(Logarithmic) Densities for Automatic Sequences along Primes and Squares

Michael Drmota

joint work with Boris Adamczewski and Clemens Müllner

supported by the Austrian Science Foundation FWF, project F5502

Institut für Diskrete Mathematik und Geometrie
Technische Universität Wien

One World Numeration Seminar
March 30, 2021

Summary

- ★ Thue-Morse and Rudin-Shapiro sequence
- ★ Thue-Morse sequence along subsequences
- ★ Automatic sequences
- ★ Logarithmic densities along subsequences
- ★ Automatic sequences along primes
- ★ Automatic sequences along squares

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

0

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

01

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

0110

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

01101001

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

0110100110010110

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

0110100110010110100101100110 1001

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$011010011001011010010110011010011001011001101\cdots$$

$$t_0 = 0, \qquad t_{2^n+k} = 1 - t_k \quad (0 \leqslant k < 2^n)$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$011010011001011010010110011010011001011001101\cdots$$

$$t_0 = 0, \qquad t_{2^n+k} = 1 - t_k \quad (0 \leqslant k < 2^n)$$

$$\boxed{t(n) = s_2(n) \bmod 2}$$

$$n = \sum_{i=0}^{\ell-1} \varepsilon_i(n) q^i \qquad \varepsilon_i(n) \in \{0, 1, \ldots, q-1\}, \qquad s_q(n) = \sum_{i=0}^{\ell-1} \varepsilon_i(n)$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

$$01101001100101101001011001101001100101100110 1 \cdots$$

$$t_0 = 0, \qquad t_{2^n+k} = 1 - t_k \quad (0 \leqslant k < 2^n) \qquad \text{or} \qquad t_{2k} = t_k,\ t_{2k+1} = 1 - t_k$$

$$\boxed{t(n) = s_2(n) \bmod 2}$$

$$n = \sum_{i=0}^{\ell-1} \varepsilon_i(n) q^i \qquad \varepsilon_i(n) \in \{0, 1, \ldots, q-1\}, \qquad s_q(n) = \sum_{i=0}^{\ell-1} \varepsilon_i(n)$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$0110100110010110100101100110100110010110011 01\cdots$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$0110100110010110100101100110100110010110\cdots$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

$$0110\mathbf{10}0110010110100101100110100110010110 01101\cdots$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$011010011001011010010110011010011001011001101\cdots$$

$$\#\{0 \leqslant n < N : t(n) = 0\} \sim \frac{N}{2}$$

★ Thue-Morse sequence

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$01101001100101101001011001101001100101100110 1\cdots$$

$$\#\{0\leqslant n<N : t(n)=0\}\sim\frac{N}{2}$$

The letters 0 and 1 appear with asymptotic frequency $\frac{1}{2}$:

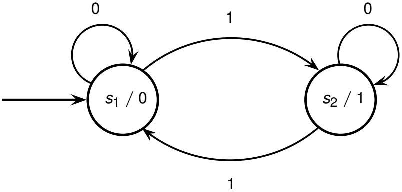
$$\mathrm{dens}(t(n),0)=\mathrm{dens}(t(n),1)=\frac{1}{2}.$$

★ Thue-Morse sequence

- TM sequence is **not periodic** and **cubeless**.
- TM sequence is **almost periodic**:
  Every appearing consecutive block appears infinitely many times with bounded gaps.
- **Subword complexity is linear**: $p_k \leqslant \frac{10}{3} k$
  p_k ... subword complexity (*number of different consecutive blocks of length k that appear in the TM sequence*).
- **Zero topological entropy** of the corresponding dynamical system:
  $$h = \lim_{k\to\infty} \frac{1}{k} \log p_k = 0$$
- **Linear subsequences** $(t_{an+b})_{n\geqslant 0}$ have the same properties.
- The TM sequence and its linear subsequences are **automatic sequences**.

★ Thue-Morse sequence

Automaton that generates the Thue-Morse sequence:
$t(n) = \sum_{j \geqslant 0} \varepsilon_j(n) \bmod 2$



★ Rudin-Shapiro sequence

Rudin-Shapiro sequence $(r(n))_{n\geqslant 0}$:

★ Rudin-Shapiro sequence

Rudin-Shapiro sequence $(r(n))_{n\geqslant 0}$:

$$000100100001110100010010111000100001001000011101111\cdots$$

★ Rudin-Shapiro sequence

Rudin-Shapiro sequence $(r(n))_{n\geqslant 0}$:

$$000100100001110100010010111000100001001000011101111\cdots$$

$$r_0 = 0, \qquad r_{2k} = r_k, \qquad r_{2k+1} = \begin{cases} r_k & \text{if } k \text{ is even,} \\ 1 - r_k & \text{if } k \text{ is odd.} \end{cases}$$

★ Rudin-Shapiro sequence

Rudin-Shapiro sequence $(r(n))_{n\geqslant 0}$:

$$0001001000011101000100101110001000010010000011101111\cdots$$

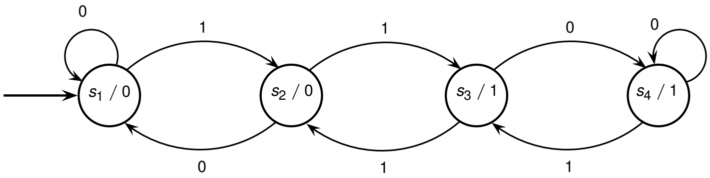
$$r_0 = 0, \qquad r_{2k} = r_k, \qquad r_{2k+1} = \begin{cases} r_k & \text{if } k \text{ is even,} \\ 1 - r_k & \text{if } k \text{ is odd.} \end{cases}$$

$$\boxed{r(n) = \sum_{i\geqslant 0} \varepsilon_i(n)\varepsilon_{i+1}(n) \bmod 2}$$

$$n = \sum_{i=0}^{\ell-1} \varepsilon_i(n) q^i \qquad \varepsilon_i(n) \in \{0, 1, \ldots, q-1\}$$

★ Rudin-Shapiro sequence

Automaton that generates the Rudin-Shapiro sequence:
$r(n) = \sum_{j \geqslant 0} \varepsilon_j(n)\varepsilon_{j+1}(n) \bmod 2$



★ Thue-Morse sequence along primes

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence along primes

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$0110100110010110100101100110100110010110011010\cdots$$

★ Thue-Morse sequence along primes

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

10 0 1 1 1 0 1 0 0 1 1 1 0 $\cdots$

Mauduit and Rivat (2010):

$$\# \{0 \leqslant n < N : t(p_n) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(p_n), 0) = \frac{1}{2}.$$

★ Thue-Morse sequence along primes

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

10 0 1 1 1 0 1 0 0 1 1 1 0 $\cdots$

Mauduit and Rivat (2010):

$$\#\{0 \leqslant n < N : t(p_n) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(p_n), 0) = \frac{1}{2}.$$

Solution of a **Conjecture of Gelfond** (1968)

★ Thue-Morse sequence along primes

Thue-Morse sequence $(t(n))_{n \geqslant 0}$:

10 0 1 1 1 0 1 0 0 1 1 1 0 ⋯

Mauduit and Rivat (2010):

$$\#\{0 \leqslant n < N : t(p_n) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(p_n), 0) = \frac{1}{2}.$$

Solution of a **Conjecture of Gelfond** (1968)

The same property holds for the Rudin-Shapiro sequence $r(n)$ (Mauduit and Rivat (2015)).

★ Thue-Morse sequence along squares

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$0110100110010110100101100110100110010110011010\cdots$$

★ Thue-Morse sequence along squares

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

$$0110100110010110100101100110100110010110 \cdots$$

★ Thue-Morse sequence along squares

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

01 1 0 1 1 0 $\cdots$

Mauduit and Rivat (2009):

$$\#\{0 \leqslant n < N : t(n^2) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(n^2), 0) = \frac{1}{2}$$

★ Thue-Morse sequence along squares

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

01 1 0 1 1 0 ...

Mauduit and Rivat (2009):

$$\#\{0 \leqslant n < N : t(n^2) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(n^2), 0) = \frac{1}{2}$$

Solution of a **Conjecture of Gelfond** (1968)

★ Thue-Morse sequence along squares

Thue-Morse sequence $(t(n))_{n\geqslant 0}$:

01 1 0 1 1 0 $\cdots$

Mauduit and Rivat (2009):

$$\#\{0 \leqslant n < N : t(n^2) = 0\} \sim \frac{N}{2} \quad \text{or} \quad \mathrm{dens}(t(n^2), 0) = \frac{1}{2}$$

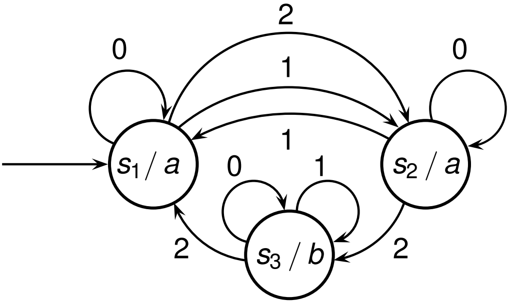
Solution of a **Conjecture of Gelfond** (1968)

The same property holds for the Rudin-Shapiro sequence $r(n)$ (Mauduit and Rivat (2018)).

★ Automatic sequences

Definition

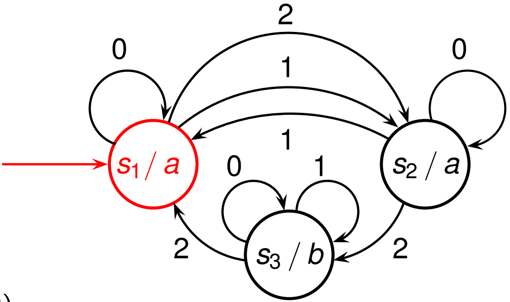
A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.



★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

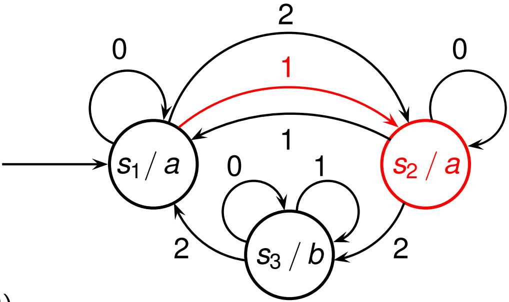


$32 = (1012)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n \geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

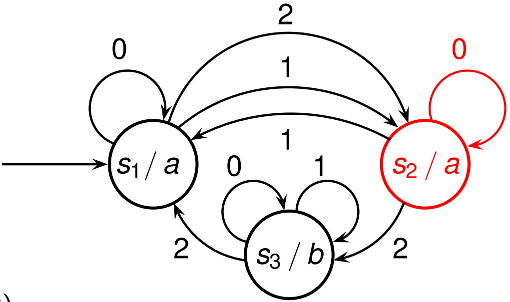


$32 = (1012)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

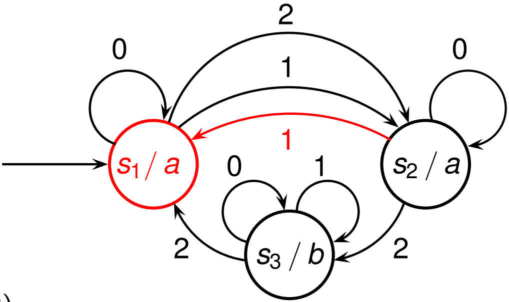


$32 = (1012)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

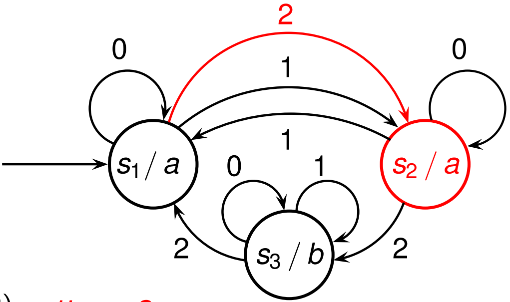


$32 = (1012)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

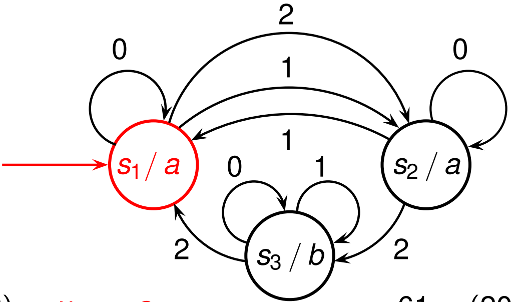


$32 = (1012)_3 \quad u_{32} = a,$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

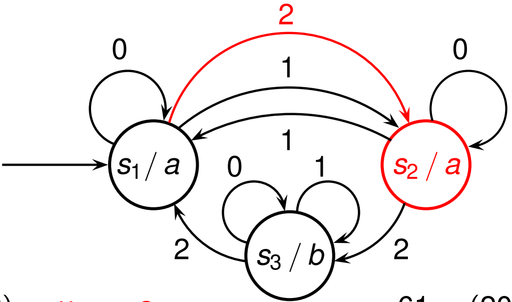


$32 = (1012)_3$ $\quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

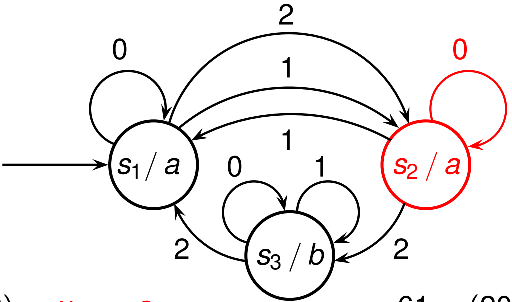


$32 = (1012)_3$ $\quad u_{32} = a$, $\qquad\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n \geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

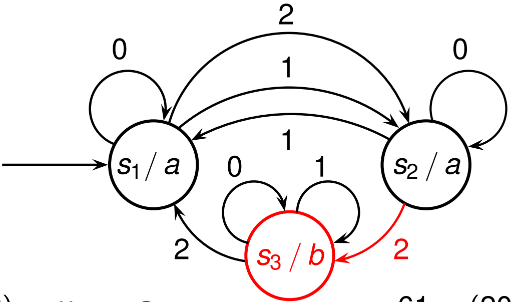


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

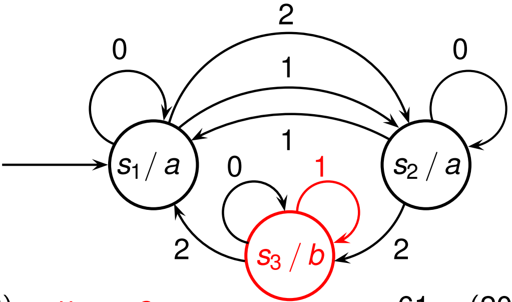


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.

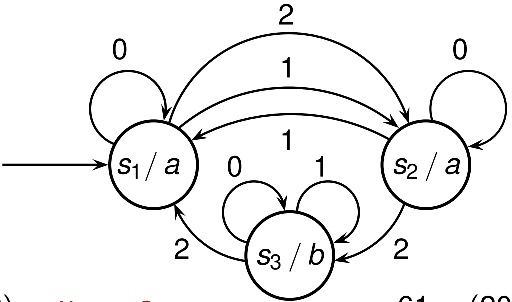


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3 \quad u_{61} = b$

★ Automatic sequences

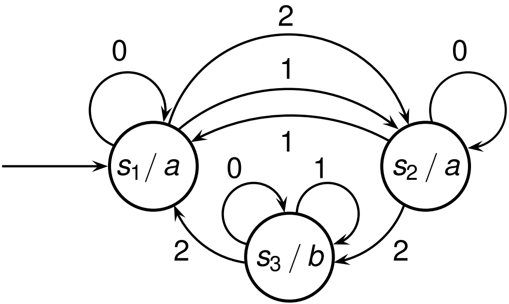
Definition

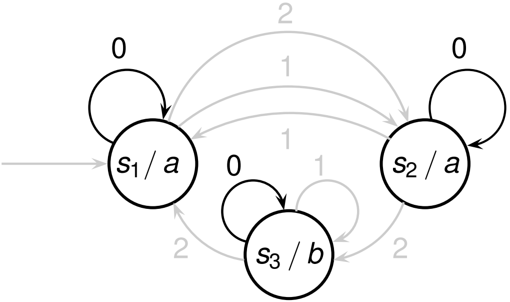
A sequence $(a(n))_{n\geqslant 0}$ is called a *q-automatic sequence*, if $a(n)$ is the output of an automaton when the input is the q-ary expansion of n.



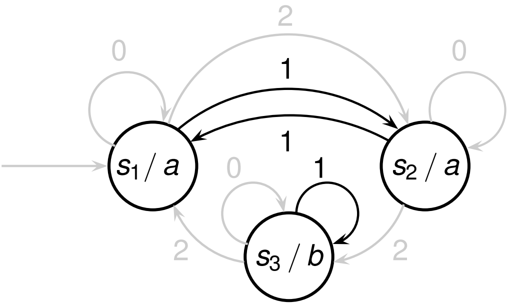
$32 = (1012)_3 \quad u_{32} = a$, $\qquad$ $61 = (2021)_3 \quad u_{61} = b$

$(a(n))_{n\geqslant 0}$: $aaaaabaabaabaaabbaaabaaabbaaabaaabbaaaaaaba\ldots$

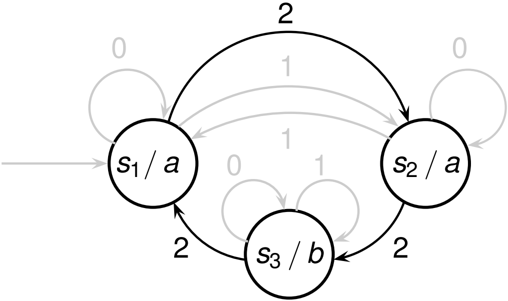

2
0
1
0
1
s_1 / a
0
1
s_2 / a
2
s_3 / b
2



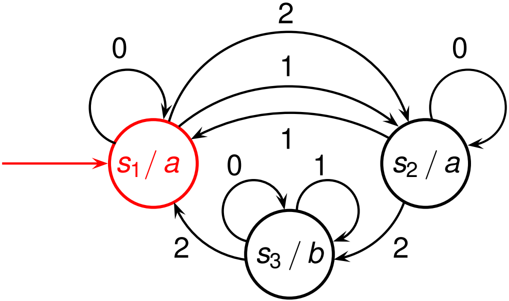
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

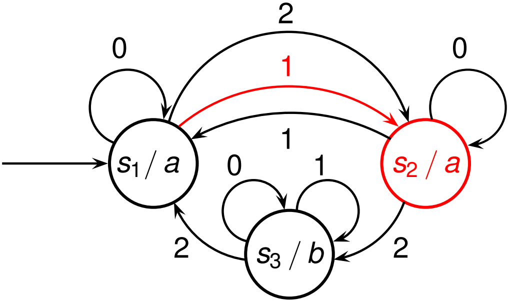


$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$



$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

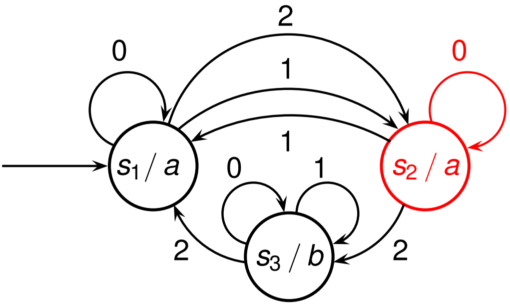
$$32 = (1012)_3 : \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$



$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

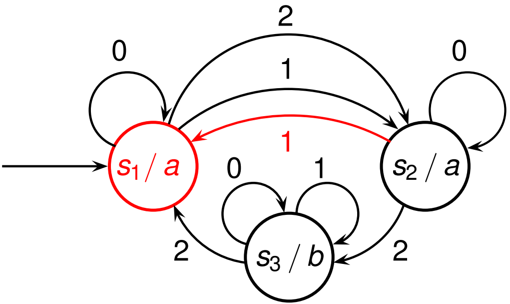
$32 = (1012)_3$:

$$M_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$



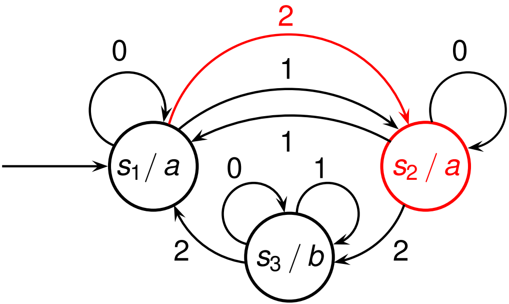
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$32 = (1012)_3 : \qquad M_0 \circ M_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$



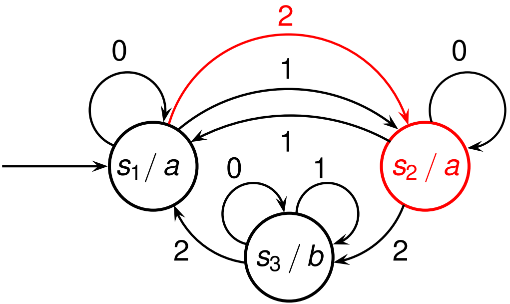
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$32 = (1012)_3 : \qquad M_1 \circ M_0 \circ M_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$



$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$32 = (1012)_3 : \qquad M_2 \circ M_1 \circ M_0 \circ M_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$



$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\boxed{S(n) := M_{\varepsilon_0(n)} M_{\varepsilon_1(n)} \cdots M_{\varepsilon_{\ell-1}(n)}}$$

$$\boxed{a(n) = f(S(n)\mathbf{e}_1)} \qquad \mathbf{e}_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$$

★ Automatic sequences

- For every q-automatic sequence $a(n)$ (on an alphabet $\mathcal{A}$) there exists the **logarithmic density** (for every letter $\alpha \in \mathcal{A}$)

$$\mathrm{logdens}(a(n),\alpha) = \lim_{N\to\infty} \frac{1}{\log N} \sum_{1\leqslant n\leqslant N} \frac{1}{n}\cdot \mathbf{1}_{[a(n)=\alpha]}$$

  which is also computable.
- If the **densities**

$$\mathrm{dens}(a(n),\alpha) = \lim_{N\to\infty} \frac{1}{N}\#\{n\leqslant N : a(n)=\alpha\}$$

  exist then they coincide with the logarithmic densities.
- It can be effectively decided whether densities exist (as for the Thue-Morse sequence $t(n)$ and the Rudin-Shapiro sequence $r(n)$).
- The matrix $M = M_0 + \cdots + M_{q-1}$ plays an important role:

$$\sum_{n<q^\lambda} S(n) = \sum_{n<q^\lambda} M_{\varepsilon_0(n)} M_{\varepsilon_1(n)} \cdots M_{\varepsilon_{\ell-1}(n)} = \left(M_0 + \cdots + M_{q-1}\right)^\lambda$$

★ Automatic sequences

We will say that an automatic sequence is **primitive and prolongable** if the directed graph of the corresponding minimal automaton is strongly connected and the initial state has a 0-labeled loop.

Lemma

For every primitive and prolongable automatic sequence $a(n)$ the densities

$$\mathrm{dens}(a(n),\alpha) = \lim_{N\to\infty} \frac{1}{N} \#\{n \leqslant N : a(n) = \alpha\}$$

exist

Proof with matrix product representation and Perron-Frobenius theory.

★ Automatic sequences

Proof

$\mathbf{b}(n) = S(n)\mathbf{e}_1$. The j-component of $\mathbf{b}(n)$ equals 1 if we stop in state j.

$M = M_0 + \cdots + M_{q-1}$ is primitive matrix;
q is dominant eigenvalue;
P projection to dominant eigen-direction.

$$\sum_{n<q^\lambda} \mathbf{b}(n) = M^\lambda \mathbf{e}_1 = q^\lambda P\mathbf{e}_1 + O(q^{\lambda(1-\varepsilon)})$$

$\Longrightarrow \quad P\mathbf{e}_1 \quad$ is vector of (rational) frequencies.

★ Automatic sequences

Example. Leading digit, $q \geqslant 3$

$a(n)$ ... leading digits of n in the q-ary expansion.

$$a(n) = \ell \quad \Longleftrightarrow \quad \ell q^k \leqslant n < (\ell+1)q^k \quad \text{for some } k \geqslant 0$$

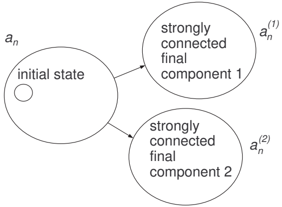
$$\frac{1}{\ell q^k} \#\{n < \ell q^k : a(n) = \ell\} = \frac{1}{\ell q^k} \frac{q^k - 1}{q-1} \sim \frac{1}{\ell(q-1)}$$

$$\frac{1}{(\ell+1)q^k} \#\{n < (\ell+1)q^k : a(n) = \ell\} = \frac{1}{(\ell+1)q^k} \frac{q^{k+1} - 1}{q-1} \sim \frac{q}{(\ell+1)(q-$$

Hence, no densities exist!!

$$\sum_{\ell q^k \leqslant n < (\ell+1)q^k} \frac{1}{n} \sim \log\left(1 + \frac{1}{\ell}\right) \quad \Longrightarrow \quad \mathrm{logdens}(a(n), \ell) = \log_q\left(1 + \frac{1}{\ell}\right)$$

★ Automatic sequences



Suppose that $a_n^{(1)}$ and $a_n^{(2)}$ are produced by the two final components.

- If the densities for $a_n^{(1)}$ and $a_n^{(2)}$ exist and are equal then a_n has the same densities.
- If the densities for $a_n^{(1)}$ and $a_n^{(2)}$ exist but are not equal then a_n has logarithmic densities.

★ Automatic sequences

Theorem

For every automatic sequence $a(n)$ the logarithmic densities $\mathrm{logdens}(a(n),\alpha)$ exist and are computable.
Furthermore, if the densities of those automatic sequences that correspond to the final strongly connected components coincide then the densities $\mathrm{dens}(a(n),\alpha)$ exist and are computable rational numbers.

★ Automatic sequences along primes

Theorem (Adamczewski+D.+Müllner 2020+)

For every automatic sequence $a(n)$ the logarithmic densities $\mathrm{logdens}(a(p_n),\alpha)$ of the subsequence along **prime numbers** *exist and are computable.*
Furthermore, if the densities along primes on those automatic sequences that correspond to the final strongly connected components coincide then the densities $\mathrm{dens}(a(p),\alpha)$ exist and are computable rational numbers.

Theorem (Adamczewski+D.+Müllner 2020+)

For any automatic sequence $a(n)$ there exists a computable positive integer m such that, for all α, $\mathrm{logdens}(a(p_n),\alpha)$ is equal to the logarithmic density of $a(n)$ along the integers n satisfying $(n,m)=1$.

★ Automatic sequences along squares

Theorem (Adamczewski+D.+Müllner 2020+)

For every automatic sequence $a(n)$ the logarithmic densities $\mathrm{logdens}(a(n^2), \alpha)$ of the subsequence along **squares** *exist and are computable.*
Furthermore, if the densities along squares on those automatic sequences that correspond to the final strongly connected components coincide then the densities $\mathrm{dens}(a(n^2), \alpha)$ exist and are also computable. If the input base q is prime, then these densities are rational numbers.

★ Automatic sequences along primes squares

Equivalently one can say that for all automatic sequences $a(n)$ the following two limits exist:

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n\leqslant N} \frac{1}{n} a(n)\Lambda(n) \quad \text{and} \quad \lim_{N\to\infty} \frac{1}{\log N} \sum_{n\leqslant N} \frac{1}{n} a(n^2).$$

And it can be decided when the logarithmic means can be replaced by the usual means:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} a(n)\Lambda(n) \quad \text{and} \quad \lim_{N\to\infty} \frac{1}{N} \sum_{n\leqslant N} a(n^2).$$

$\Lambda(n)$ denotes the von Mangoldt Λ-function
($\Lambda(n) = \log p$ for prime powers $n = p^k$ and $\Lambda(n) = 0$ else.)

★ Paper folding sequence

$$(a(n)) = 1|1|0|1|100|1|1100100|\cdots$$

- **primes**: $m = 2$ (it suffices to consider odd numbers)

$$\mathrm{dens}(a(p_n), 0) = \mathrm{dens}(a(p_n), 1) = \frac{1}{2}$$

- **squares**:

$$\mathrm{dens}(a(n^2), 1) = 1, \qquad \mathrm{dens}(a(n^2), 0) = 0.$$

★ Automatic sequences along subsequences

We call a strictly increasing subsequences $(n_\ell)_{\ell\geqslant 0}$ of the positive integers **regularly varying sequences**, if

$$\boxed{n_\ell = \ell^\gamma L(\ell)},$$

where $\gamma \geqslant 1$ and $L(n)$ is slowly varying in the sense that

$$\lim_{\ell\to\infty} \frac{L(\lceil \delta\ell \rceil)}{L(\ell)} = 1 \qquad (\text{for all } 0 < \delta < 1)$$

The sequence of **primes**, **polynomial sequences**, and **Piatetski-Shapiro sequences** (*i.e.,* $\lfloor n^c \rfloor$, where $c > 1$) are regularly varying sequences.

★ Automatic sequences along subsequences

Theorem (Adamczewski+D.+Müllner 2020+)

Suppose that $(n_\ell)_{\ell \geqslant 0}$ *is a regularly varying sequence and suppose that for any primitive and prolongable automatic sequence* $\tilde{a}(n)$ *the densities along the subsequence* (n_ℓ) *exit:*

$$\mathrm{dens}(\tilde{a}(n_\ell), \alpha) := \lim_{N\to\infty} \frac{\{\ell \leqslant N : \tilde{a}(n_\ell) = \alpha\}}{N}$$

Then the two following properties hold.

(i) *Then for every automatic sequence* $a(n)$ *the logarithmic densities* $\mathrm{logdens}(a(n_\ell), \alpha)$ *exist and can be explicitly computed.*

(ii) *Furthermore, if the densities along the subsequence* n_ℓ *corresponding to those automatic sequences that are generated by the final strongly connected components of the directed graph are all equal then the densities* $\mathrm{dens}(a(n_\ell), \alpha)$ *exist and are equal to them.*

★ Automatic sequences along subsequences

Proposition (Müllner 2017)

Let $a(n)$ be a prolongable and primitive automatic sequence. Then the density of $a(n) = \alpha$ exist along the subsequence of **primes**.

Proposition

Let $a(n)$ be a prolongable and primitive automatic sequence. Then the density of $a(n) = \alpha$ exist along the subsequence of **squares**.

★ Reduction of automatic sequences

Lemma (Müllner 2017)

Let $a(n)$ be a primitive and prolongable q-automatic sequences. Then it can be represented in the form

$$a(n) = f(s(n), T(n)),$$

where $s(n)$ is a pure synchronizing q-automatic sequence and $T(n)$ takes values in a finite group G with the following property. For every $j < q$ and every q there exists $g_{j,q} \in G$ such that

$$T(n \cdot q + j) = T(n) \cdot g_{j,s(n)}$$

holds for all $n \in \mathbb{N}$.

The synchronizing part can be handled with the help of **residue class** considerations, however, for the group structure we need a **representation theoretic analysis.**

★ Carry propery

$\mathbb{U}_d$ ... set of unitrary $d \times d$ matrices, $f_\lambda(n) = f(n \bmod q^\lambda)$

Definition

A function $f : \mathbb{N} \to \mathbb{U}_d$ has the *Carry property* if there exists $\eta > 0$ such that uniformly for $(\lambda, \alpha, \rho) \in \mathbb{N}^3$ with $\rho < \lambda$, the number of integers $0 \le \ell < k^\lambda$ such that there exists $(n_1, n_2) \in \{0, \ldots, q^\alpha - 1\}^2$ with

$$f(\ell q^\alpha + n_1 + n_2)^H f(\ell q^\alpha + n_1) \ne f_{\alpha+\rho}(\ell q^\alpha + n_1 + n_2)^H f_{\alpha+\rho}(\ell q^\alpha + n_1)$$

is at most $O(q^{\lambda - \eta\rho})$ where the implied constant may depend only on q and f.

★ Fourier propery

Definition

Given a non-decreasing function $\gamma : \mathbb{R} \to \mathbb{R}$ satisfying $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$ and $c > 0$ we let $F_{\gamma,c}$ denote the set of functions $f : \mathbb{N} \to \mathbb{U}_d$ such that for $(\alpha, \lambda) \in \mathbb{N}^2$ with $\alpha \le c\lambda$ and $t \in \mathbb{R}$:

$$\left\| q^{-\lambda} \sum_{u<q^\lambda} f(uq^\alpha)\mathrm{e}(-ut) \right\|_F \le q^{-\gamma(\lambda)}.$$

We say in this case that f has the *Fourier property*.

★ Automatic sequences along primes

Theorem (Müllner 2017, generalization of Mauduit+Rivat 2015)

Let $\gamma : \mathbb{R} \to \mathbb{R}$ *be a non-decreasing function satisfying* $\lim_{\lambda\to\infty} \gamma(\lambda) = +\infty$, *and* $f \in \mathcal{F}_{\gamma,c}$ *be a function satisfying the Carry property for some* $\eta \in (0,1]$ *and the Fourier property for some* $c \geq 10$*. Then for any* $\theta \in \mathbb{R}$ *we have*

$$\left\| \sum_{n\le x} \Lambda(n) f(n) \mathrm{e}(\theta n) \right\|_2 \ll c_1(q) (\log x)^{c_2(q)} x q^{-\eta\gamma(2\lfloor (\log x)/(80\log q)\rfloor)/20}.$$

★ Automatic sequences along squares

Theorem (Generalization of Mauduit+Rivat 2018)

Let $\gamma : \mathbb{R} \to \mathbb{R}$ *be a nondecreasing function satisfying* $\lim_{\lambda\to\infty} \gamma(\lambda) = \infty$, *and let* $f \in \mathcal{F}_{\gamma,c}$ *be a function satisfying the Carry property for some* $\eta \in (0,1]$ *and the Fourier property for some* $c \geq 18$. *Then for any* $\theta \in \mathbb{R}$, *we have*

$$\left\| \sum_{0<n\leq x} f(n^2)\mathrm{e}(n\theta) \right\|_2 \ll_{d,f,q} (\log x)^{\omega(q)+2} \left(xq^{-\frac{\eta\gamma(2\lfloor (3\log x)/(100\log q)\rfloor)}{56}} \right),$$

where the absolute implied constant only depends on d, f *and* q.

★ Automatic sequences along primes and squares

Application for

$$f(n) = D(T(n)),$$

where $D : G \to \mathbb{U}_d$ is a unitary and irreducible representation of the group G.

By Fourier analysis on G and by considering residue classes (for $s(n)$) leads to the (limiting) **distributional behavior** of

$$a(n) = f(s(n), T(n))$$

along **primes** and **squares**.

Thank you!

(Logarithmic) Densities for Automatic Sequences along Primes and Squares

Michael Drmota

joint work with Boris Adamczewski and Clemens Müllner

supported by the Austrian Science Foundation FWF, project F5502

Institut für Diskrete Mathematik und Geometrie
Technische Universität Wien

One World Numeration Seminar
March 30, 2021