On multiplicative automatic sequences
Jakub Konieczny

To cite this version:

HAL Id: hal-03919573
https://hal.science/hal-03919573
Submitted on 3 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ON MULTIPLICATIVE AUTOMATIC SEQUENCES

JAKUB KONIECZNY

ABSTRACT. We show that any automatic multiplicative sequence either coincides with a Dirichlet character or is identically zero when restricted to integers not divisible by small primes. This answers a question of Bell, Bruin and Coons. A similar result was obtained independently by Klurman and Kurlberg.

1. Introduction

Automatic sequences — that is, sequences computable by finite automata — give rise to one of the most basic models of computation. As such, for any class of sequences it is natural to inquire into which sequences in it are automatic. In particular, the question of classifying automatic multiplicative sequences has been investigated by a number of authors, including [SP11], [BBC12], [Hu17], [AG18], [Li19] and [KK19a]. The interplay between multiplicative automatic sequences is studied also in [Yaz01], [SP03], [BCH14] and [LM19], among others.

The two most recent papers [Li19], [KK19a] listed above give a classification of completely multiplicative automatic sequences, but until now the question remained open for sequences which are multiplicative but not completely so. In particular, the authors of [BBCT12] conjectured that a multiplicative automatic sequence agrees with an eventually periodic sequence on the primes.

We confirm this conjecture and give some additional structural results. A similar result is also obtained in an upcoming preprint of Klurman and Kurlberg [KK19b].

Theorem 1.1. If \( a : \mathbb{N}_0 \to \mathbb{C} \) is an automatic multiplicative sequence then there exists a threshold \( p^* \) and sequence \( \chi \) which is either a Dirichlet character or identically zero such that \( a(n) = \chi(n) \) for all \( n \) not divisible any prime \( p < p^* \).

The proof naturally splits into two cases, depending on how often \( a \) vanishes. These cases are addressed in Sections 2 and 3 respectively.

Remark 1.2. Not all multiplicative sequences satisfying the conclusion of the above theorem are automatic. A full classification of automatic multiplicative sequences appears to still be out of reach of the available techniques, even if barely so. In principle, combining a slightly more precise version of this theorem discussed in subsequent sections and the classification of multiplicative periodic sequences in [LW76], we could completely classify multiplicative \( k \)-automatic sequences which vanish on all integers not coprime to \( k \). The behaviour of \( a \) on powers of primes dividing \( k \) remains problematic, as evidenced by the fact that when \( k \) is prime then for any Dirichlet character with modulus \( k^* \) and any root of unity \( \xi \), the sequence \( a_\chi(n) := \xi^{\nu_\chi(n)} \chi(n/k^{\nu_k(n)}) \) is multiplicative and \( k \)-automatic. The last sequence is a mock Dirichlet character, investigated in [BBCT12].

2010 Mathematics Subject Classification. Primary: 11B85; Secondary: 11N64, 68R15.
1.1. Basics and notion. Throughout, \( \mathbb{N} \) denotes the positive integers and \( \mathbb{N}_0 := \mathbb{N} \cup \{0\} \). A sequence \( a : \mathbb{N}_0 \to \mathbb{C} \) is multiplicative if \( a(nm) = a(n)a(m) \) for any coprime \( m, n \in \mathbb{N} \), and it is completely multiplicative the assumption of coprimality can be dropped.

For \( n \in \mathbb{N}_0 \) we let \([n] = \{1, 2, \ldots, n\}\) (in particular, \([0] = \emptyset\)). If \( p \) is a prime, \( \alpha \in \mathbb{N}_0 \) and \( n \in \mathbb{N}_0 \) then \( \nu_p(n) \) denotes the largest power of \( p \) which divides \( n \) and \( p^\alpha \mid n \) means that \( \alpha = \nu_p(n) \) (or, equivalently, that \( p^\alpha \mid n \) but \( p^{\alpha+1} \nmid n \)). If \( n, m \in \mathbb{N} \) then \( n \perp m \) is shorthand for \( \gcd(n, m) = 1 \). For two quantities \( X \) and \( Y \) we write \( X = O(Y) \) or \( X \ll Y \) if there exists an absolute constant \( c \) such that \( |X| < cY \).

We let \( \Sigma_k = \{0, 1, \ldots, k-1\} \) denote the set of digits in base \( k \). For a set \( X \) we let \( X^* \) denote the set of words over \( X \), including the empty word \( \epsilon \). If \( u \in \Sigma_k^* \) then \([u]_k \in \mathbb{N}_0 \) denotes the integer obtained by interpreting \( u \) as a digital expansion in base \( k \). Conversely, if \( n \in \mathbb{N}_0 \) then \((n)_k \in \Sigma_k^* \) denotes the expansion of \( n \) in base \( k \) without any leading zeros. More generally, for \( l \in \mathbb{N}_0 \) we let \((n)_k^l\) denote the suffix of the word \( 0^\infty(n)_k \) of length \( l \).

A sequence \( a : \mathbb{N}_0 \to \mathbb{C} \) is \( k \)-automatic if there exists finite automaton \( \mathcal{A} = (\mathcal{S}, s_0, \delta, \tau) \) which produces \( a \). Here, \( \mathcal{S} \) is a finite set of states, \( s_0 \in \mathcal{S} \) is the initial state, \( \delta \) is the transition function \( \Sigma_k \times \mathcal{S} \to \mathcal{S}, (u, s) \mapsto \delta_u(s) \), extended to a map \( \Sigma_k^* \times \mathcal{S} \to \mathcal{S} \) by \( \delta_{uv} = \delta_u \circ \delta_v \), \( \tau \) is an output function \( \tau : \mathcal{S} \to \mathbb{C} \), and finally the sequence produced by \( \mathcal{A} \) is \([u]_k \mapsto \tau(\delta_u(s_0))\), where \( u \in \Sigma_k \) does not begin with any initial zeros. A sequence is automatic if it is \( k \)-automatic for some \( k \in \mathbb{N} \).

We fix from now on the automatic multiplicative sequence \( a : \mathbb{N}_0 \to \mathbb{C} \) and an automaton \( \mathcal{A} = (\mathcal{S}, s_0, \delta, \tau) \) which produces it. It is well-known that if \( k, l \in \mathbb{N} \) are multiplicatively dependent, i.e., \( \log(k)/\log(l) \in \mathbb{Q} \setminus \{0\} \), then \( k \)-automatic sequences are the same as \( l \)-automatic sequences. Hence, we may assume without loss of generality that \( k \) is not a perfect power.

Acknowledgements. The author is grateful to O. Klurman for sharing the aforementioned preprint and for helpful comments. This research is supported by ERC grant ErgComNum 682150.

2. Sparse case

Throughout this section, make the following assumption:

(1) There exists infinitely many primes \( p \) such that \( a(p^\alpha) = 0 \) for some \( \alpha \in \mathbb{N} \).

We let \( Z \subseteq \mathbb{N}_0 \) denote the set of \( n \in \mathbb{N} \) such that \( a(n) \neq 0 \).

Proposition 2.1. The set \( Z \) is a finite union of (possibly degenerate) geometric progressions with ratio \( k^l \) \((l \in \mathbb{N}_0)\).

From here, it easily follows that \( a(p) = 0 \) for large primes. We also note that the fact that \( a \) is \( k \)-multiplicative imposes further restrictions on \( Z \) and on the behaviour of \( a \) on \( Z \). In fact, it is not hard to show that when \( k \) is composite, \( Z \) needs to be finite. For lack of other nontrivial observations we do not delve further into this subject and devote the remainder of this section to proof of the above Proposition 2.1.
2.1. Reduction to arid sets. Our first step is to show that \( Z \) is, using the terminology borrowed from [BK19], an arid set. 

**Definition 2.2.** A set \( A \subseteq \mathbb{N}_0 \) is a basic arid set of rank \( \leq r \) if it takes the form

\[
A = \left\{ \left[ u_r v_r^i u_{r-1} \ldots u_1 v_1^i u_0 \right]_k \mid l_1, \ldots, l_r \in \mathbb{N}_0 \right\}.
\]

for some \( u_0, \ldots, u_r \in \Sigma_k^* \) and \( v_1, \ldots, v_r \in \Sigma_k^* \). A set \( A \subseteq \mathbb{N}_0 \) is arid of rank \( \leq r \) if it is a union of finitely many basic arid sets of rank \( \leq r \). If \( A \subseteq \mathbb{N}_0 \), then the rank of \( A \) is the smallest \( r \) such that \( A \) is contained in an arid set of rank \( \leq r \), or \( \infty \) if no such \( r \) exists.

**Proposition 2.3** ([BK19 Proposition 3.4.]). One of the following is true:

(i) The set \( Z \) is arid.

(ii) There exists \( c \neq 0 \) and words \( w, v_1, v_2, u \in \Sigma_k^* \) such that \( |v_1| = |v_2| \), \( v_1 \neq v_2 \) and \( a([wwu]_k^*)) = c \) for all \( v \in \{v_1, v_2\}^* \).

In the following argument it will be convenient to use the notion of an IP\(^+ \) set, that is, the set of the form

\[
A = \left\{ n_0 + \sum_{i \in I} n_i \mid I \subseteq [r] \right\}.
\]

where \( n_0 \in \mathbb{N}_0 \) and \( n_1, \ldots, n_r \in \mathbb{N} \). We refer to \( n_1, \ldots, n_r \) as the sidelengths of \( A \).

**Lemma 2.4.** Let \( m \in \mathbb{Z} \) and let \( A \subseteq \mathbb{N}_0 \) be an IP\(^+ \) set with sidelengths coprime to \( m \). Then \( \#(A \mod m) \geq \min(m, r+1) \).

**Proof.** We proceed by induction on \( r \), the case \( r = 0 \) being trivial. If \( r \geq 1 \) then we can construe \( A \) as the subset \( A + \{0, u_r\} \) of an IP\(^+_r \) set \( A' \) and a two-element set. Either \( \#(A' \mod m) = m \), in which case we are done, or there is \( n \in A' \mod m \) such that \( n + n_r \notin A' \mod m \), in which case \( \#(A \mod m) \geq \#(A' \mod m) + 1 \geq r+1 \) so we are also done. \( \square \)

**Remark 2.5.** The final step uses a trivial case of the Cauchy–Davenport theorem.

**Proposition 2.6.** The set \( Z \) is arid.

**Proof.** For the sake of contradiction, suppose that \( Z \) was not arid, and let \( c \) and \( w, v_1, v_2, u \) be as in Proposition 2.3. Assumption 1 guarantees that we can find a prime power \( q = p^a \) such that \( a(q) \neq 0 \) and \( p \mid |v_1| - |v_2| \) and \( p \nmid k \). Pick \( r := p^{a+1} \) and consider the set

\[
A := \left\{ [wwu]_k^* \mid v \in \{v_1, v_2\}^* \right\}.
\]

It follows directly from the defining conditions in 2.3 that \( a(n) = c \) for all \( n \in A \). On the other hand, \( A \) is an IP\(^+ \) set with sidelengths of the form \( ([v_1]_k - [v_2]_k)k^l \) with \( l \in \mathbb{N} \), which are not divisible by \( p \). By Lemma 2.4, \( A \) covers all residues modulo \( r \). In particular, \( A \) contains an integer \( n \) exactly divisible by \( q \), whence

\[
0 \neq c = a(n) = a(n/q)a(q) = 0,
\]

which is the sought for contradiction. \( \square \)

We are now left with the task of showing that arid sets cannot be the level sets of multiplicative sequences, except for arguably trivial cases of geometric progressions whose ratio is a power of \( k \).
2.2. Base $k$ geometric progressions. While arid sets are well-adjusted for studying combinatorial properties of base $k$ expansions, in order to study arithmetic properties it is convenient to work in a slightly more general setup. We define a \textit{generalised geometric progression} of rank $\leq r$ as a set of the form

\[
A = \left\{ x_0 + \sum_{i=1}^{r} x_i k^{\alpha_i} \mid \alpha_1, \ldots, \alpha_r \in \mathbb{N}_0 \right\},
\]

where $x_0, x_1, \ldots, x_r \in \mathbb{Q}$. For the sake of uniformity, define also $\alpha_0 := 0$ Likewise, we define a \textit{restricted generalised geometric progression} of rank $\leq r$ as a set of the form

\[
A = \left\{ x_0 + \sum_{i=1}^{r} x_i k^{\alpha_i} \mid \alpha_1 \in F_1, \alpha_2 \in F_2(\alpha_1), \ldots, \alpha_r \in F_r(\alpha_1, \ldots, \alpha_{r-1}) \right\},
\]

where $x_1, \ldots, x_r \in \mathbb{Q}$ and $F_i$ are maps $\mathbb{N}_0^{i-1} \rightarrow \mathcal{P}_{\text{inf}}(\mathbb{N}_0)$ for each $i \in [r]$. Here, $\mathcal{P}_{\text{inf}}(X)$ denotes the set of all infinite subsets of a set $X$. In a fully analogous manner we define \textit{restricted arid sets} of rank $\leq r$ as sets of the form

\[
A = \left\{ [u_r v_r^l u_{r-1} \ldots u_1 v_1^l u_0]_k \mid l_1 \in F_1, l_2 \in F_2(l_1), \ldots, l_r \in F_r(l_1, \ldots, l_{r-1}) \right\},
\]

where $u_0, \ldots, u_r \in \Sigma_k^*$ and $v_1, \ldots, v_r \in \Sigma_k^*$ and $F_i$ are like above.

Given sequences $F_i$ as in (5), let us call a vector $(\alpha_0, \alpha_1, \ldots, \alpha_s)$ \textit{admissible} if $\alpha_0 = 0$ and $\alpha_i \in F_i(\alpha_1, \ldots, \alpha_{i-1})$ for $i \in [s]$, $0 \leq s \leq r$. The elements of the restricted generalised geometric progression $A$ given by (5) can naturally be indexed by the leaves of a regular rooted tree with vertex degree $\infty$, whose vertices are admissible sequences $(\alpha_0, \ldots, \alpha_s)$, whose root is $(0)$ and whose edges are given by $(\alpha_0, \ldots, \alpha_s) \rightarrow (\alpha_0, \ldots, \alpha_s, \alpha_{s+1})$. It follows from basic Ramsey theory that if the leaves of such a tree are coloured by finitely many colours then there exists an infinite regular subtree of depth $r$ with monochromatic leaves. As a consequence, restricted generalised geometric progressions of a given rank are partition regular.

The same observation, mutatis mutandis, applies to restricted arid sets.

It is clear that any (restricted) arid set is a (restricted) generalised geometric progression. The following lemma provides a partial converse to this statement.

\textbf{Lemma 2.7.} For any $x_0, x_1, \ldots, x_r \in \mathbb{Q}$ there exists $B \in \mathbb{N}$ and $C > 0$ such that the following holds. Suppose that $\alpha_1, \ldots, \alpha_r \in \mathbb{N}_0$ is a sequence such that

\[
x_0 + \sum_{i=1}^{r} x_i k^{\alpha_i} \in \mathbb{N}
\]

and $\alpha_i \geq \alpha_{i-1} + C$ for all $i \in [r]$, where $\alpha_0 := 0$. Then there exist words $u_0 \in \Sigma_k^B$, $u_1, \ldots, u_r \in \Sigma_k^{3B}$ and $v_1, \ldots, v_r \in \Sigma_k^B$ such that

\[
x_0 + \sum_{i=1}^{r} x_i k^{\alpha_i} = [u_r v_r^l u_{r-1} \ldots u_1 v_1^l u_0]_k,
\]

where for $i \in [r]$ the lengths $l_i \in \mathbb{N}_0$ are uniquely determined by

\[
0 \leq \alpha_i - B \left( \sum_{j=1}^{i} l_j + 3i - 1 \right) < B.
\]
If additionally \( x_i \neq 0 \) for all \( i \in [r] \) then additionally the expansion in (8) is nondegenerate in the sense that \( u_r \neq 0^B \) and there is no \( i \in [r] \) such that \( v^3_i = u_i = v^3_{i-1} \).

Proof. This follows by inspection of the standard long addition procedure. Suppose first that \( x_0, x_1, \ldots, x_r \) were all positive integers. Then the conclusion would hold with \( u_i = 0^B \) and \( u_i = 0^*(x_i)_k0^* \), where 0* denotes an unspecified string of zeros. If we drop the assumption of positivity, then the same conclusion holds except \( v_i \) can also take the form \((k-1)^Bi \) and \( u_i \) need to be modified accordingly. Finally, if \( x_i \) are rational then apply the above reasoning to the sequence \( Mx_0, Mx_1, \ldots, Mx_r \) where \( M \) is multiplicatively rich enough that the latter sequence consists of only integers, and use the fact that division by \( M \) takes periodic digital expansions to periodic digital expansions.

\[ \square \]

Remark 2.8. (i) The definition of \( l_i \) in (9) is arranged so that if the right hand side of (8) is construed as the \( B \)-block expansion of the sum on the left hand side (with each \( v_i \) occupying one block and \( u_i \) occupying three blocks) then the position \( \alpha_i \) falls in the middle block of \( u_i \) for all \( i \in [r] \).

(ii) The constant \( B \) can be replaced by any multiple, and the constant \( C \) can always be enlarged. We could have required that \( B = C \), but we believe that would decrease the intuitive appeal of the result.

Our definition rank guarantees that if \( A \) is a (restricted) arid set of rank \( \leq r \) then \( \text{rank} \ A \leq r \). It follows from Lemma 2.7 that if \( A \subseteq \mathbb{N}_0 \) is a (restricted) generalised geometric progression of rank \( \leq r \) then also \( \text{rank} \ A \leq r \). In order to estimate rank of certain sets from below, we record some observations which essentially say that the above inequalities cannot be strict except for degenerate cases. In fact, the following lemma would suffice.

Lemma 2.9. Let \( B \in \mathbb{N} \) and let \( u_0 \in \Sigma^B_k, u_1, \ldots, u_r \in \Sigma^3B_k \) and \( v_1, \ldots, v_r \in \Sigma^B_k \) be nondegenerate in the sense of Lemma 2.7 and let \( A \) be the corresponding arid set given by (9). Then \( \text{rank} \ A = r \).

This result is elementary and there are several ways to prove it. One possibility is to run a combinatorial argument relying on the intuition that given \( n = [u_0, u_1, u_2, \ldots, u_{r-1}, u^B_0, u^B_0] \) it is essentially possible to recover \( r, u_0, \ldots, u_r, v_1, \ldots, v_r \) and \( l_1, \ldots, l_r \), except for some basic operations which do not change the rank and except for some border cases corresponding to small \( l_i \). Another option is to use estimates on the rate of growth of arid sets. We choose yet another route, and derive Lemma 2.9 from the following stronger statement, whose proof we delegate to the appendix.

Lemma 2.10. Let \( A \) be a restricted generalised geometric progression of rank \( \leq r \) given by (9) with \( x_1, \ldots, x_r \neq 0 \). Then \( \text{rank} \ A = r \).

2.3. Multiplication and arid sets. Recall that the set \( Z \) of nonzero places of \( a \) is closed under products of coprime elements. More generally, if \( n, m \in Z \) and \( d \in \mathbb{N} \) is such that \( d \mid n, n/d \perp n \) and \( n/d \perp m \) then also \( mn/d \in Z \). This motivates the interest in the following lemma.

Lemma 2.11. For any \( u, v, w \in \Sigma^B_k \) with \([u]_k, [w]_k \neq 0 \) there exists \( D \in \mathbb{N} \) such that the following is true. For any prime \( p \) there exists \( Q \in \mathbb{N} \) such that if \( l \in \mathbb{N} \) and \( Q \mid l \) then \( \nu_p ([wv'u]_k) \leq \nu_p (D) \).
Proof. For reasons which will become clear in the course of the argument, we will take $D := D_0 D_1$ where $D_1 := [wv]_k$ and $D_0 := [k^{wv} - (k^{v} - 1)[u]]_k$. Note that $D_1 \neq 0$ since $[w]_k \neq 0$ and $D_0 \neq 0$ since $[k^{wv} - 1][u]_k \equiv 0 \mod{k^{wv}}$. The argument splits into two cases, depending on whether $p$ divides $k$. Note that in full generality we have

$$[wv]_k = [w]_k k^{[wv]+[u]} + [v]_k k^{[wv]} - 1 - [u]_k.$$

Suppose first that $p \nmid k$. For any $\omega \in \mathbb{N}_0$ there exists $Q_\omega$ such that

$$[wv]_k \equiv [w]_k k^{[u]} + [v]_k k^{[w]} - 1 - [u]_k \not\equiv 0 \mod{p^\omega}$$

for all $l \in \mathbb{N}$ divisible by $Q_\omega$. In particular, letting $\omega > \nu_p(D_1)$ we conclude that

$$\nu_p([wv]_k) \leq \nu_p(D_1) \leq \nu_p(D)$$

for all $l$ divisible by $Q_\omega$.

Secondly, suppose that $p | k$. Then for any $\omega \in \mathbb{N}_0$ there exists $Q_\omega$ such that

$$(k^{wv} - 1)[wv]_k \equiv -[v]_k k^{[u]} + (k^{wv} - 1)[u]_k \equiv \pm D_0 \not\equiv 0 \mod{p^\omega}$$

for all $l \in \mathbb{N}$ with $l \geq Q_\omega$ (which in particular holds if $Q_\omega | l$). Letting $\omega > \nu_p(D_0)$ we conclude that

$$\nu_p([wv]_k) \leq \nu_p(D_0) \leq \nu_p(D)$$

for all $l$ divisible by $Q_\omega$. \qed

To simplify notation in the following result, for $n, m \in \mathbb{N}$ let $\text{gcd}(m^{\infty}, n)$ denote the limit $\lim_{n \to \infty} \text{gcd}(m^n, n)$, or equivalently the product $\prod_{p \mid \text{gcd}(m^n, \nu_p(n))} p^{\nu_p(n)}$. Note we do not attribute any independent meaning to the symbol $m^{\infty}$ outside of $\text{gcd}$. It follows directly from Lemma 2.11 that, with notation therein, for each $m \in \mathbb{N}$ there exists an integer $Q$ such that for any $l \in \mathbb{N}$ divisible by $Q$ we have $\text{gcd}(m^{\infty}, [wv]_k) | D$.

**Proposition 2.12.** Let $u, v, w \in \Sigma^*$, $v \neq \epsilon$, and $[u]_k, [w]_k \neq 0$ or $[v]_k \neq 0$. Then there exists $l \in \mathbb{N}_0$ such that $[wv]_k \not\in Z$.

Proof. For the sake of contradiction suppose that $[wv]_k \in Z$ for all $l \in \mathbb{N}_0$. Replacing $w$ and $u$ with $wv$ and $vu$, we may assume that $[u]_k, [w]_k \neq 0$. Let $t \in \mathbb{N}$ be a large parameter. Our strategy is to show that the elements of $Z$ which can be constructed taking products of $t$ terms of the form $[wv]_k$ ($l \in \mathbb{N}_0$) give rise to an arid set of rank $\geq 2^t - 1$, which leads to contradiction since $Z < \infty$.

For $l \in \mathbb{N}_0$ let $n(l) := [wv]_k$. It follows from Lemma 2.11 that there exists $D \in \mathbb{N}$ such that for any $m \in \mathbb{N}$ there exist $Q \in \mathbb{N}$ such that if $l \in \mathbb{N}$ and $Q | l$ then $\text{gcd}(m^{\infty}, n(l)) | D$. Using this observation iteratively we can find a sequence of infinite sets $F_1, F_2(l_1), F_3(l_1, l_2), \ldots, F_t(l_1, \ldots, l_t) \subseteq \mathbb{N}_0$ such that for any sequence $l_1, \ldots, l_t \in \mathbb{N}_0$ with $l_i \in F_i(l_1, \ldots, l_{i-1})$ for all $i \in [t]$ we have

$$d_i := \text{gcd} \left( \prod_{j=1}^{i-1} n(l_j) \right) | D$$

for all $i \in [t]$. Using partition regularity, we may further assume that $1 = d_1, d_2, \ldots, d_t$ are independent of the choice of $l_1, \ldots, l_t$. Hence, for any admissible $l_1, \ldots, l_t \in \mathbb{N}_0$ we have $d_i | n(l_i), n(l_j)/d_i \perp d_i$ and $n(l_i)/d_i \perp n(l_j)/d_j$ for all $i, j \in [t]$ with $i \neq j$. Hence, $Z$ contains the set

$$A := \left\{ \prod_{i=1}^{t} n(l_i)/d_i \left| \begin{array}{l} l_1 \in F_1, \ldots, l_t \in F(l_1, \ldots, l_{t-1}) \end{array} \right. \right\}.$$
It is clear from the definition of $n(l)$ that there exists $z, y \in \mathbb{Q} \setminus \{0\}$ and $c \in \mathbb{N}$ such that $n(l) = zk^d + y$. Hence, for any admissible $l_1, \ldots, l_t$ we can expand the product in the definition of $A$ as

$$\prod_{i=1}^{t} n(l_i)/d_i = \sum_{I \subseteq [t]} x_I k^{n_I},$$

where $x_I = x^{|I|} y^{|I|-|I|} \neq 0$ and $n_I = \sum_{i \in I} l_i$ (in particular, $n_s = 0$). Let $s := 2^t - 1$. We may identify $\{0, 1\}^t$ with $\{0, 1, \ldots, s\}$ in a standard way. Replacing $F_I$ with smaller sets if necessary we may assume that the sequence $\{\alpha_j\}_{j=0}^{s-1}$ is increasing, and indeed that $\alpha_j' > \alpha_{j-1}' + C$ for all $j \in [s]$ where $C > 0$ is an arbitrary constant. In particular, letting $C'$ and $B$ be the constants from Lemma 2.7 we conclude that there exists words $u_0 \in \Sigma_B^C$, $u_1, \ldots, u_s \in \Sigma_B^C$ and $v_1, \ldots, v_s \in \Sigma_B^C$ obeying the nondegeneracy conditions $[u_s] \neq 0$ and $\# \{v_j, u_j^3, v_{j-1} \} > 1$ for all $j \in [s]$, such that

$$\prod_{i=1}^{t} n(l_i)/d_i = x_0 + \sum_{j=1}^{s} x_j k^{n_j} = [u_s v_s^m u_{s-1} \ldots u_1 v_1^m u_0]_k,$$

where $m_1, \ldots, m_s \in \mathbb{N}$ obey the asymptotic relation $m_j = (\alpha_j - \alpha_{j-1})/B + O(1)$ for all $j \in [s]$. Note that we could assume that $u_j, v_j$ are independent of $l_1, \ldots, l_t$ by partition regularity. By the same token, we may also assume that $\overline{m}_j := m_j \text{ mod } M$ ($j \in [s]$) are independent of $l_1, \ldots, l_t$, where $M$ is a multiplicatively rich constant such that $\delta_{v}^{M}$ is idempotent for each $v \in \Sigma_k$ (we can take $M = \#S!$). If now follows that $Z$ contains the arid set

$$\{[u_s v_s^m u_{s-1} \ldots u_1 v_1^m u_0]_k \mid m_j \in \overline{m}_j + MN \text{ for all } j \in [s]\},$$

whose rank is equal to $s$. In particular, rank $Z \geq s$, as needed. \hfill \Box

**Corollary 2.13.** The set $Z$ is a finite union of geometric progressions of the form $\{zk^d \mid l \in \mathbb{N}_0\}$ where $x \in \mathbb{N}_0$ and $c \in \mathbb{N}$.

**Proof.** It is enough to notice that the only basic arid sets not containing patterns forbidden by Proposition 2.12 take the form $\{[w] \mid l \in \mathbb{N}_0\}$ with $c \in \mathbb{N}_0$. \hfill \Box

3. Dense case

We now assume that there exists a threshold $p_0$ such that the following holds:

(10) \hspace{1cm} \text{For all primes } p \geq p_0 \text{ and all } \alpha \in \mathbb{N} \text{ we have } a(p^\alpha) \neq 0.

Our main aim is to show that $a(n)$ coincides with a Dirichlet character for $n$ devoid of small prime factors. We also record some observations concerning the behaviour of $a$ on small primes.

3.1. Large primes. We first deal with large primes. From this point onwards, we let $m, \chi$ and $p_1$ denote the objects in the following result; we may assume that $p_1 > m$ and that $p_1 > k$.

**Proposition 3.1.** There exists a Dirichlet character $\chi$ with modulus $m$ and a threshold $p_1$ such that $a(n) = \chi(n)$ for all $n \in \mathbb{N}$ which are products of primes $\geq p_1$. 

Relying on the following result, we can prove Proposition 3.1 by essentially the same methods which were used by Klurman and Kurlberg [KK19a] for completely multiplicative sequences. We will also need the fact that the \( k \)-kernel of any \( k \)-automatic sequence is finite. Here, the \( k \)-kernel of \( a \) is the set of all the sequences \( n \mapsto a(k^\alpha n + r) \) with \( \alpha \in \mathbb{N}_0 \) and \( 0 \leq r < k^\alpha \).

**Proposition 3.2.** There exists a threshold \( p_2 \) such that if \( p \geq p_2 \) is a prime then \( a(p^\alpha) = a(p)^\alpha \neq 0 \) for all \( \alpha \in \mathbb{N} \).

**Proof of Prop. 3.1 assuming Prop. 3.2.** Suppose for the sake of clarity that \( p_2 \geq p_0 \). Because the \( k \)-kernel of \( a \) is finite, we can find integers \( \beta < \gamma \) such that \( a(k^\beta n + 1) = a(k^\gamma n + 1) \) for all \( n \in \mathbb{N} \). Denote by \( \tilde{a} \) the multiplicative function given by \( \tilde{a}(p^\alpha) = a(p^\alpha) = a(p)^\alpha \) if \( p \geq p_2 \) and \( \tilde{a}(p^\alpha) = 1 \) if \( p < p_2 \), \( \alpha \in \mathbb{N} \). Clearly, \( \tilde{a} \) is totally multiplicative. Because \( a \) takes on finitely many values, \( a(p) \) is a root of unity for each \( p \geq p_2 \). Letting \( Q \) be the product of all primes \( < p_2 \) we obtain

\[
\tilde{a}(k^\beta Qn + 1) = a(k^\beta Qn + 1) = a(k^\gamma Qn + 1) = \tilde{a}(k^\gamma Qn + 1)
\]

for all \( n \in \mathbb{N}_0 \), whence \( \tilde{a} \) is a Dirichlet character by [EK17, Thm. 2]. \( \square \)

In order to prove Proposition 3.2 it will be convenient to introduce an equivalence relation on \( \Sigma_k^* \) where \( u \sim v \) if \( |u| = |v| \) and words \( u, v \) give rise to the same transition function in the automaton \( A \), that is, \( \delta_u = \delta_v \). Since transition functions are self-maps of \( S \), the number of equivalence classes \( # (\Sigma_k^*/\sim) \) is bounded uniformly with respect to \( l \in \mathbb{N} \). Likewise, consider the equivalence relation on \( \mathbb{N}_0 \) given by \( n_1 \sim n_2 \) if \( (n_1)_k^l \sim (n_2)_k^l \) for all sufficiently large \( l \in \mathbb{N} \), or — equivalently — if \( (n_1)_k^l \sim (n_2)_k^l \) for at least one \( l \in \mathbb{N} \) with \( n_1, n_2 < k^l \). Crucially, there are only finitely many equivalence classes: \( # (\mathbb{N}_0/\sim) < \infty \).

**Lemma 3.3.** There exists a threshold \( p_3 \) such that for any \( p > p_3 \) there exists a pair \( n_1, n_2 \in \mathbb{N}_0 \) with \( n_1 \neq n_2 \) (mod \( p \)) such that \( n_1 \sim n_2 \) and \( pn_1 \sim pn_2 \) for all sufficiently large \( l \in \mathbb{N} \).

**Proof.** Since \( # (\mathbb{N}_0/\sim) < \infty \), this follows from the pigeonhole principle. \( \square \)

For the sake of brevity, in the following argument and elsewhere we will say that a statement \( \varphi(n) \) is true for almost all \( n \) if the set of \( n \) for which it fails has Banach density 0:

\[
\lim_{N \to \infty} \sum_{M \geq 0} \frac{1}{N} \# \{ n \in [M, M + N) \mid \neg \varphi(n) \} = 0.
\]

**Proof of Prop. 3.2**. Take any prime \( p \) with \( p > p_3 \), where \( p_3 \) is the threshold in Lemma 3.3 and let \( n_1, n_2 \) be the pair whose existence is guaranteed by said lemma. Let \( l \) be a large integer, to be determined in the course of the argument, and put \( u_i := (n_i)_k^l \), \( u_i' := (pn_i)_k^l \). It is a general fact that if \( w \in \Sigma_k^* \) the for almost all \( n \) the expansion \( (n)_k \) contains \( w \). Hence, for almost all \( n \) there exists a decomposition

\[
(n)_k = x_n y_n u_n,
\]

for some \( x_n, y_n \in \Sigma_k^* \) where \( x_n \) is nonempty and does not start with any zeros and \( y_n \) starts with at least \( l \) zeros. Letting \( x_n' \) and \( y_n' \) denote the expansions of \( p|x_n|_k \) and \( p|y_n|_k \) with \( x_n' \) not starting with any zeros and \( |y_n'| = |y_n| \), we get the decomposition

\[
(pn)_k = x_n' y_n u'_n.
\]
If \( p \nmid n \) then clearly \( a(pn) = a(p)a(n) \). On the other hand, if \( p \mid n \) then
\[
a(pn) = a([x_n u_1 y_n]_k) = a([x_n u_2 y_n]_k) = a(p)a([x_n u_2 y_n]_k) = a(p)a([x_n u_1 y_n]_k) = a(p)a(n).
\]
Hence, we have shown that \( a(pn) = a(p)a(n) \) for almost all \( n \).

Let \( \alpha \in \mathbb{N} \). Integers \( n \) such that \( p^\alpha \parallel n \) and \( n \perp q \) for all \( q < p_0 \) constitute a positive proportion of all integers, whence there exists many \( n \) such that
\[
a(p^{\alpha+1}) = \frac{a(pm)}{a(n/p^\alpha)} = \frac{a(p)a(n)}{a(n/p^\alpha)} = a(p)a(p^\alpha).
\]
It now follows by induction that \( a(p^\alpha) = a(p)^\alpha \). \( \square \)

3.2. Small primes. In this section we address the behaviour of \( a \) on small primes. Unfortunately, we can only obtain a weaker analogue of Proposition 3.1.

**Proposition 3.4.** For any prime \( p \nmid k \), the sequence \( a(p^\alpha) \) is eventually periodic.

**Proof.** Recall that there exists (many) pairs of distinct integers \( n_1, n_2 \in \mathbb{N}_0 \) such that \( n_1 \sim n_2 \). Note also that if \( n_1 \sim n_2 \) and \( n_1' \sim n_2' \) then also \( k^\alpha n_1 + n_1' \sim k^\alpha n_2 + n_2' \) for sufficiently large \( \alpha \). Hence, we can assume that \( d := n_1 - n_2 \) is divisible by any prime \( p < p_1 \) and also by a large power of \( k \). Let \( v_1 = (n_1)_k \) and \( v_2 = (n_2)_k \) where \( l \) is a large integer.

For any \( \alpha \in \mathbb{N} \) and any \( \beta \) sufficiently large in terms of \( \alpha \), there exists a prime \( q \) such that \( (qp^\alpha)_k \in 1v_1 \Sigma^\beta_k \), that is, the expansion of \( qp^\alpha \) starts with \( 1v_1 \) and contains \( \beta \) other digits. (This follows from the classical fact for any \( \varepsilon > 0 \) and any sufficiently large \( N \), there exists a prime between \( N \) and \( N + \varepsilon N \); in fact, by the Prime Number Theorem there are \( \sim \varepsilon N / \log N \) such primes.) Let \( \delta = \nu_p(d) \) and suppose that \( \alpha > \delta \). Then
\[
a(p^\alpha) = a(qp^\alpha)/a(q) = a(qp^\alpha + dk^\delta)/a(q)
\]
\[
= a(qp^{\alpha-\delta} + dk^\delta/p^\delta)a(p^\delta)/a(q) = \chi(qp^{\alpha-\delta} + dk^\delta/p^\delta)a(p^\delta)/\chi(q),
\]
where in the last transition we use the fact that any prime \( < p_1 \) divides exactly one of \( qp^{\alpha-\delta} \) and \( (d/p^\delta)k^\delta \). We may also assume (using the Prime Number Theorem in arithmetic progressions) that \( q \equiv 1 \mod m \) and \( dk^\delta/p^\delta \equiv d/p^\delta \mod m \), whence
\[
a(p^\alpha) = \chi(p^{\alpha-\delta} + d/p^\delta)a(p^\delta).
\]
It remains to notice that the expression on the right hand side is periodic in \( p^\alpha \). \( \square \)

**Corollary 3.5.** There exists a periodic sequence \( b: \mathbb{N}_0 \to \mathbb{C} \) and threshold \( n_0 \) such that \( a(n) = b(n) \) for all \( n \geq n_0 \) coprime to \( k \).

**Proof.** Partitioning \( \mathbb{N}_0 \) into arithmetic progressions, we may assume that for each prime \( p < p_1 \), either \( n \) is divisible by a large power of \( p \) or \( n \) is not divisible by \( p \). Repeating the same reasoning as in Proposition 3.4 we conclude that
\[
a(n) = \chi \left( (n + d)/\prod_p p^{\delta_p} \right) a(p^{\delta_p}),
\]
where \( \delta_p = \nu_p(d) \) and the product runs over all primes \( p < p_1 \) with \( p \nmid k \). \( \square \)
Appendix A. Rank of Generalised Geometric Progressions

Lemma A.1. For any \( x_1, \ldots, x_r \in \mathbb{Q} \) there exists a constant \( C \geq 0 \) such that the following is true. Let \( \alpha_1, \ldots, \alpha_r \in \mathbb{N}_0 \) and suppose that

\[
\sum_{i=1}^r x_i k^{\alpha_i} = 0.
\]

Then there exists a partition \([r] = I_1 \cup \cdots \cup I_s\) and \( \gamma_1, \ldots, \gamma_s \in \mathbb{N}_0 \) such that for each \( i \) there exists a partition described above. Then there exists a sequence \( (\alpha_1', \ldots, \alpha_r') \) for all \( n \) and \( |\gamma_j - \alpha_i| < C \) for all \( i \in I_j, j \in [s] \).

Proof. We proceed by induction on \( r \), the case \( r = 1 \) being trivial. Suppose that the claim was false and let \( C(\alpha_1, \ldots, \alpha_r) \) denote the smallest value of \( C \) such that there exists a partition described above. Then there exists a sequence \( (\alpha_1', \ldots, \alpha_r') \) for all \( n \) such that \( C(\alpha_1', \ldots, \alpha_r') \to \infty \) as \( n \to \infty \).

Put \( \gamma_n := \max_{i \in [r]} \alpha_i' \). Passing to a subsequence if necessary, we may assume that for each \( i \in [r] \), the limit \( \delta_i := \lim_{n \to \infty} (\gamma_n - \alpha_i') \) exists in \( \mathbb{N}_0 \cup \{\infty\} \). Let \( I_s \) be the set of those \( n \) for which \( \delta_i < \infty \). Note that if \( i \in I_s \) then \( \alpha_i' = \gamma_n - \delta_i \) for all sufficiently large \( n \); passing to a subsequence again we may assume that the above holds for all \( n \). It follows that

\[
\sum_{i \in I_1} x_i k^{\alpha_i} = k^{\gamma_n} \lim_{m \to \infty} \sum_{i=1}^r x_i k^{\alpha_i} - \gamma_n = 0
\]

for all \( n \). Rearranging the indices if necessary, we may assume without loss of generality that \( I_1 \) is the terminal segment \([r' + 1, \ldots, r]\). The above construction guarantees that

\[
\sum_{i=1}^{r'} x_i k^{\alpha_i} = 0
\]

for all \( n \). Hence, by the inductive assumption there exists a constant \( C' \) such that for each \( n \) there exists a decomposition \([r'] = I_1 \cup \cdots \cup I_{s'}\) and \( \gamma_1', \ldots, \gamma_{s'}' \in \mathbb{N}_0 \) such that

\[
\sum_{i \in I_j} x_i k^{\alpha_i} \text{ for all } j \in [s'],
\]

and \( |\gamma_j' - \alpha_i'| < C' \) for all \( i \in I_j, j \in [s'] \). Passing to a subsequence again, we may assume that the decomposition \( I_1, \ldots, I_{s'} \) is independent of \( n \). Letting \( s = s' + 1 \), \( I_s = I_{s'} \) and \( \gamma_s = \gamma_{s'} \) we conclude that

\[
C(\alpha_1, \ldots, \alpha_n) \leq \max \{ \{ \delta_i \mid i \in I_s \} \cup \{ C' \} \}.
\]

This contradicts the choice of \( \alpha_1', \ldots, \alpha_r' \) and finishes the argument.

Proof of Lemma 2.9. By partition regularity, it will suffice to show that \( A \) is not contained in any generalised geometric progression \( B \) of rank \( \leq s < r \). For the sake of contradiction suppose otherwise and write \( B \) as

\[
B = \left\{ y_0 + \sum_{j=1}^s y_j k^{\beta_j} \mid \beta_1, \ldots, \beta_s \in \mathbb{N}_0 \right\}.
\]

Assume also that \( s \) is as small as possible. For each admissible sequence \( \bar{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_r) \) pick \( \bar{\beta}(\bar{\alpha}) = (\beta_0(\bar{\alpha}), \ldots, \beta_s(\bar{\alpha})) \) such that

\[
\sum_{i=0}^r x_i k^{\alpha_i} = \sum_{j=0}^s y_j k^{\beta_j(\bar{\alpha})}.
\]
By Lemma [A.1], there exists a constant $C > 0$ such that for each admissible $\alpha$ there exists a partition $[-s, r] = I_1 \cup I_2 \cup \ldots I_t$ such that for each $l \in [t]$ we have
\[\sum_{i \in I_l \cap \mathbb{N}_0} x_i k^{\alpha_i} = \sum_{j \in -I_l \cap \mathbb{N}_0} y_j k^{\beta_j(\alpha)},\]
and the diameter of the set $\{\alpha_i \mid i \in I_l \cap \mathbb{N}_0\} \cup \{\beta_j \mid j \in -I_l \cap \mathbb{N}_0\}$ is $\leq C$. Using partition regularity again, we may assume that the partition $I_1, \ldots, I_t$ is independent of $\alpha$.

Since the sets $F_i(\alpha_1, \ldots, \alpha_{i-1})$ are infinite, we can always assume that $|\alpha_i - \alpha_{i'}| > C$ for all $i \neq i'$. In particular, $|(I_l \cap \mathbb{N}_0)| \leq 1$ for each $l \in [t]$. On the other hand, if $I_l \cap \mathbb{N}_0 = \emptyset$ for some $l \in [t]$ then $A$ would be contained in the generalised geometric progression of rank $s$ obtained by setting $\beta_j = 0$ for $j \in -I_l \cap \mathbb{N}_0$, which contradicts the choice of $s$. Hence, $|(I_l \cap \mathbb{N}_0)| = 1$ for each $l \in [t]$. Since $x_i \neq 0$ for $i \in [r]$ it follows that $#(-I_l \cap \mathbb{N}_0) \geq 1$ for each $l \in [t]$. This contradicts the assumption that $s < r$ and finishes the argument.

\[\square\]

**Remark A.2.** A careful inspection of the above proof shows that the restricted generalised geometric progression $A$ given by (5) is contained in the generalised geometric progression $B$ given by (13) if and only if there exists a partition $[0, r] = J_0 \cup J_1 \cup \ldots J_s$ and integers $0 = \delta_0, \delta_1, \ldots, \delta_s$ such that $x_i = \sum_{j \in J_i} y_j k^{\delta_j(i)}$ for each $i \in [0, r]$.

### References


(J. Konieczny) **Einstein Institute of Mathematics Edmond J. Safra Campus, The Hebrew University of Jerusalem Givat Ram. Jerusalem, 9190401, Israel**

**Faculty of Mathematics and Computer Science, Jagiellonian University in Kraków, Lojasiewicza 6, 30-348 Kraków, Poland**

*Email address: jakub.konieczny@gmail.com*