STUDIA MATHEMATICA 231 (1) (2015)

Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture

by

JEAN-MARC DESHOUILLERS (Bordeaux), MICHAEL DRMOTA (Wien) and CLEMENS MÜLLNER (Wien)

Abstract. We prove that automatic sequences generated by synchronizing automata satisfy the full Sarnak conjecture. This is of particular interest, since Berlinkov proved recently that almost all automata are synchronizing.

1. Introduction. In 2009, Peter Sarnak stated the following conjecture [17]:

CONJECTURE 1.1. *Let μ be the Möbius function. For any sequence $\xi(n)$ observed by a deterministic flow (X,T),*

$$\sum_{n\le N}\xi(n)\mu(n)=o(N). \tag{1.1}$$

There are some classes of functions for which the Sarnak conjecture has already been proved, for example periodic sequences, quasiperiodic sequences [7], nilsequences [12] and horocycle flows [5].

The purpose of this paper is to add a further class of sequences $\xi(n)$ that fulfill the Sarnak conjecture, namely automatic sequences that are generated by synchronizing automata.

Let us be more precise. Suppose that $\mathcal{A}$ is a complete deterministic finite automaton with output (DFAO) with input alphabet $\Sigma=\{0,\dots,k-1\}$, transition function δ, state set Q with the initial state q_0 and output mapping $\tau: Q\to\Delta$. We say a sequence $(a_n)_{n\in\mathbb{N}}$ is *automatic* if there exists a DFAO $\mathcal{A}$ such that $a_n=\tau(\delta(q_0,\mathbf{w}_n))$ for all $n\ge 0$, where $\mathbf{w}_n$ is the representation of n in base k. As in the definition given in [1, p. 152], we assume the input starts with the most significant digit.

2010 *Mathematics Subject Classification*: Primary 11B85; Secondary 37B10, 11L20.
Key words and phrases: Sarnak conjecture, automatic sequences, synchronizing automata.
Received 11 December 2015; revised 22 January 2016.
Published online 10 February 2016.

DOI: 10.4064/sm8479-2-2016

A DFAO $\mathcal{A}$ is called *synchronizing* if there exists a *synchronizing word* $\tilde{\mathbf{w}} \in \Sigma^*$ whose action resets $\mathcal{A}$, i.e. $\tilde{\mathbf{w}}$ leaves the automaton in one specific state, no matter which state in Q it is applied to: $\delta(q, \tilde{\mathbf{w}}) = \delta(q_0, \tilde{\mathbf{w}})$ for all $q \in Q$. Note that the output of a synchronizing automaton for an input word only depends on the last occurrence of the synchronizing word and the part thereafter.

We fix a synchronizing DFAO $\mathcal{A}$ and denote by $\mathbf{a}$ the corresponding automatic sequence $(a_n)_{n\in\mathbb{N}}$. We denote by T the shift operator on the sequences in Δ. We then define $X := \overline{\{T^n(\mathbf{a}) : n \in \mathbb{N}_0\}}$. Since the subword complexity of an automatic sequence is at most linear, it follows that the topological entropy of the dynamical system (X, T) is zero. Thus, any automatic sequence (as well as any other sequence observed by (X, T)) is observed by a deterministic flow. It is therefore natural to study the Sarnak conjecture for automatic sequences and their related dynamical system.

Actually a Möbius randomness law has already been established for the Thue–Morse sequence [6, 15], the Rudin–Shapiro sequence [16] and sequences related to invertible automata [8, 10].

As already mentioned, the main focus of this paper is to study automatic sequences that are generated by synchronizing automata. It is worth mentioning that Berlinkov [2, 3] established a result showing that "almost all" automata are synchronizing. In particular this means that we are studying —more or less—almost all automatic sequences and only *exceptional cases* are not covered. (By the way, the Thue–Morse sequence as well the Rudin–Shapiro sequence are *exceptional* from this point of view; this is due to the symmetry of their generating automata as described in [1, pp. 153–154]: for each of them there exist two states q_0 and q_1 such that $\delta(q_0, \mathbf{w}) \neq \delta(q_1, \mathbf{w})$ for any word $\mathbf{w}$.) Nevertheless the following general result can be proved by elementary methods.

Theorem 1.2. *Let $h : \mathbb{N} \to \mathbb{C}$ be a bounded function fulfilling*

$$\sum_{n\le N} h(an+b) = o(N) \quad \text{for all positive integers } a, b. \tag{1.2}$$

Let, furthermore, (X, T) be the dynamical system related to an automatic sequence $(a_n)_{n\in\mathbb{N}}$ that is generated by a synchronizing automaton, where the input starts with the most significant digit of n. Then for all sequences $\xi(n) := f(T^n(x))$ (with $x \in X$ and $f \in C(X, \mathbb{C})$) we have

$$\sum_{n\le N} \xi(n)h(n) = o(N).$$

This implies the full Sarnak conjecture for this dynamical system.

In Section 3 we address the case where the sequence under consideration is produced by a synchronizing automaton, where the input starts with the least significant digit. Finally, in Section 4 we discuss Gelfond-like problems for sequences that are generated by synchronizing automata.

2. Proof of Theorem 1.2. We recall that we work with a fixed synchronizing DFAO $\mathcal{A}$ and denote by $\mathbf{a}$ the corresponding automatic sequence $(a_n)_{n\in\mathbb{N}}$. We denote by T the shift operator on the sequences in Δ and define $X := \overline{\{T^n(\mathbf{a}) : n \in \mathbb{N}_0\}}$. In this case we can use the metric $d(x,y) = \sum_{n=0}^{\infty} 2^{-n-1} d_n(x_n, y_n)$ on X, where d_n denotes the discrete metric on Δ.

First we present a simple lemma that reduces the Sarnak conjecture for the dynamical system (X,T) (i.e. the full Sarnak conjecture for $\mathbf{a}$) to a Möbius random law on the corresponding automatic sequence $\mathbf{a} = (a_n)_{n\in\mathbb{N}}$.

LEMMA 2.1. *Suppose that for every $j \geq 1$ and for every function $g : \Delta^j \to \mathbb{C}$ we have*

$$\sum_{n\leq N} g(a_{n+\ell}, \dots, a_{n+\ell+j-1}) h(n) = o(N),$$

uniformly for $\ell \in \mathbb{N}$, where $h : \mathbb{N} \to \mathbb{C}$ satisfies (1.2)*. Then also*

$$\sum_{n\leq N} \xi(n) h(n) = o(N)$$

for all sequences of the form $\xi(n) = f(T^n x)$, where $x \in X$ and $f \in C(X, \mathbb{C})$.

Proof. Let $f \in C(X,\mathbb{C})$ and $\varepsilon > 0$ and suppose that $h : \mathbb{N} \to \mathbb{C}$ satisfies (1.2) and without loss of generality $|h(n)| \leq 1$. Furthermore we assume that Δ is linearly ordered.

By continuity there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon/2$ if $d(x,y) < \delta$. Next fix $j \geq 1$ with $2^{-j} \leq \delta$. For every $x = (x_0, x_1, \dots) \in X$ set

$$x^{(j)} = (x_0, x_1, \dots, x_{j-1}, y_j, y_{j+1}, \dots)$$

such that $(y_j, y_{j+1}, \dots)$ is the lexicographically smallest sequence in Δ with $(x_0, x_1, \dots, x_{j-1},\ y_j, y_{j+1}, \dots) \in X$. Then $d(x, x^{(j)}) < 2^{-j} \leq \delta$ and consequently $|f(x) - f(x^{(j)})| < \varepsilon/2$. Hence

$$\Big|\sum_{n\leq N} f(T^n x) h(n) - \sum_{n\leq N} f((T^n x)^{(j)}) h(n)\Big| \leq \frac{\varepsilon}{2} N.$$

Since $x^{(j)}$ only depends on $x_0, \dots, x_{j-1}$, so does $f(x^{(j)})$, which can therefore be written as $f(x^{(j)}) = g(x_0, \dots, x_{j-1})$ for some function $g : \Delta^j \to \mathbb{C}$. By assumption there exists $N_0 = N_0(\varepsilon)$ such that

$$\Big|\sum_{n\leq N} g(a_{n+\ell}, \dots, a_{n+\ell+j-1}) h(n)\Big| \leq \frac{\varepsilon}{2} N$$

for all $N \geq N_0$ and all $\ell \geq 0$.

Finally we fix $x \in X$. By definition this means that for all $N \in \mathbb{N}$ there exists $\ell \in \mathbb{N}$ with

$$x_n = a_{n+\ell} \quad \text{for } n = 1, \ldots, N + j. \tag{2.1}$$

In particular it follows that

$$f((T^n x)^{(j)}) = g(a_{n+\ell}, \ldots, a_{n+\ell+j-1}) \quad \text{for } n = 1, \ldots, N,$$

and thus

$$\Big|\sum_{n\le N} f((T^n x)^{(j)})h(n)\Big| = \Big|\sum_{n\le N} g(a_{n+\ell}, \ldots, a_{n+\ell+j-1})h(n)\Big| \le \frac{\varepsilon}{2}N.$$

Consequently, for $N \ge N_0(\varepsilon)$,

$$\Big|\sum_{n\le N} f(T^n x)h(n)\Big| \le \varepsilon N,$$

which completes the proof of the lemma. ■

Next we find specific bounds on how often synchronizing words occur.

LEMMA 2.2. *Let $\mathcal{A}$ be a synchronizing DFAO with synchronizing word $\tilde{\mathbf{w}} \in \Sigma^{m_0}$. There exists $\eta > 0$ depending only on m_0 and k such that the number of synchronizing words of length n is bounded from below by $k^n(1 - k^{m_0-\eta n})$ for all $n \in \mathbb{N}$, that is, at most $O(k^{n(1-\eta)})$ words of length n are not synchronizing.*

The same kind of statement holds if we delete leading zeros of the words in Σ^n and consider only those words in Σ^n for which the reduced word is synchronizing.

Proof. We note that if $\tilde{\mathbf{w}}$ occurs as a (consecutive) subword of some $\mathbf{w} \in \Sigma^*$, then $\mathbf{w}$ is also a synchronizing word. We now decompose every word $\mathbf{w} \in \Sigma^n$ into $\lfloor n/m_0\rfloor$ distinct parts, each of length m_0, and one part of length $n \bmod m_0$. If $\mathbf{w}$ is not synchronizing, it follows directly that no block of length m_0 can coincide with $\tilde{\mathbf{w}}$. Thus we can bound the number of words of length n that are not synchronizing by $(k^{m_0} - 1)^{\lfloor n/m_0\rfloor} \cdot k^{n \bmod m_0}$. We define

$$\eta := \frac{\log(k^{m_0}/(k^{m_0} - 1))}{\log(k_0^m)} > 0,$$

and the statement follows directly.

It is easy to extend the proof to the reduced case, where we delete leading zeros. ■

Now we are prepared to prove the following proposition which leads to the main theorem of this paper, Theorem 1.2.

PROPOSITION 2.3. *Let $(a_n)_{n\in\mathbb{N}}$ be an automatic sequence generated by a synchronized automaton. Then for all $j \in \mathbb{N}, h : \mathbb{N} \to \mathbb{C}$ fulfilling* (1.2) *and $g : \Delta^j \to \mathbb{C}$,*

$$\sum_{n\le N} g(a_{n+\ell},\dots,a_{n+\ell+j-1})h(n) = o(N)$$

uniformly for $\ell \in \mathbb{N}$.

Proof. Without loss of generality, we assume that g is absolutely bounded by 1. We have to show that for every $\varepsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$ and all $\ell \ge 0$,

$$\Big|\sum_{n\le N} g(a_{n+\ell},\dots,a_{n+\ell+j-1})h(n)\Big| \le \varepsilon N,$$

in particular, N_0 does not depend on ℓ. We choose n_1 such that there exist at least $k^{n_1}(1-\varepsilon/(3k^j))$ synchronizing words of length n_1. (Lemma 2.2 shows that

$$n_1 = \frac{1}{\eta}\left(m_0 - \frac{\log(\varepsilon/(3k^j))}{\log k}\right)$$

is a valid choice.) Let M denote the set of integers $0 \le m < k^{n_1}$ such that the j input sequences $\mathbf{w}_{m+i \bmod k^{n_1}}$, $0 \le i < j$, are synchronizing. Clearly we have $|M| \ge k^{n_1}(1-\varepsilon/3)$. We then choose N_0 such that for all $a < k^{n_1}$ and $N \ge N_0$,

$$\Big|\sum_{\substack{n\le N\\ n\equiv a \bmod k^{n_1}}} h(n)\Big| \le \frac{\varepsilon}{3}\,\frac{N}{k^{n_1}}.$$

Furthermore we assume that $N_0 \ge k^{n_1}$. Note that N_0 only depends on n_1, k and ε (and h).

Furthermore we note that if $n \equiv m \bmod k^{n_1}$ for some $m < k^{n_1}$ for which the input word $\mathbf{w}_m$ is synchronizing, then $a_n = a_m$, since $\mathbf{w}_n = \mathbf{v}\mathbf{w}_m$ for some $\mathbf{v} \in \{0,\dots,k-1\}^*$.

Hence, we obtain

$$\begin{aligned}
&\Big|\sum_{n\le N} g(a_{n+\ell},\dots,a_{n+\ell+j-1})h(n)\Big| \\
&\quad\le \sum_{m\in M}\Big|\sum_{\substack{n\le N\\ n+\ell\equiv m \bmod k^{n_1}}} g(a_{n+\ell},\dots,a_{n+\ell+j-1})h(n)\Big| \\
&\qquad + \#\{n \le N : (n+\ell \bmod k^{n_1}) \notin M\} \\
&\quad\le \sum_{m\in M} |g(a_m,\dots,a_{m+j-1})|\Big|\sum_{\substack{n\le N\\ n\equiv m-\ell \bmod k^{n_1}}} h(n)\Big| + \left\lceil \frac{N}{k^{n_1}}\right\rceil \cdot (k^{n_1} - |M|) \\
&\quad\le \frac{\varepsilon}{3}N + \left(\frac{N}{k^{n_1}} + 1\right)\frac{k^{n_1}\varepsilon}{3} \le \varepsilon N,
\end{aligned}$$

which completes the proof. ∎

Of course this implies Theorem 1.2.

3. Reading the digits in reverse order. We made the usual convention that we start with the most significant digit of $\mathbf{w}_n$ for the input in $\mathcal{A}$. We could also start with the least significant digit but then we will get (in general) another automatic sequence which can also be considered as an automatic sequence generated by a synchronizing automaton.

If we start with the most significant digit, then the analysis of Section 2 shows that the density of the pull-back of an element of the output set exists (see also Section 4). However, if we start with the least significant digit, the situation is completely different. As we will see in the analysis below, we will typically get no density (just a logarithmic density, which exists for all automatic sequences; see for example [1, Corollary 8.4.9]). Nevertheless we can also say something in the direction of the Sarnak conjecture for sequences of this kind (however, in a slightly different form).

Theorem 3.1. *Let $h : \mathbb{N} \to \mathbb{C}$ be a bounded function fulfilling*

$$\sum_{n\le N} h(n) = o(N). \tag{3.1}$$

Let, furthermore, $(a_n)_{n\in\mathbb{N}}$ be an automatic sequence that is generated by a synchronizing automaton, where the input starts with the least significant digit of n; without loss of generality assume that $\Delta \subseteq \mathbb{C}$. Then

$$\sum_{n\le N} a_n h(n) = o(N).$$

This implies that $(a_n)_{n\in\mathbb{N}}$ is orthogonal to the Möbius function in the sense of (1.1).

In order to prove Theorem 3.1 we will use the following lemma.

Lemma 3.2. *Let $k \ge 2$, $f : \mathbb{N} \to \mathbb{C}$ a bounded function and $\mathcal{B}$ a subset of $\mathbb{N}$ with density 1. Then the following properties are equivalent:*

$$\sum_{n\le N} f(n) = o(N) \qquad (N \to \infty). \tag{3.2}$$

$$\textit{For all } b \in B, \quad \sum_{0\le n<k^t} f(bk^t + n) = o(k^t) \quad (t \to \infty). \tag{3.3}$$

Proof. Without loss of generality, we may assume that f is bounded by 1.

First, we assume (3.2) and prove (3.3) for any non-negative integer b not necessarily belonging to $\mathcal{B}$. For $\varepsilon > 0$, we select T_ε such that for all $N \ge k^{T_\varepsilon}$,

$$\Big|\sum_{m\le N} f(m)\Big| \le \frac{\varepsilon}{3b} N.$$

For $t \geq T_\varepsilon$, we have

$$\Big|\sum_{n \leq k^t} f(bk^t + n)\Big| \leq \Big|\sum_{m \leq (b+1)k^t} f(m)\Big| + \Big|\sum_{m \leq b\,k^t} f(m)\Big| \leq \frac{\varepsilon}{3b}((b+1)k^t + bk^t) \leq \varepsilon k^t,$$

which proves (3.3).

Secondly, we assume that (3.3) holds true. We fix $\varepsilon > 0$ and construct an integer V such that

$$\text{for all } N \geq k^V, \quad \Big|\sum_{n \leq N} f(n)\Big| \leq \varepsilon N. \tag{3.4}$$

To this end we first select v such that

$$\text{for all } M \geq k^v, \quad \#([1, M) \cap \mathcal{B}) \geq \Big(1 - \frac{\varepsilon}{4}\Big)M \quad \text{and} \quad k^{-v} \leq \frac{\varepsilon}{4}. \tag{3.5}$$

Condition (3.3) says that for all $b \in \mathcal{B}$ and all $\eta > 0$ there exists $T(b, \eta)$ such that for all $t \geq T(b, \eta)$,

$$\Big|\sum_{0 \leq n < k^t} f(bk^t + n)\Big| \leq \eta k^t. \tag{3.6}$$

We define

$$V = v + \max_{b \in \mathcal{B},\, b \leq k^{v+1}} T\Big(b, \frac{\varepsilon}{4k^{2v+1}}\Big). \tag{3.7}$$

Let $N \geq k^V$; we let $u = \lfloor \log N / \log k \rfloor$ so that $k^u \leq N < k^{u+1}$. We set $A = \lfloor N/k^{u-v} \rfloor$ so that $k^v \leq A < k^{v+1}$. We split the interval $[1, N]$ into the disjoint union of $A + 1$ intervals $I_0, I_1, \ldots, I_A$ given by

$$\begin{aligned} I_0 &= [1, k^{u-v}), \\ I_a &= [ak^{u-v}, (a+1)k^{u-v}) \quad \text{for } 1 \leq a < A, \\ I_A &= [Ak^{u-v}, N]. \end{aligned}$$

Since the length of any interval I_i is at most k^{u-v}, we have

$$\Big|\sum_{n \in I_0 \cup I_A} f(n)\Big| \leq |I_0| + |I_A| \leq 2k^{u-v} \leq 2Nk^{-v} \leq \frac{\varepsilon}{2}N. \tag{3.8}$$

Since $A \geq k^v$, by (3.5) we have

$$\Big|\sum_{1 \leq a < A,\, a \notin \mathcal{B}} \sum_{n \in I_a} f(n)\Big| \leq \sum_{1 \leq a < A,\, a \notin \mathcal{B}} |I_a| \leq \frac{\varepsilon}{4}Ak^{u-v} \leq \frac{\varepsilon}{4}N. \tag{3.9}$$

If $a \in \mathcal{B}$, by (3.6) and (3.7) we have

$$\Big|\sum_{m \in I_a} f(m)\Big| = \Big|\sum_{0 \leq n < k^{u-v}} f(ak^{u-v} + n)\Big| \leq \frac{\varepsilon}{4k^{2v+1}}k^{u-v} \leq \frac{\varepsilon}{4k^{v+1}}N.$$

Since $A \le k^{v+1}$, we have

$$(3.10)\qquad \Big|\sum_{1\le a<A,\, a\in\mathcal{B}}\ \sum_{m\in I_a} f(m)\Big| \le \sum_{1\le a<A,\, a\in\mathcal{B}} \Big|\sum_{m\in I_a} f(m)\Big| \le \frac{A\varepsilon}{4k^{v+1}}N \le \frac{\varepsilon}{4}N.$$

Putting together (3.8)–(3.10), we obtain (3.4), which completes the proof of Lemma 3.2. ∎

It is now easy to complete the proof of Theorem 3.1.

Proof of Theorem 3.1. First we observe that (3.1) and Lemma 3.2 applied in the special case $\mathcal{B} = \mathbb{N}$ imply

$$\sum_{0\le n<k^t} h(bk^t + n) = o(k^t)$$

for all positive integers b.

Next let $\mathcal{B}$ be the set of positive integers b for which the k-ary expansions $\mathbf{w}_b$ are synchronizing (when we start with the least significant digit). Hence for every integer of the form $m = bk^t + n$ (with $b \in \mathcal{B}$ and $0 \le n < k^t$) the output a_m of the automaton is constant, namely a_b. Thus it follows that

$$\sum_{0\le n<k^t} a_{bk^t+n} h(bk^t + n) = a_b \sum_{0\le n<k^t} h(bk^t + n) = o(k^t) \qquad (t \to \infty).$$

By Lemma 2.2 the set $\mathcal{B}$ has density 1. Consequently, by Lemma 3.2,

$$\sum_{n\le N} a_n h(n) = o(N) \qquad (N \to \infty),$$

which finishes the proof. ∎

We remark that we (usually) get no densities of the pull-back of the elements of the output set. Namely, for every synchronizing b the set $\bigcup_{t\ge 0}\{bk^t + n : 0 \le n < k^t\}$ (on which the output sequence equals a_b) has no density. Indeed, if this sequence has a density δ, we must have $\delta(bk^t + k^t) = \delta(bk^t) + k^t + o(k^t)$, whence $\delta = 1$, which occurs only when our automatic sequence is almost constant.

Finally, we note that we do not see how we could generalize Theorem 3.1 to the dynamical system related to $(a_n)_{n\in\mathbb{N}}$. We can show that for any $j \ge 1$, $g : \Delta^j \to \mathbb{C}$ and any $\ell \ge 1$, one has

$$\sum_{n\le N} g(a_{n+\ell}, \ldots, a_{n+\ell+j-1}) h(n) = o(N) \qquad (N \to \infty),$$

but it is not clear how to show the uniformity in ℓ.

4. Gelfond problems. Let $s_k(n)$ denote the k-ary sum-of-digits function. Then it is well known that sequences of the form $a_n = s_k(n) \bmod m$ are

k-automatic. The most prominent sequence of this form is the Thue–Morse sequence $t_n = s_2(n) \bmod 2$.

Without using (or even knowing) the notion of automatic sequences, Gelfond [11] proved that for every $\ell \in \{0, 1, \ldots, m-1\}$,

$$\#\{n \le N : s_k(an+b) \equiv \ell \bmod m\} = \frac{N}{m} + O(N^{1-\eta})$$

for some $\eta > 0$ (provided that $(k-1, m) = 1$), that is, linear subsequences of the k-automatic sequence $a_n = s_k(n) \bmod m$ are asymptotically uniformly distributed on the output alphabet $\{0, 1, \ldots, m-1\}$.

What is even more interesting, in the same paper [11] Gelfond formulated three conjectures, which are usually called *Gelfond problems*:

(1) If $k_1, k_2 \ge 2$ are coprime integers and $(k_1 - 1, m_1) = (k_2 - 1, m_2) = 1$ then

$$\#\{n \le N : s_{k_1}(n) \equiv \ell_1 \bmod m_1,\ s_{k_2}(n) \equiv \ell_2 \bmod m_2\} = \frac{N}{m_1 m_2} + O(N^{1-\eta})$$

for some $\eta > 0$.

(2) If $k \ge 2$ with $(k-1, m) = 1$ then

$$\#\{p \le N : p \in \mathbb{P},\ s_k(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

for some $\eta > 0$. (Here $\pi(x)$ denotes the number of primes $\le x$.)

(3) If $k \ge 2$ with $(k-1, m) = 1$ then for every integer polynomial $P(x)$,

$$\#\{n \le N : s_k(P(n)) \equiv \ell \bmod m\} = \frac{N}{m} + O(N^{1-\eta})$$

for some $\eta > 0$.

Whereas the first problem was almost immediately solved by Bésineau [4] (without an explicit error term; the error term was finally proved by Kim [13]), it took more than 40 years till the other two problems were solved or came close to a solution. Actually, the second problem on the subsequence along the primes was solved by Mauduit and Rivat [15]. The third problem was completely solved for quadratic polynomials by Mauduit and Rivat [14] and partially solved for general polynomials by Drmota, Mauduit and Rivat [9] (it is assumed that the base k is prime and sufficiently large with respect to the degree of $P(x)$).

It is immediate to translate the Gelfond problems to arbitrary automatic sequences. The only difference is that we cannot expect a uniform distribution on the output alphabet but just densities (or only logarithmic densities). The general question is maybe out of reach, nevertheless we can give a complete answer for automatic sequences that are generated by synchronizing automata. (Note that such sequences need not be uniformly distributed. For example, the period-doubling sequence, described in [1, Example 6.3.4], is synchronizing and a quick computation shows that q_0 has density $2/3$ and q_1 has density $1/3$.)

THEOREM 4.1. *Suppose that $k_1, k_2 > 1$ are coprime integers and $(a_n)_{n\in\mathbb{N}} \in \Delta_1^{\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}} \in \Delta_2^{\mathbb{N}}$ are automatic sequences that are generated by a synchronizing k_1-automaton and a synchronizing k_2-automaton, respectively. Then for every pair $(\alpha,\beta) \in (\Delta_1, \Delta_2)$ the density*

$$\delta(\alpha,\beta) = \lim_{N\to\infty} \frac{1}{N} \#\{n \le N : (a_n, b_n) = (\alpha,\beta)\}$$

exists and equals the product of the densities of α in $(a_n)_{n\in\mathbb{N}}$ and β in $(b_n)_{n\in\mathbb{N}}$.

THEOREM 4.2. *Suppose that $(a_n)_{n\in\mathbb{N}} \in \Delta_1^{\mathbb{N}}$ is an automatic sequence that is generated by a synchronizing automaton. Then for every $\alpha \in \Delta$ the density*

$$\delta_{\mathbb{P}}(\alpha) = \lim_{N\to\infty} \frac{1}{\pi(N)} \#\{p \in \mathbb{P},\, p \le N : a_p = \alpha\}$$

exists.

THEOREM 4.3. *Suppose that $(a_n)_{n\in\mathbb{N}} \in \Delta^{\mathbb{N}}$ is an automatic sequence that is generated by a synchronizing automaton and $P(x)$ is a non-zero and positive integer valued polynomial. Then for every $\alpha \in \Delta$ the density*

$$\delta_P(\alpha) = \lim_{N\to\infty} \frac{1}{N} \#\{n \le N : a_{P(n)} = \alpha\}$$

exists.

Before we start the proof of Theorems 4.1–4.3 we show that every element of the output alphabet Δ of an automatic sequence that is generated by a synchronizing automaton has a density.

LEMMA 4.4. *Suppose that $(a_n)_{n\in\mathbb{N}} \in \Delta^{\mathbb{N}}$ is an automatic sequence that is generated by a synchronizing automaton. Then for every $\alpha \in \Delta$ the density*

$$\delta(\alpha) = \lim_{N\to\infty} \frac{1}{N} \#\{n \le N : a_n = \alpha\}$$

exists.

Proof. Let $\varepsilon > 0$. Then by using a simplified version of the arguments in the proof of Proposition 2.3 there exists $n_1 \ge 1$ and a set M of integers $0 \le m < k^{n_1}$ such that $|M| \ge k^{n_1}(1-\varepsilon)$ and that all words $\mathbf{w}_m$ with $m \in M$ are synchronizing (in the automaton related to $(a_n)_{n\in\mathbb{N}}$). We also recall that

$a_n = a_m$ if $n \equiv m \bmod k^{n_1}$. Hence

$$\begin{aligned}\#\{n \le N : a_n = \alpha\} &= \sum_{m \in M,\, a_m = \alpha} \#\{n \le N : n \equiv m \bmod k^{n_1}\} + O(\varepsilon N) \\ &= \sum_{m \in M,\, a_m = \alpha} \left(\frac{N}{k^{n_1}} + O(1) \right) + O(\varepsilon N) \\ &= \frac{\#\{m \in M : a_m = \alpha\}}{k^{n_1}} N + O(k^{n_1}) + O(\varepsilon N).\end{aligned}$$

Consequently, the sequence

$$\frac{1}{N} \#\{n \le N : a_n = \alpha\}$$

is a Cauchy sequence, and thus convergent. ∎

Proof of Theorem 4.1. As in the proof of Lemma 4.4 we construct for a given $\varepsilon > 0$ a set M_1 of integers $0 \le m_1 < k_1^{n_1}$ such that $|M_1| \ge k^{n_1}(1-\varepsilon)$ and all k_1-ary words $\mathbf{w}_{m_1}$ with $m_1 \in M_1$ are synchronizing in the automaton related to $(a_n)_{n\in\mathbb{N}}$. In the same way we can handle the sequence $(b_n)_{n\in\mathbb{N}}$. Therefore we have to construct a corresponding set M_2 of integers $0 \le m_2 < k_2^{n_2}$ such that $|M_2| \ge k_2^{n_2}(1-\varepsilon)$ and all k_2-ary words $\mathbf{w}_{m_2}$ with $m_2 \in M_2$ are synchronizing in the automaton related to $(b_n)_{n\in\mathbb{N}}$.

Similarly to the above we obtain

$$\begin{aligned}&\#\{n \le N : (a_n, b_n) = (\alpha, \beta)\} \\ &\quad = \sum_{\substack{m_1 \in M_1 \\ a_{m_1} = \alpha}} \sum_{\substack{m_2 \in M_2 \\ b_{m_2} = \beta}} \#\{n \le N : n \equiv m_1 \bmod k_1^{n_1},\, n \equiv m_2 \bmod k_2^{n_2}\} + O(\varepsilon N) \\ &\quad = \sum_{m_1 \in M_1,\, a_{m_1} = \alpha} \; \sum_{m_2 \in M_2,\, b_{m_2} = \beta} \left(\frac{N}{k_1^{n_1} k_2^{n_2}} + O(1) \right) + O(\varepsilon N) \\ &\quad = \frac{\#\{m_1 \in M_1 : a_{m_1} = \alpha\}}{k_1^{n_1}} \cdot \frac{\#\{m_2 \in M_2 : b_{m_2} = \beta\}}{k_2^{n_2}} N + O(k_1^{n_1} k_2^{n_2}) \\ &\qquad + O(\varepsilon N),\end{aligned}$$

where we have used the assumption that k_1 and k_2 are coprime. This shows that $\delta(\alpha, \beta)$ exists, and also that it is equal to $\delta(\alpha) \cdot \delta(\beta)$. ∎

Proof of Theorem 4.2. For every $\varepsilon > 0$ we define a set M as in the proof of Lemma 4.4. Furthermore we denote by $\pi(N)$ the number of primes $p \le N$ and by $\pi(a, k; N)$ the number of primes $p \le N$ with $p \equiv a \bmod k$. By the prime number theorem in arithmetic progressions we have $\pi(a, k; N) \sim \pi(N)/\varphi(k)$ if $(a, k) = 1$.

Here we have

$$\begin{aligned}
&\#\{p \in \mathbb{P},\, p \le N : a_p = \alpha\} \\
&\qquad = \sum_{m \in M,\, a_m = \alpha} \#\{p \in \mathbb{P},\, p \le N : p \equiv m \bmod k^{n_1}\} + O(\varepsilon \pi(N)) \\
&\qquad = \sum_{m \in M,\, (m,k^{n_1})=1,\, a_m = \alpha} \frac{\pi(N)}{\varphi(k^{n_1})} + o(\pi(N)) + O(k_1^{n_1}) + O(\varepsilon \pi(N)),
\end{aligned}$$

which shows that the sequence

$$\frac{1}{\pi(N)} \#\{p \in \mathbb{P},\, p \le N : a_p = \alpha\}$$

is a Cauchy sequence, and thus convergent. ∎

Proof of Theorem 4.3. The proof is (again) very similar to the proof of Lemma 4.4. With the notation

$$c(m,K) = \#\{0 \le n < K : P(n) \equiv m \bmod K\}$$

we have

$$\begin{aligned}
&\#\{n \le N : a_{P(n)} = \alpha\} \\
&\qquad = \sum_{m \in M,\, a_m = \alpha} \#\{n \le N : P(n) \equiv m \bmod k^{n_1}\} + O(\varepsilon N) \\
&\qquad = \sum_{m \in M,\, a_m = \alpha} \left(c(m,k^{n_1})\, \frac{N}{k^{n_1}} + O(c(m,k^{n_1})) \right) + O(\varepsilon N) \\
&\qquad = \sum_{m \in M,\, a_m = \alpha} \frac{c(m,k^{n_1})}{k^{n_1}} N + O(k^{2n_1}) + O(\varepsilon N).
\end{aligned}$$

Again this proves convergence and completes the proof of the theorem. ∎

Acknowledgements. All authors are supported by the binational research project MuDeRa which is funded by the French and Austrian Science Funds ANR and FWF. The last two authors are also supported by the project F55-02 of the Austrian Science Fund FWF which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

References

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences. Theory, Applications, Generalizations*, Cambridge Univ. Press, Cambridge, 2003.

[2] M. V. Berlinkov, *On the probability to be synchronizable*, arXiv:1304.5774 (2013).

[3] M. V. Berlinkov, *Testing for synchronization*, arXiv:1401.2553 (2014).

[4] J. Bésineau, *Indépendance statistique d'ensembles liés à la fonction "somme des chiffres"*, Acta Arith. 20 (1972), 401–416.

[5] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Moebius from horocycle flows*, in: From Fourier Analysis and Number Theory to Radon Transforms and Geometry, Dev. Math. 28, Springer, New York, 2013, 67–83.

[6] C. Dartyge et G. Tenenbaum, *Sommes des chiffres de multiples d'entiers*, Ann. Inst. Fourier (Grenoble) 55 (2005), 2423–2474.

[7] H. Davenport, *On some infinite series involving arithmetical functions (II)*, Quart. J. Math. 8 (1937), 313–320.

[8] M. Drmota, *Subsequences of automatic sequences and uniform distribution*, in: Uniform Distribution and Quasi-Monte Carlo Methods, Radon Ser. Comput. Appl. Math. 15, de Gruyter, Berlin, 2014, 87–104.

[9] M. Drmota, C. Mauduit, and J. Rivat, *The sum-of-digits function of polynomial sequences*, J. London Math. Soc. 84 (2011), 81–102.

[10] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, and C. Mauduit, *Substitutions and Möbius disjointness*, arXiv:1507.01123v1 (2015).

[11] A. Gelfond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith. 13 (1967/1968), 259–265.

[12] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) 175 (2012), 541–566.

[13] D.-H. Kim, *On the joint distribution of q-additive functions in residue classes*, J. Number Theory 74 (1999), 307–336.

[14] C. Mauduit et J. Rivat, *La somme des chiffres des carrés*, Acta Math. 203 (2009), 107–148.

[15] C. Mauduit et J. Rivat, *Sur un problème de Gelfond: la somme des chiffres des nombres premiers*, Ann. of Math. (2) 171 (2010), 1591–1646.

[16] C. Mauduit and J. Rivat, *Prime numbers along Rudin–Shapiro sequences*, J. Eur. Math. Soc. 17 (2015), 2595–2642.

[17] P. Sarnak, *Three Lectures on the Möbius Function Randomness and Dynamics*, http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures%282%29.pdf.

Jean-Marc Deshouillers
Institut Mathématique de Bordeaux
Bordeaux INP
33405 Talence, France
E-mail: jean-marc.deshouillers@math.u-bordeaux.fr

Michael Drmota, Clemens Müllner
Institut für Diskrete Mathematik
und Geometrie
TU Wien
Wiedner Hauptstr. 8–10
1040 Wien, Austria
E-mail: michael.drmota@tuwien.ac.at
clemens.muellner@tuwien.ac.at