equals $g = a'_{p_1} = a_{p_1}$, 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 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 equals 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 the Proposition 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 2.

So we have proved that the sequence $(a_n)_{n\in\mathbf{N}}$ must have finitely many prime factors. However, Corollary 2 of [Hu17] 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.

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 couple 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 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 Firstly, it is easy to check that, there is an integer r such that a_p is r-th root of unity for all but finitely many primes p (see Lemma 1 [SP11]). If $(a_n)_{n\in\mathbf{N}}$ is not strongly aperiodic, then because of Proposition 6.1 in [Fra18], 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 6, $(a_n\overline{X(n)})_{n\in\mathbf{N}}$ is a trivial Dirichlet-like character.

Proposition 8 *Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS and $X_t(n)_{n\in\mathbf{N}}$ 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

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

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

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

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

But on the other hand, 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 [AFP00]. By comparing the two sets of possible location of poles for the same function, we can see that there is at most one $a_{p_i} \neq 0$.

6 Conclusion

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

Definition Let $(a_n)_{n\in\mathbf{N}}$ be a sequence, the 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 9 *If $(a_n)_{n\in\mathbf{N}}$ is a CMAS then it is not strongly aperiodic.*

Proof From Theorem 2 in ([FH19]) and the following remark, the block complexity of sequence $(a_n)_{n\in\mathbf{N}}$ should satisfy the propriety that $\lim_{n\to\infty}\frac{p(n)}{n}=\infty$, which contradicts to the fact that the block complexity of an automatic sequence is bounded by a linear function [Cob72]. So the non-existence of strongly aperiodic CMAS.

Theorem 1 *Let $(a_n)_{n\in\mathbf{N}}$ be a CMAS, then it can be written in the 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.

7 Acknowledgement

In recent papers, we found some results in the literature on similar topics, which have applications to the classification of CMAS. In [LM18], the authors prove 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. Remarking 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 is stronger than Proposition 10 in this article by restricting the automatic sequence to the primitive case, however, this result could probably be generated to non-primitive case.

In [KM17], the authors consider general multiplicative functions with the condition $\liminf_{N\to\infty}|b_{n+1}-b_n|>0$. They prove that if $(b_n)_{n\in\mathbf{N}}$ is a completely multiplicative sequence, than most primes, at a fixed power, gives the same values as a Dirichlet character.

References

[AFP00] J.-P. Allouche, M. Mendès France, and J. Peyrière. Automatic Dirichlet series. *Journal of Number Theory*, 81(2):359 – 373, 2000.

[AG18] J.-P. Allouche and L. Goldmakher. Mock characters and the Kronecker symbol. *Journal of Number Theory*, 192:356 – 372, 2018.

[Cob72] A. Cobham. Uniform tag sequences. *Mathematical systems theory*, 6(1):164–192, Mar 1972.

[FH19] 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.

[Fra18] N. Frantzikinakis. An Averaged Chowla and Elliott Conjecture Along Independent Polynomials. *International Mathematics Research Notices*, 2018(12):3721–3743, 2018.

[Hu17] Y. Hu. Subword complexity and non-automaticity of certain completely multiplicative functions. *Advances in Applied Mathematics*, 84:73 – 81, 2017.

[KM17] O. Klurman and A. P. Mangerel. Rigidity theorems for multiplicative functions. *Mathematische Annalen*, 07 2017.

[LM18] M. Lemańczyk and C. Müllner. Automatic sequences are orthogonal to aperiodic multiplicative functions. 11 2018. preprint, http://arxiv.org/abs/1811.00594.

[Ruz77] I. Z. Ruzsa. General multiplicative functions. *Acta Arithmetica*, 32(4):313–347, 1977.

[SP11] J.-C. Schlage-Puchta. Completely multiplicative automatic functions. *Integers*, 11:8, 2011.