Subsequences of Automatic Sequences and Uniform Distribution

Michael Drmota

Institut für Diskrete Mathematik und Geometrie
Technische Universität Wien

Uniform Distribution and Quasi-Monte Carlo Methods, RICAM Linz,
October 14–18, 2013

Summary

- ★ Thue-Morse sequence
- ★ Gelfond problems
- ★ Automatic sequences
- ★ Subsequences along $\lfloor n^c \rfloor$
- ★ Subsequences along squares
- ★ Subsequences along polynomial values
- ★ Subsequences along primes

★ 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 $(t_n)_{n \geqslant 0}$:

01101001100101101001011001101001

★ 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

- 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 their linear subsequences are **automatic sequences**.

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$01101001100101101001011001101001100101100110 1\cdots$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$0110100110010110100101100110100110010110011010 \cdots$$

★ Thue-Morse sequence

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

$$0110100110010110100101100110100110010110011 01\cdots$$

$$\#\{0 \leqslant n < N : t_n = 0\} \sim \frac{N}{2}$$

★ 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 $(t_n)_{n\geqslant 0}$:

$$0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad \cdots$$

$$\#\{0 \leqslant n < N : t_n = 0\} \sim \frac{N}{2}$$

$$\#\{0 \leqslant n < N : t_{3n} = 0\} \sim \frac{N}{2}$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

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

Mauduit and Rivat (2010):

$$\#\{0 \leqslant p < N : t_p = 0\} \sim \frac{\pi(N)}{2}$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$0110100110010110100101100110100110010110 0110 1001 \cdots$$

★ Thue-Morse sequence

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

01 1 0 1 1 0 ...

Mauduit and Rivat (2009):

$$\#\{0 \leqslant n < N : t_{n^2} = 0\} \sim \frac{N}{2}$$

★ Thue-Morse sequence

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

$$01 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0 \quad \cdots$$

Mauduit and Rivat (2009):

$$\#\{0 \leqslant n < N : t_{n^2} = 0\} \sim \frac{N}{2}$$

D., Mauduit and Rivat (2013+):

$$\#\{0 \leqslant n < N : t_{n^2} = b_0, t_{(n+1)^2} = b_1, \ldots, t_{(n+k-1)^2} = b_{k-1}\} \sim \frac{N}{2^k}$$

★ Gelfond Problems

Gelfond 1967/1968

a, m ... positive integers, b, ℓ ... non-neg. integer, $(m, q-1) = 1$.

$$\Longrightarrow \quad \boxed{\#\{n < N : \; s_q(an+b) \equiv \ell \bmod m\} = \frac{N}{m} + O(N^\lambda)}$$

with $0 < \lambda < 1$.

★ Gelfond Problems

Gelfond 1967/1968

a, m ... positive integers, b, ℓ ... non-neg. integer, $(m, q-1) = 1$.

$$\Longrightarrow \quad \boxed{\#\{n < N :\ s_q(an+b) \equiv \ell \bmod m\} = \frac{N}{m} + O(N^\lambda)}$$

with $0 < \lambda < 1$.

In particular:

$$\begin{aligned}\#\{n < N : t_{an+b} = 0\} &= \#\{n < N : s_2(an+b) \equiv 0 \bmod 2\} \\ &= \frac{N}{2} + O(N^\lambda)\end{aligned}$$

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \neq j$, $(m_j, q_j - 1) = 1$:

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \neq j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \ne j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

2. $(m, q - 1) = 1$:

$$\#\{\text{primes } p < N : s_q(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \ne j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

2. $(m, q - 1) = 1$: Mauduit, Rivat 2010

$$\#\{\text{primes } p < N : s_q(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \ne j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

2. $(m, q-1) = 1$: Mauduit, Rivat 2010

$$\#\{\text{primes } p < N : s_q(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

3. $(m, q-1) = 1$, $P(x) \in \mathbb{N}[x]$:

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

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \neq j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

2. $(m, q-1) = 1$: Mauduit, Rivat 2010

$$\#\{\text{primes } p < N : s_q(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

3. $(m, q-1) = 1$, $P(x) \in \mathbb{N}[x]$: Mauduit, Rivat 2009 for $P(x) = x^2$

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

★ Gelfond Problems

1. $q_1, q_2, \ldots, q_d \geqslant 2$, $(q_i, q_j) = 1$ for $i \neq j$, $(m_j, q_j - 1) = 1$: Kim 1999

$$\#\{n < N : s_{q_j}(n) \equiv \ell_j \bmod m_j,\ 1 \leqslant j \leqslant d\} = \frac{N}{m_1 \cdots m_d} + O(N^{1-\eta})$$

2. $(m, q-1) = 1$: Mauduit, Rivat 2010

$$\#\{\text{primes } p < N : s_q(p) \equiv \ell \bmod m\} = \frac{\pi(N)}{m} + O(N^{1-\eta})$$

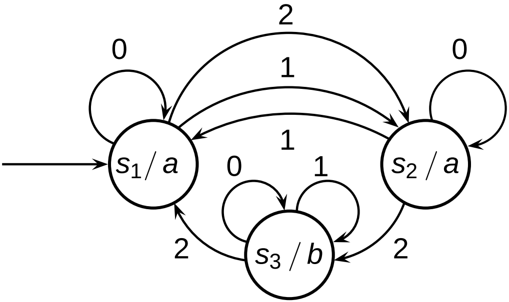
3. $(m, q-1) = 1$, $P(x) \in \mathbb{N}[x]$: Mauduit, Rivat 2009 for $P(x) = x^2$
Drmota, Mauduit, Rivat 2011 for **large bases** $q \geqslant q_0(\deg(P))$

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

★ Automatic sequences

Definition

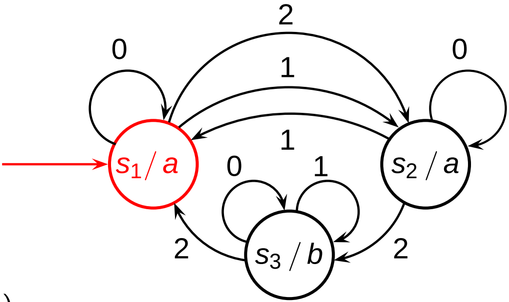
A sequence $(u_n)_{n \geqslant 0}$ is called a *q-automatic sequence*, if u_n is the output of an automaton when the input is the q-ary expansion of n.



★ Automatic sequences

Definition

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

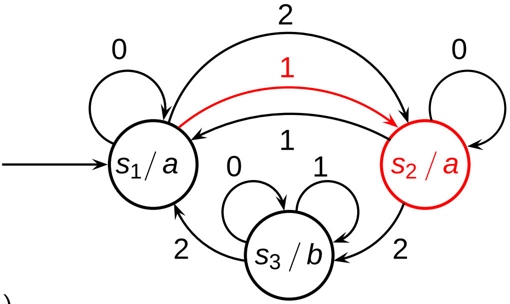


$32 = (1012)_3$

★ Automatic sequences

Definition

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

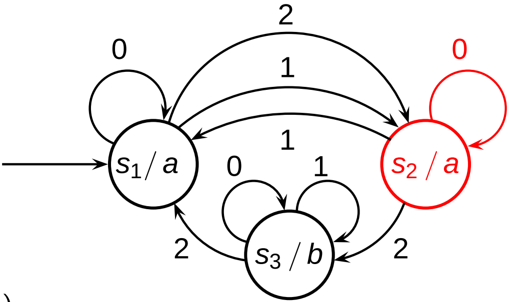


$32 = (1012)_3$

★ Automatic sequences

Definition

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

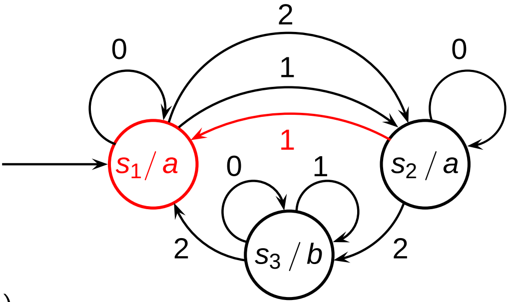


$32 = (1012)_3$

★ Automatic sequences

Definition

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

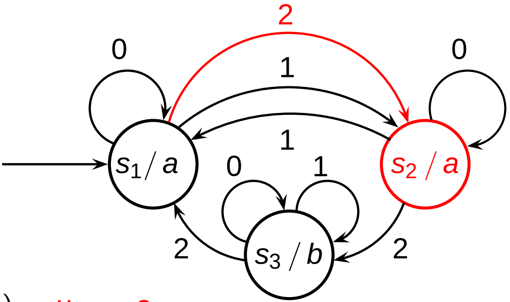


$32 = (1012)_3$

★ Automatic sequences

Definition

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

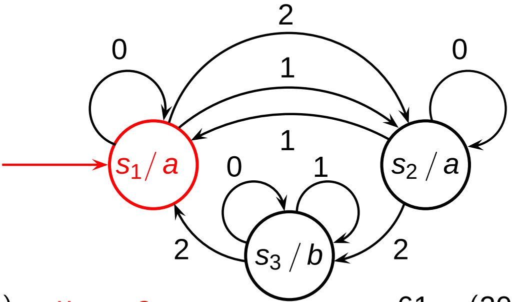


$32 = (1012)_3 \quad u_{32} = a,$

★ Automatic sequences

Definition

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

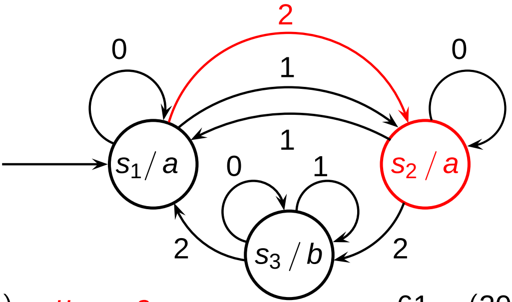


$32 = (1012)_3$ $\quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

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

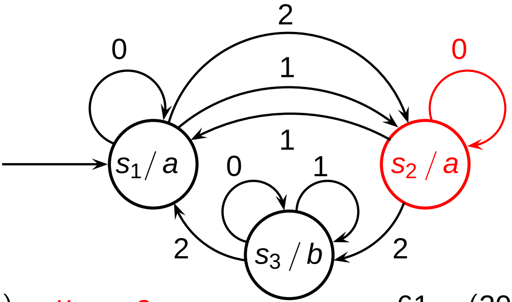


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

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

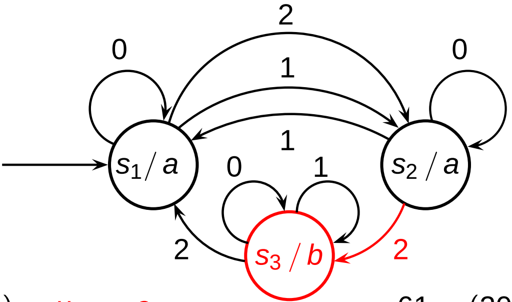


$32 = (1012)_3$ $\quad u_{32} = a$, $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

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

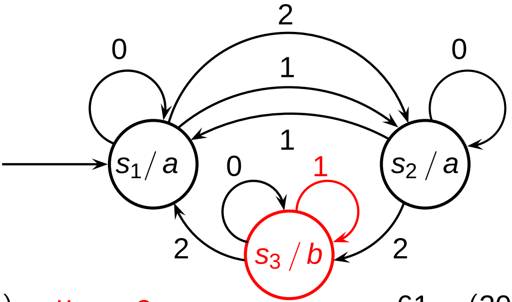


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3$

★ Automatic sequences

Definition

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

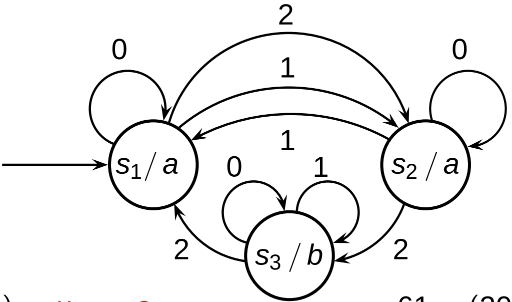


$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3 \quad u_{61} = b$

★ Automatic sequences

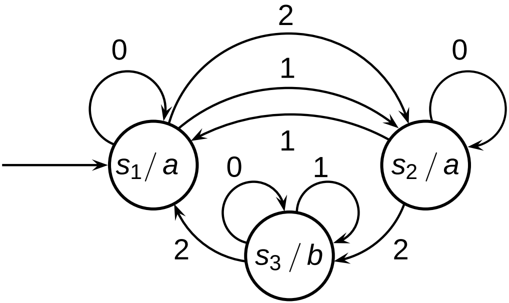
Definition

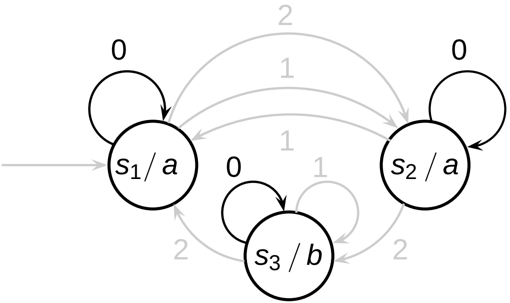
A sequence $(u_n)_{n\geqslant 0}$ is called a *q-automatic sequence*, if u_n is the output of an automaton when the input is the q-ary expansion of n.



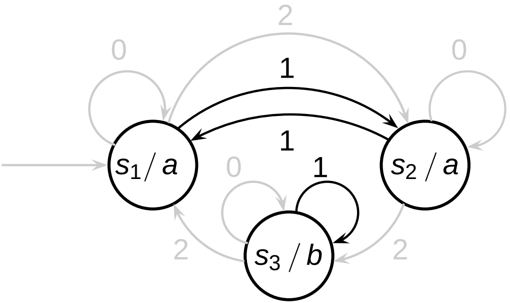
$32 = (1012)_3 \quad u_{32} = a,$ $\qquad 61 = (2021)_3 \quad u_{61} = b$

$(u_n)_{n\geqslant 0}$: *aaaaabaabaabaaabbaaabaaabbaaabaaabbaaaaaaba*...

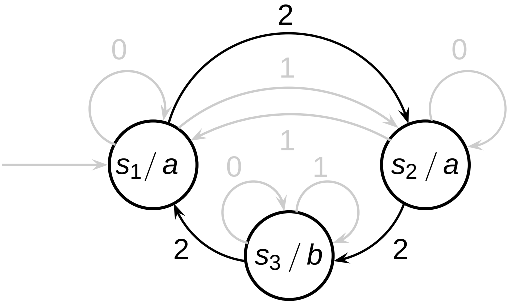

2
0
1
0
1
s_1 / a
0
1
s_2 / a
2
s_3 / b
2



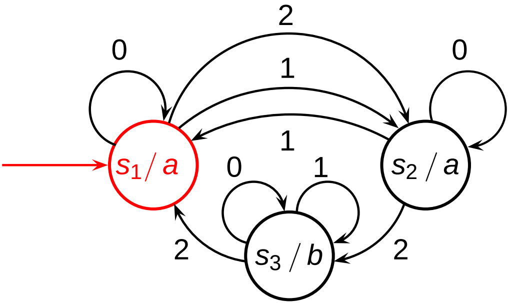
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



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

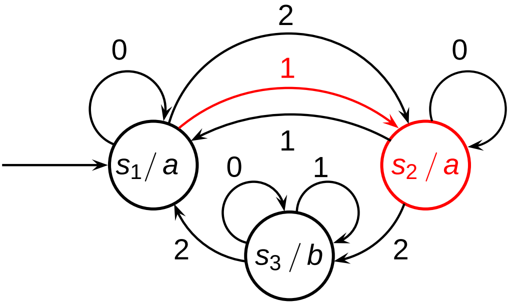


$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$



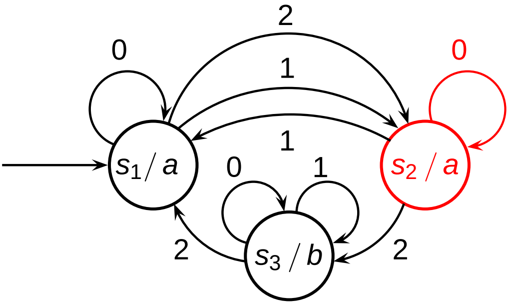
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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



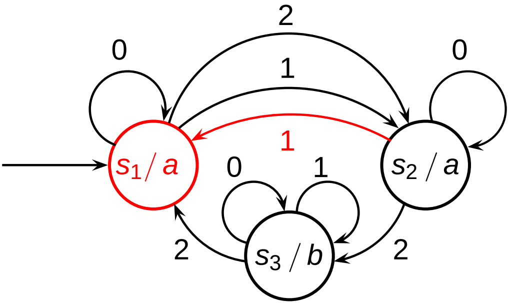
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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



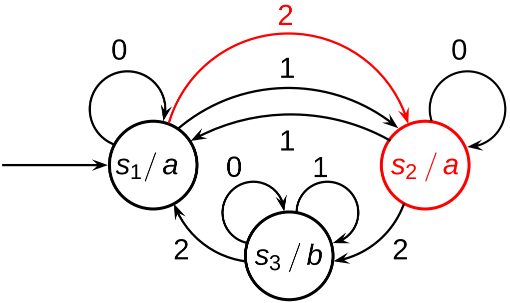
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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



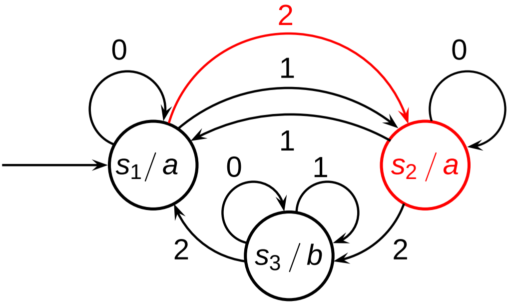
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

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



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

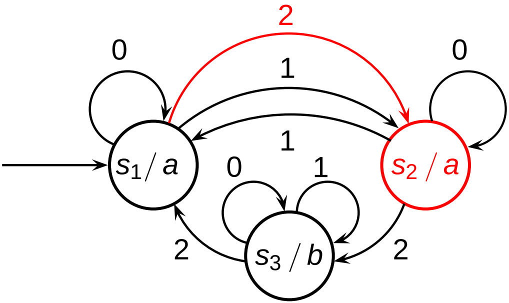
$$32 = (1012)_3 : \qquad M_2 \circ M_1 \circ M_0 \circ M_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$



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

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

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



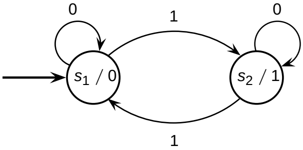
$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Definition

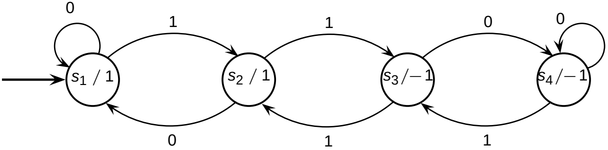
A q-automatic sequence is called *invertible* if there exists an automaton such that all transition matrices are invertible and M_0 is the identity matrix.

Examples of automatic sequences

Thue-Morse sequence (invertible): $t_n = \sum_{j \geqslant 0} \varepsilon_j(n) \bmod 2$



Rudin-Shapiro sequence (**not** invertible): $r_n = \sum_{j \geqslant 0} \varepsilon_j(n)\varepsilon_{j+1}(n) \bmod 2$



Examples of automatic sequences

Sum-of-digits-function (invertible): $u_n = s_q(n) \bmod m$

q-additive function modulo m (invertible): $u_n = f(n) \bmod m$

$$f(n) = \sum_{j \geqslant 0} f(\varepsilon_j(n)) \quad \text{and} \quad f(0) = 0.$$

★ Subsequences of Automatic Sequences

★ **General concept:**

1. Start with an **automatic sequence** u_n that is uniformly distributed on the output alphabet.
   (Recall: u_n has at most linear subword complexity)
2. Consider a relatively sparse **subsequence** u_{n_k} that has the same asymptotic frequencies.
   (It is assumed that the average size of the gaps increases sufficiently fast so that one can expect **random properties**)
3. This subsequence should be **normal** on the output alphabet

This leads to a new class of pseudo-random sequences!!

★ Subsequences along $\lfloor n^c \rfloor$

Theorem (Deshouillers, D. and Morgenbesser, 2012)

Let u_n be a q-automatic sequence (on an alphabet $\mathcal{A}$) and

$$1 < c < 7/5.$$

Then for each $a \in \mathcal{A}$ then asymptotic density $\mathrm{dens}(u_{\lfloor n^c \rfloor}, a)$ of a in the subsequence $u_{\lfloor n^c \rfloor}$ exists if and only if the asymptotic density of α in u_n exists and we have

$$\mathrm{dens}(u_{\lfloor n^c \rfloor}, a) = \mathrm{dens}(u_n, a).$$

★ Subsequences along $\lfloor n^c \rfloor$

Theorem (Deshouillers, D. and Morgenbesser, 2012)

Let u_n be a q-automatic sequence (on an alphabet $\mathcal{A}$) and

$$1 < c < 7/5.$$

Then for each $a \in \mathcal{A}$ then asymptotic density $\mathrm{dens}(u_{\lfloor n^c \rfloor}, a)$ of a in the subsequence $u_{\lfloor n^c \rfloor}$ exists if and only if the asymptotic density of α in u_n exists and we have

$$\mathrm{dens}(u_{\lfloor n^c \rfloor}, a) = \mathrm{dens}(u_n, a).$$

In particular it follows that for $\boxed{u_n = s_q(n) \bmod m}$:

$$\mathrm{dens}(u_{\lfloor n^c \rfloor}, a) = \lim_{N \to \infty} \frac{1}{N} \# \{ n < N : s_q(\lfloor n^c \rfloor) \equiv a \bmod m \} = \frac{1}{m}.$$

It is **conjectured** that this holds for all non-integers $c > 1$.

★ Subsequences of the form $\lfloor n^c \rfloor$

★ **Partial answer:**

Theorem (Morgenbesser, 2011)

For every non-integer $c > 0$ *there exists* $q_0(c) \geqslant 2$ *such that for all* $\boxed{q \geqslant q_0(c)}$ *the densities of the* $\lfloor n^c \rfloor$*-subsequences of* $\boxed{u_n = s_q(n) \bmod m}$ *are given by*

$$\mathrm{dens}(u_{\lfloor n^c \rfloor}, a) = \frac{1}{m}.$$

★ Subsequences along $\lfloor n^c \rfloor$

Theorem (Deshouillers, D. and Morgenbesser, 2012)

Let $\boxed{1 < c < 10/9}$ *and let* t_n *denote the* **Thue-Morse sequence**. *Then for every pair* $(a, b) \in \{0, 1\}^2$ *we have*

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leqslant n < N : (t_{\lfloor n^c \rfloor}, t_{\lfloor (n+1)^c \rfloor}) = (a, b)\} = \frac{1}{4}.$$

★ Subsequences along $\lfloor n^c \rfloor$

Theorem (Deshouillers, D. and Morgenbesser, 2012)

Let $\boxed{1 < c < 10/9}$ *and let* t_n *denote the* **Thue-Morse sequence***. Then for every pair* $(a, b) \in \{0, 1\}^2$ *we have*

$$\lim_{N\to\infty} \frac{1}{N} \#\{1 \leqslant n < N : (t_{\lfloor n^c \rfloor}, t_{\lfloor (n+1)^c \rfloor}) = (a, b)\} = \frac{1}{4}.$$

Conjecture

For every non-integer $c > 1$ the sequence $t_{\lfloor n^c \rfloor}$ is **normal**.
(Every block $(b_0, b_1, \ldots, b_{k-1})$ of length $k \geqslant 1$ appears with asymptotic frequency 2^{-k}.)

★ Subsequences along squares

Theorem (D. and Morgenbesser, 2012)

Let $q \geqslant 2$ and $(u_n)_{n \geqslant 0}$ an **invertible q-automatic sequence***. Then the densities* $\boxed{\mathrm{dens}(u_{n^2}, a)}$ *exist for each letter $a \in \mathcal{A}$.*

★ Subsequences along squares

Theorem (D. and Morgenbesser, 2012)

Let $q \geqslant 2$ and $(u_n)_{n\geqslant 0}$ an **invertible** q**-automatic sequence**. *Then the densities* $\boxed{\mathrm{dens}(u_{n^2}, a)}$ *exist for each letter* $a \in \mathcal{A}$.

Theorem (D., Mauduit and Rivat, 2013+)

Let $(t_n)_{n\geqslant 0}$ denote the **Thue-Morse sequence**. *Then the sequence $(t_{n^2})_{n\geqslant 0}$ is* **normal** *on the alphabet $\{0, 1\}$.*

★ Subsequences along squares

Theorem (D. and Morgenbesser, 2012)

Let $q \geqslant 2$ and $(u_n)_{n \geqslant 0}$ an **invertible** q**-automatic sequence**. *Then the densities* $\boxed{\mathrm{dens}(u_{n^2}, a)}$ *exist for each letter* $a \in \mathcal{A}$.

Theorem (D., Mauduit and Rivat, 2013+)

Let $(t_n)_{n \geqslant 0}$ denote the **Thue-Morse sequence**. *Then the sequence $(t_{n^2})_{n \geqslant 0}$ is* **normal** *on the alphabet $\{0, 1\}$.*

Theorem (Mauduit and Rivat, 2013+)

Let $r(n)$ denote the **Rudin-Shapiro sequence**. *Then* $\boxed{\mathrm{dens}(r(n^2), a) = 1/2}$ *for* $a \in \{0, 1\}$.

★ Fourier estimates

Truncated sum-of-digits function

$$s_{2,\lambda}(n + k2^\lambda) = s_2(n), \quad 0 \leqslant n < 2^\lambda,\ k \geqslant 0.$$

$s_{2,\lambda}$ is periodic with period 2^λ

Discrete Fourier transform

$$F_\lambda(h,\alpha) = \frac{1}{2^\lambda} \sum_{0 \leqslant u < 2^\lambda} e(\alpha s_{2,\lambda}(u) - h2^{-\lambda})$$

of the function $n \mapsto e(\alpha s_{q,\lambda}(n))$; $e(x) = \exp(2\pi i x)$.

★ Fourier estimates

Uniform upper bounds.

$$|F_\lambda(h,\alpha)| \leqslant 2^{-c\|\alpha\|^2\lambda}$$

(for some constant $c > 0$, $\|x\| = \min\{|x-k| : k \in \mathbb{Z}\}$).

This upper bound (together with two applications of the Van-der-Corput inequality, a proper Fourier analysis and estimates for quadratic exponential sums) leads to

$$\left|\sum_{n<N} e(\alpha s_2(n^2))\right| \ll N^{1-c'\|\alpha\|^2}$$

(for some constant $c' > 0$) and consequently to

$$\#\{0 \leqslant n < N : t_{n^2} = 0\} \sim \frac{N}{2}$$

★ Fourier estimates

Fourier term with correlations in oder to handle blocks of length > 1:

$$G_\lambda(h,d) = \frac{1}{2^\lambda} \sum_{0 \leqslant u < 2^\lambda} e\left(\frac{1}{2} \sum_{\ell=0}^{k-1} \alpha_\ell s_{2,\lambda}(u + \ell d) - h 2^{-\lambda} \right)$$

$(\alpha_0, \ldots, \alpha_{k-1} \in \{0,1\})$

Uniform upper bounds.

$$|G_\lambda(h,d)| \leqslant 2^{-c''\lambda}$$

(for some constant $c'' > 0$)

★ Fourier estimates

After several quite technical steps (in particular with a subtle **Fourier analysis** this leads to upper bounds for the exponential sums

$$\sum_{n<N} e\left(\frac{1}{2}\sum_{\ell=0}^{k-1}\alpha_\ell s_2((n+\ell)^2)\right) \ll N^{1-\eta}$$

and consequently to the proof that t_{n^2} is normal.

★ Subsequences along polynomials

Theorem (D., Mauduit and Rivat, 2011)

For every $d \geqslant 2$ there exists $q_0(d) \geqslant 2$ such that for all prime $\boxed{q \geqslant q_0(c)}$ and all integer polynomials $P(x)$ of degree d (where the leading coefficient if coprime to q)

$$\#\{1 \leqslant n < N : s_q(P(n)) \equiv a \bmod m\} = \frac{N}{m} + O\left(N^{1-\eta}\right)$$

for some $\eta > 0$ and all integers m with $(m, q-1) = 1$.

★ Subsequences along polynomials

Theorem (D., Mauduit and Rivat, 2011)

For every $d \geqslant 2$ there exists $q_0(d) \geqslant 2$ such that for all prime $\boxed{q \geqslant q_0(c)}$ and all integer polynomials $P(x)$ of degree d (where the leading coefficient if coprime to q)

$$\#\{1 \leqslant n < N : s_q(P(n)) \equiv a \bmod m\} = \frac{N}{m} + O\left(N^{1-\eta}\right)$$

for some $\eta > 0$ and all integers m with $(m, q-1) = 1$.

Conjecture

Suppose that $(q-1, m) = 1$. Then for every integer polynomial $P(x)$ of degree $\geqslant 2$ the sequence $\boxed{s_q(P(n)) \bmod m}$ is **normal** on the alphabet $\{0, 1, \ldots, q-1\}$.

★ Subsequences along primes

Theorem (Mauduit and Rivat, 2010)

Let $u_n = s_q(n) \bmod m$ *(and suppose that* $(q-1,m)=1$*). Then we have*

$$\mathrm{dens}((u_p)_{p\in\mathbb{P}}, a) = \frac{1}{m}$$

for $a \in \{0,1,\ldots,m-1\}$.

★ Subsequences along primes

Theorem (Mauduit and Rivat, 2010)

Let $u_n = s_q(n) \bmod m$ *(and suppose that* $(q-1, m) = 1$*). Then we have*

$$\mathrm{dens}((u_p)_{p\in\mathbb{P}}, a) = \frac{1}{m}$$

for $a \in \{0, 1, \ldots, m-1\}$*.*

Theorem (Mauduit and Rivat, 2013+)

Let r_n *denote the* **Rudin-Shapiro sequence***. Then we have*

$$\mathrm{dens}((r_p)_{p\in\mathbb{P}}, 0) = \mathrm{dens}((r_p)_{p\in\mathbb{P}}, 1) = \frac{1}{2}.$$

★ Subsequences along primes

Theorem (D., 2013)

Let $q \geqslant 2$ *and* $(u_n)_{n\geqslant 0}$ *an* **invertible** q**-automatic sequence**. *Then the densities* $\mathrm{dens}((u_p)_{p\in\mathbb{P}}, a)$ *exist for each letter* $a \in \mathcal{A}$.

★ Subsequences along primes

Theorem (D., 2013)

Let $q \geqslant 2$ and $(u_n)_{n \geqslant 0}$ an **invertible q-automatic sequence***. Then the densities $\mathrm{dens}((u_p)_{p\in\mathbb{P}}, a)$ exist for each letter $a \in \mathcal{A}$.*

Conjecture

For every automatic sequence $(u_n)_{n\geqslant 0}$ the logarithmic densities $\mathrm{log\text{-}dens}(u_p)_{p\in\mathbb{P}}, a)$ exist for all letters $a \in \mathcal{A}$.

★ Local results the sum-of-digits function on primes

Theorem (D., Mauduit and Rivat, 2009)

Suppose that $(q, k-1) = 1$. *Then*

$$\#\{primes\ p < N : s_q(p) = k\}$$
$$= \frac{q-1}{\varphi(q-1)} \frac{\pi(N)}{\sqrt{2\pi\sigma_q^2 \log_q N}} \left(\exp\left(-\frac{(k - \mu_q \log_q N)^2}{2\sigma_q^2 \log_q N} \right) + O((\log N)^{-\frac{1}{2}+$$

where

$$\mu_q := \frac{q-1}{2}, \quad \sigma_q^2 := \frac{q^2-1}{12}.$$

Remark: This asymptotic expansion is only significant if

$$\left| k - \mu_q \log_q N \right| \leqslant C(\log N)^{\frac{1}{2}}$$

Note that $\frac{1}{\pi(N)} \sum_{p<N} s_q(p) \sim \mu_q \log_q N$.

Corollary

Theorem (D., Mauduit, and Rivat, 2011)

$s_2(n)$ ... *number of powers of 2 in the binary expansion of n*

$$\#\{primes\ p < 2^{2k} : s_2(p) = k\} \sim \frac{2^{2k}}{\sqrt{2\pi}\log 2\, k^{\frac{3}{2}}}$$

★ Binary Representation of Primes

Corollary

Theorem (D., Mauduit, and Rivat, 2011)

$s_2(n)$ ... *number of powers of 2 in the binary expansion of n*

$$\#\{primes\ p < 2^{2k} : s_2(p) = k\} \sim \frac{2^{2k}}{\sqrt{2\pi}\log 2\, k^{\frac{3}{2}}}$$

In particular for every (sufficiently large) positive integer k there exist a prime p with

$$s_2(p) = k.$$

Thank you!