The Digital Expansion of Squares and Primes

Michael Drmota

Institut für Diskrete Mathematik und Geometrie
Technische Universität Wien

Conference in Number Theory and Applications in Memory of
Pr. Christian Mauduit (1959–2019), 9-12 Apr 2026 Hammamet,
Tunesia

Christian Mauduit 1959–2019

Summary

- ★ Thue-Morse sequence
- ★ Gelfond-Problems
- ★ Sarnak Conjecture
- ★ Combined Problems
- ★ Zeckendorf Expansion
- ★ Methods
- ★ Open Problems

★ 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

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

$$011010011001011010010110011010011001011001101\cdots$$

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

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

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

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$01\,1010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

$$\#\{0\leqslant n<N : t_n=0\}=\#\{0\leqslant n<N : s_2(n)\equiv 0 \bmod 2\}\sim\frac{N}{2}$$

★ Thue-Morse sequence

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

$$011010011001011010010110011010011001011001101\cdots$$

$$\#\left\{0\leqslant n<N : t_n=0\right\}=\#\left\{0\leqslant n<N : s_2(n)\equiv 0 \bmod 2\right\}\sim\frac{N}{2}$$

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

★ Thue-Morse sequence

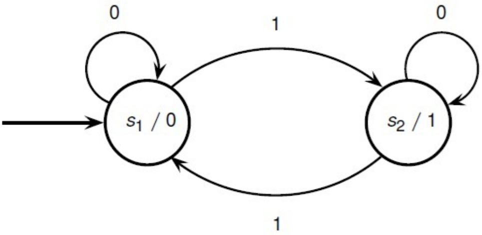
- TM sequence is **not periodic** and **cubeless**.
- TM sequence is **almost periodic**:
  Every appearing consecutive block appears infinitely many times with bounded gaps.
- **Subword complexity is linear**: $p_k \leqslant \frac{10}{3}k$
  p_k ... subword complexity (*number of different consecutive 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 their linear subsequences are **automatic sequences**.

★ Thue-Morse sequence

Automaton that generates the Thue-Morse sequence:
$t_n = \sum_{j\geqslant 0} \varepsilon_j(n) \bmod 2 = s_2(n) \bmod 2$



★ Background 1: Gelfond Problems

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

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

★ Background 1: 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

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

★ Background 1: 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

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

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

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

★ Background 1: 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

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

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

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

★ Background 1: 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

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

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

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

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

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

★ Background 1: 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

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

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

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

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

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

★ Background 1: 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

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

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

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

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

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

★ Background 1: 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

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

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

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

3. $(m, q - 1) = 1$, $P(x) \in \mathbb{N}[x]$: Mauduit, Rivat 2009: $P(x) = x^2$
D., Mauduit, Rivat 2011: large bases, Spiegelhofer: $P(x) = x^3$

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

★ Background 1: Gelfond Problems

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Background 1: Gelfond Problems

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

$$0110100110010110100101100110100110010110011 01\cdots$$

★ Background 1: Gelfond Problems

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

10 0 1 1 1 0 1 0 0 1 1 1 0 · · ·

★ Background 1: Gelfond Problems

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

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

Theorem (Mauduit, Rivat 2010)

Suppose that $q \geqslant 2$ and $m \geqslant 2$ withe $(m, q-1) = 1$. Then for all ℓ

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

for some $\eta > 0$.

★ Background 1: Gelfond Problems

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

$$011010011001011010010110011010011001011001101\cdots$$

★ Background 1: Gelfond Problems

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

$011010011001011010010110011010011001011001101\cdots$

★ Background 1: Gelfond Problems

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

01 1 0 1 1 0 $\cdots$

★ Background 1: Gelfond Problems

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

01 1 0 1 1 0 ⋯

Theorem (Mauduit, Rivat 2009)

Suppose that $q \geqslant 2$ *and* $m \geqslant 2$ *withe* $(m, q-1) = 1$*. Then for all* ℓ

$$\#\{n < N : s_q(n^2) \equiv \ell \bmod m\} = \frac{N}{m} + O(N^{1-\eta})$$

for some $\eta > 0$*.*

★ Background 1: Gelfond Problems

Theorem (D., Mauduit and Rivat, 2011)

For every $d \geqslant 2$ *there exists* $q_0(d) \geqslant 2$ *such that for all prime* $\boxed{q \geqslant q_0(d)}$ *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* $(m, q-1) = 1$*.*

★ Background 1: Gelfond Problems

Theorem (D., Mauduit and Rivat, 2011)

For every $d \geqslant 2$ *there exists* $q_0(d) \geqslant 2$ *such that for all prime* $\boxed{q \geqslant q_0(d)}$ *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* $(m, q-1) = 1$*.*

Theorem (Spiegelhofer)

$$\#\{1 \leqslant n < N : s_2(n^3) \equiv 0 \bmod 2\} = \frac{N}{2} + O\left(N^{1-\eta}\right)$$

for some $\eta > 0$*.*

★ Background 1: Gelfond Problems

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*.

★ Background 1: Gelfond Problems

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.

For example, the sequences $u_n = s_q(n) \bmod m$ are automatic.

★ Background 1: Gelfond Problems

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.

For example, the sequences $u_n = s_q(n) \bmod m$ are automatic.

Lemma

For every automatic sequence $(u_n)_{n\geqslant 0}$ *the* **densities**

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

or, if not, the **logarithmic densities**

$$\mathrm{logdens}(u_n; a) = \lim_{N\to\infty} \frac{1}{\log N} \sum_{n\leqslant N} \frac{1}{n} \mathbf{1}_{u_n = a}$$

exist.

★ Background 1: Gelfond Problems

Theorem (Adamczewski+D.+Müllner 2020)

For every automatic sequence $(u_n)_{n\geqslant 0}$ *the logarithmic densities*

$$\boxed{\mathrm{logdens}(u_{p_n}, a)}$$

of the subsequence along **prime numbers** *and the logarithmic densities*

$$\boxed{\mathrm{logdens}(u_{n^2}, a)}$$

of the subsequence along **squares** *exist and are computable. Furthermore, the densities* $\boxed{\mathrm{dens}(u_{p_n}, a)}$ *and* $\boxed{\mathrm{dens}(u_{n^2}, a)}$ *exist if and only if the densities* $\mathrm{dens}(u_n, a)$ *exist.*

★ Background 1: Gelfond Problems

Theorem (Adamczewski+D.+Müllner 2020)

For every automatic sequence $(u_n)_{n\geqslant 0}$ *the logarithmic densities*

$$\boxed{\mathrm{logdens}(u_{p_n}, a)}$$

of the subsequence along **prime numbers** *and the logarithmic densities*

$$\boxed{\mathrm{logdens}(u_{n^2}, a)}$$

of the subsequence along **squares** *exist and are computable. Furthermore, the densities* $\boxed{\mathrm{dens}(u_{p_n}, a)}$ *and* $\boxed{\mathrm{dens}(u_{n^2}, a)}$ *exist if and only if the densities* $\mathrm{dens}(u_n, a)$ *exist.*

This is the most general result related to the Gelfond problems so far.

★ Background 2: Sarnak Conjecture

A bounded complex valued sequence $u(n)$, $n \in \mathbb{N}$, is said to be **deterministic** if for every $\epsilon > 0$ the set $\{(u(n+1), \ldots, u(n+m)) : n \in \mathbb{N}\}$ can be covered by $O(\exp(o(m)))$ balls of radius ϵ (as $m \to \infty$).

For example, if $\boxed{u(n) := f(T^n x)}$ for a minimal topological dynamical system (X, T) with **zero topological entropy** (and a continuous function f) then $u(n)$ is deterministic.

★ Background 2: Sarnak Conjecture

A bounded complex valued sequence $u(n)$, $n \in \mathbb{N}$, is said to be **deterministic** if for every $\epsilon > 0$ the set $\{(u(n+1), \ldots, u(n+m)) : n \in \mathbb{N}\}$ can be covered by $O(\exp(o(m)))$ balls of radius ϵ (as $m \to \infty$).

For example, if $\boxed{u(n) := f(T^n x)}$ for a minimal topological dynamical system (X, T) with **zero topological entropy** (and a continuous function f) then $u(n)$ is deterministic.

In particular **automatic sequences** $u(n)$ are deterministic.
And every sequence $u(n) := f(T^n x)$, where $x \in X$ and (X, T) is generated from an **automatic sequence** is deterministic.

★ Background 2: Sarnak Conjecture

Definition (Möbius-function)

$$\mu(n) = \begin{cases} 1 & n = 1, \\ (-1)^r & n = p_1 p_2 \cdots p_r, \\ 0 & n \text{ is not square free.} \end{cases}$$

★ Background 2: Sarnak Conjecture

Definition (Möbius-function)

$$\mu(n) = \begin{cases} 1 & n = 1, \\ (-1)^r & n = p_1 p_2 \cdots p_r, \\ 0 & n \text{ is not square free.} \end{cases}$$

$\mu(3) = \mu(5) = -1$, $\mu(6) = \mu(35) = 1$

★ Background 2: Sarnak Conjecture

Theorem

$$\sum_{n=1}^{N} \mu(an+b) = o(N) \qquad (N \to \infty)$$

About **half** *of the (square-free) numbers of the form $an+b$ with $n \leqslant N$ have an* **even number** *of has prime divisors and the other* **half** *has an* **odd number** *of prime divisors.*

★ Background 2: Sarnak Conjecture

Theorem

$$\sum_{n=1}^{N} \mu(an+b) = o(N) \qquad (N \to \infty)$$

About **half** *of the (square-free) numbers of the form $an+b$ with $n \leqslant N$ have an* **even number** *of has prime divisors and the other* **half** *has an* **odd number** *of prime divisors.*

The relation $\sum_{n=1}^{N} \mu(n) = o(N)$ is equivalent to the **prime number theorem** $\pi(x) \sim x/\log x$ and the general property $\sum_{n=1}^{N} \mu(an+b) = o(N)$ is equivalent to the **Dirichlet prime number theorem** $\pi(x;a,m) \sim x/(\varphi(m)\log x)$.

★ Background 2: Sarnak Conjecture

A sequence $u(n)$ is **orthogonal to the Möbius function** $\mu(n)$ if

$$\sum_{n\le N} \mu(n)u(n) = o(N) \qquad (N \to \infty).$$

★ Background 2: Sarnak Conjecture

A sequence $u(n)$ is **orthogonal to the Möbius function** $\mu(n)$ if

$$\sum_{n\le N} \mu(n)u(n) = o(N) \qquad (N \to \infty).$$

For example, a constant sequence and periodic sequences are orthorgonal to the Möbius function

★ Background 2: Sarnak Conjecture

A sequence $u(n)$ is **orthogonal to the Möbius function** $\mu(n)$ if

$$\sum_{n\le N} \mu(n)u(n) = o(N) \qquad (N \to \infty).$$

For example, a constant sequence and periodic sequences are orthorgonal to the Möbius function

Conjecture (Sarnak conjecture)

Every deterministic bounded complex valued sequence $u(n)$, $n \in \mathbb{N}$ is orthogonal to the Möbius function.

★ Background 2: Sarnak Conjecture

The following sequences are orthogonal to $\mu(n)$:

- For constant sequences.
- For periodic sequences.
- For quasiperiodic sequences such as $f(n) = F(\alpha n \bmod 1)$ for some continuous F.
- For nilsequences (Green+Tao).
- For horocycle flows (Bourgain+Sarnak+Ziegler).
- For the Thue-Morse sequence (Dartyge-Tenenbaum).
- For the Rudin-Shapiro sequence (Tao, Mauduit+Rivat).
- For all **automatic sequences** (Müllner).
- ...

★ Background 2: Sarnak Conjecture

Definition (Von Mangoldt Λ-Function)

$$\Lambda(n) = \begin{cases} \log p & n = p^k \text{ for some prime } p \text{ and some } k \geqslant 1, \\ 0 & \text{else.} \end{cases}$$

★ Background 2: Sarnak Conjecture

Definition (Von Mangoldt Λ-Function)

$$\Lambda(n) = \begin{cases} \log p & n = p^k \text{ for some prime } p \text{ and some } k \geqslant 1, \\ 0 & \text{else.} \end{cases}$$

$$\sum_{n \leqslant N} \Lambda(n) u_n \approx \log N \sum_{p \leqslant N,\, p \in \mathbb{P}} u_p$$

★ Background 2: Sarnak Conjecture

The analysis of

$$\sum_{n\leqslant N} \Lambda(n) G(u(n))$$

which leads to the sum $\sum_{p\leqslant N} G(u(p))$ and to distribution properties of the subsequence $u(p)$, $p \in \mathbb{P}$, is completely similar to the analysis of

$$\sum_{n\leqslant N} \mu(n) G(u(n))$$

which is needed for the Sarnak conjecture.

★ Background 2: Sarnak Conjecture

The analysis of

$$\sum_{n\leqslant N} \Lambda(n)G(u(n))$$

which leads to the sum $\sum_{p\leqslant N} G(u(p))$ and to distribution properties of the subsequence $u(p)$, $p \in \mathbb{P}$, is completely similar to the analysis of

$$\sum_{n\leqslant N} \mu(n)G(u(n))$$

which is needed for the Sarnak conjecture.

Thus, there is a natural relation between the **Gelfond problem on primes** and the **Sarnak conjecture**.

★ Background 2: Sarnak Conjecture

Theorem (Müllner 2017)

For every automatic sequence $(u_n)_{n\geqslant 0}$ *one has*

$$\sum_{n\leq N} \mu(n) u_n = o(N) \qquad (N \to \infty).$$

For every primitive automatic sequence $(u_n)_{n\geqslant 0}$ *one has*

$$\sum_{n\leqslant N} \Lambda(n) u_n = cN + o(N)$$

for some real c.

★ Background 2: Sarnak Conjecture

Theorem (Müllner 2017)

For every automatic sequence $(u_n)_{n\geqslant 0}$ *one has*

$$\sum_{n\le N}\mu(n)u_n = o(N) \qquad (N\to\infty).$$

For every primitive automatic sequence $(u_n)_{n\geqslant 0}$ *one has*

$$\sum_{n\leqslant N}\Lambda(n)u_n = cN + o(N)$$

for some real c.

An automatic sequence u_n is called primitive if the associated dynamical system (X,T) is minimal, that is, $\overline{\{T^n(x) : n\geqslant 0\}} = X$ for all $x\in X$, or if the associated minimal automaton is strongly connected.

★ Background 2: Sarnak Conjecture

Theorem (Adamczewski+D.+Müllner 2020)

Every automatic sequence $(u_n)_{n\geqslant 0}$ *satisfies*

$$\sum_{n\leqslant N}\frac{1}{n}\Lambda(n)u_n = c\log N + o(\log N)$$

for some real c.

★ Background 2: Sarnak Conjecture

Theorem (Adamczewski+D.+Müllner 2020)

Every automatic sequence $(u_n)_{n\geqslant 0}$ *satisfies*

$$\sum_{n\leqslant N}\frac{1}{n}\Lambda(n)u_n = c\log N + o(\log N)$$

for some real c*.*

Theorem (D.+Lemańczyk+Müllner+Rivat 2025)

Suppose that $(u_n^{(j)})_{n\geqslant 0}$ *are primitive* q_j*-automatic sequences for* $1\leqslant j\leqslant d$ *such that* $q_1,\ldots,q_d$ *are pairwise coprime. Then*

$$\sum_{n\leq N}\mu(n)u_n^{(1)}u_n^{(2)}\cdots u_n^{(d)} = o(N) \qquad (N\to\infty).$$

★ Variants and Combined Problems

Theorem (D., Mauduit and Rivat, 2009)

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

$$\#\{\textit{primes } p < N : s_q(p) = k\}$$

$$= \frac{q-1}{\varphi(q-1)} \frac{\pi(N)}{\sqrt{2\pi\sigma_q^2 \log_q N}} \left(\exp\left(-\frac{(k-\mu_q \log_q N)^2}{2\sigma_q^2 \log_q N} \right) + O((\log N)^{-\frac{1}{2}+\varepsilon}$$

where

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

★ Variants and Combined Problems

Theorem (D., Mauduit and Rivat, 2009)

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

$$\#\{primes\ p < N : s_q(p) = k\}$$

$$= \frac{q-1}{\varphi(q-1)} \frac{\pi(N)}{\sqrt{2\pi\sigma_q^2 \log_q N}} \left(\exp\left(-\frac{(k-\mu_q \log_q N)^2}{2\sigma_q^2 \log_q N} \right) + O((\log N)^{-\frac{1}{2}+\varepsilon}$$

where

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

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

★ Variants and Combined Problems

Theorem (D.+Mauduit+Rivat 2019)

The Thue-Morse sequence $t_n = s_2(n) \bmod 2$ *along squares is a* **normal sequence***:*

$$\#\{n < N : t_{n^2} = \ell_0,\ t_{(n+1)^2} = \ell_1, \ldots, t_{(n+d-1)^2} = \ell_{d-1}\} \sim \frac{N}{2^d}$$

for every $d \geqslant 1$ *and for every* $\ell_0, \ldots, \ell_{d-1} \in \{0, 1\}$*.*

★ Variants and Combined Problems

Theorem (D.+Mauduit+Rivat 2019)

The Thue-Morse sequence $t_n = s_2(n) \bmod 2$ *along squares is a* **normal sequence***:*

$$\#\{n < N : t_{n^2} = \ell_0,\ t_{(n+1)^2} = \ell_1, \ldots, t_{(n+d-1)^2} = \ell_{d-1}\} \sim \frac{N}{2^d}$$

for every $d \geqslant 1$ *and for every* $\ell_0, \ldots, \ell_{d-1} \in \{0, 1\}$*.*

Theorem (D.+Mauduit+Rivat+Spiegelhofer 2022)

Let t_n *denote the Thue–Morse sequence. Then we have, as* $N \to \infty$*,*

$$\sum_{n<N} \mu(n) t_{n^2} = o(N).$$

★ Variants and Combined Problems

Theorem (D.+Mauduit+Rivat 2020)

Suppose that q_1 and q_2 are different prime numbers such that $(q_1-1, m_1) = (q_1-1, m_2) = 1$. Then

$$\#\{\textit{primes } p < N : s_{q_1}(p) \equiv \ell_1 \bmod m_1,\ s_{q_2}(p) \equiv \ell_2 \bmod m_2\} \sim \frac{\pi(N)}{m_1 m_2}$$

★ Variants and Combined Problems

Theorem (D.+Mauduit+Rivat 2020)

Suppose that q_1 and q_2 are different prime numbers such that $(q_1-1, m_1) = (q_1-1, m_2) = 1$. Then

$$\#\{\text{primes } p < N : s_{q_1}(p) \equiv \ell_1 \bmod m_1,\ s_{q_2}(p) \equiv \ell_2 \bmod m_2\} \sim \frac{\pi(N)}{m_1 m_2}$$

This result **combines** the Gelfond problem for several bases and that for prime numbers.

★ Variants and Combined Problems

Theorem (D.+Mauduit+Rivat 2020)

Suppose that q_1 and q_2 are different prime numbers such that $(q_1-1, m_1) = (q_1-1, m_2) = 1$. Then

$$\#\{primes\ p < N : s_{q_1}(p) \equiv \ell_1 \bmod m_1,\ s_{q_2}(p) \equiv \ell_2 \bmod m_2\} \sim \frac{\pi(N)}{m_1 m_2}$$

This result **combines** the Gelfond problem for several bases and that for prime numbers.
There is now a generalization to several q's by Müllner, Spiegelhofer, and Shubin (ongoing work).

★ Variants and Combined Problems

Theorem (D.+Rivat)

Suppose that $(q-1,m)=1$ *and that* q *is a sufficiently large prime number. Then*

$$\#\{primes\ p < N : s_q(p^2) \equiv a \bmod m\} = \frac{\pi(x)}{m} + O\left(x^{1-\frac{c'}{\log q}\|(q-1)\gamma\|^2}\right)$$

★ Variants and Combined Problems

Theorem (D.+Rivat)

Suppose that $(q-1, m) = 1$ *and that* q *is a sufficiently large prime number. Then*

$$\#\{primes\ p < N : s_q(p^2) \equiv a \bmod m\} = \frac{\pi(x)}{m} + O\left(x^{1-\frac{c'}{\log q}\|(q-1)\gamma\|^2}\right)$$

This result **combines** the Gelfond problesm for primes and squares. The method does now work, for example, for $q = 2$.

★ Zeckendorf Expansion

Zeckendorf expansion

($F_0 = 0, F_1 = 1, F_{k+1} = F_k + F_{k-1}$ Fibonacci numbers)

$$\boxed{n = \sum_{j\geqslant 2} \varepsilon_{Z,j}(n) F_j}, \quad \varepsilon_{Z,j}(n) \in \{0,1\},\ \varepsilon_{Z,j}(n)\varepsilon_{Z,j+1}(n) = 0.$$

Zeckendorf Sum-of-digits function

$$s_Z(n) = \sum_{j\geqslant 2} \varepsilon_{Z,j}(n)$$

★ Zeckendorf Expansion

Theorem (D.+Müllner+Spiegelhofer 2018)

We have for non-integers α

$$\sum_{n<N} \mu(n) e(\alpha s_Z(n)) = o(N).$$

★ Zeckendorf Expansion

Theorem (D.+Müllner+Spiegelhofer 2018)

We have for non-integers α

$$\sum_{n<N} \mu(n) e(\alpha s_Z(n)) = o(N).$$

Theorem (D.+Müllner+Spiegelhofer 2025)

We have for non-integers α *(and some* $c > 0$*)*

$$\sum_{n\leqslant N} \Lambda(n) \mathrm{e}\big(\alpha\, s_Z(n)\big) \ll (\log N)^5 N^{1-c\|\alpha\|^2}.$$

★ Zeckendorf Expansion

Theorem (D.+Müllner+Spiegelhofer 2018)

We have for non-integers α

$$\sum_{n<N} \mu(n) e(\alpha s_Z(n)) = o(N).$$

Theorem (D.+Müllner+Spiegelhofer 2025)

We have for non-integers α *(and some* $c > 0$*)*

$$\sum_{n\leqslant N} \Lambda(n)\mathrm{e}\big(\alpha\, s_Z(n)\big) \ll (\log N)^5 N^{1-c\|\alpha\|^2}.$$

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

★ Zeckendorf Expansion

Theorem (D.+Müllner+Spiegelhofer 2025)

For each $\varepsilon > 0$*, we have*

$$\#\{p \leqslant x : s_Z(p) = k\} = \frac{\pi(x)}{\sqrt{2\pi\sigma^2 \log_\gamma x}} \left(e^{-\frac{(k-\mu \log_\gamma x)^2}{2\sigma^2 \log_\gamma x}} + O\left((\log x)^{-\frac{1}{2}+\varepsilon}\right) \right)$$

uniformly for all integers $k \geqslant 0$*, where*

$$\mu = \frac{1}{\gamma^2+1} \quad \text{and} \quad \sigma^2 = \frac{\gamma^3}{(\gamma^2+1)^3},$$

$\pi(x)$ *denotes the number of primes* $\leqslant x$*, and* $\log_\gamma x = \log x / \log \gamma$.

★ Zeckendorf Expansion

Theorem (D.+Müllner+Spiegelhofer 2025)

For each $\varepsilon > 0$, we have

$$\#\{p \leqslant x : s_Z(p) = k\} = \frac{\pi(x)}{\sqrt{2\pi\sigma^2 \log_\gamma x}} \left(e^{-\frac{(k-\mu \log_\gamma x)^2}{2\sigma^2 \log_\gamma x}} + O\left((\log x)^{-\frac{1}{2}+\varepsilon}\right) \right)$$

uniformly for all integers $k \geqslant 0$, where

$$\mu = \frac{1}{\gamma^2+1} \quad \textit{and} \quad \sigma^2 = \frac{\gamma^3}{(\gamma^2+1)^3},$$

$\pi(x)$ *denotes the number of primes* $\leqslant x$, *and* $\log_\gamma x = \log x / \log \gamma$.

In particular, for every sufficiently large integer k there exists a prime number p with

$$s_Z(p) = k.$$

★ Proof Methods

- Discrete Fourier techniques
- Exponential sum estimates (van der Corput inequality etc.)
- Digit detection techniques
- Diophantine approximation (Baker's theorem, *p*-adic subspace theorem)
- Discrepancy estimates
- Daboussi-Kátai-criterion
- Vaughan's method
- Level of distribution and Gowers norm estimates

★ Discrete Fourier techniques

$$F_\lambda(h) = \frac{1}{q^\lambda} \sum_{0 \leqslant u < q^\lambda} e\left(\alpha s_q(u) - hq^{-\lambda}\right)$$

Lemma

We have uniformly for all h:

$$|F_\lambda(h)| \leqslant C_1 q^{-c_1 \|(q-1)\alpha\|^2 \lambda}$$

and

$$\sum_{0 \leqslant h < q^\lambda} |F_\lambda(h)| \leqslant C_2 q^{\eta_q \lambda}$$

for some positive constants C_1, c_1, C_2 *and some* $0 < \eta_q < \frac{1}{2}$.

★ Daboussi-Kátai-criterion

Lemma

Let $f(n)$ be a bounded sequence such that

$$\sum_{n\leqslant N} f(pn)\overline{f(qn)} = o(N)$$

for all distinct prime numbers p, q. Then

$$\sum_{n\leqslant N} \mu(n)f(n) = o(N).$$

There is a similar criterion by Bourgain-Sarnak-Ziegler.

★ Vaughan's method

Lemma

Let $f:\mathbb{N}\to\mathbb{C}$ *such that* $|f(n)|\leqslant 1$ *for all* $n\geq 1$. *For all* $N,U,V\geqslant 2$ *such that* $UV\leq N$ *we have*

$$\sum_{n\leq N}\Lambda(n)\,f(n)\ll U+(\log N)\sum_{t\leq UV}\max_{w}\left|\sum_{w\leq r\leq N/t}f(rt)\right|$$

$$+\sqrt{N}(\log N)^3\max_{\substack{U\leq M\leq N/V\\ V\leq q_2\leq N/M}}\left(\sum_{V<q_1\leq N/M}\left|\sum_{\substack{M<m\leq 2M\\ m\leq\min(N/q_1,N/q_2)}}f(mq_1)\overline{f(mq_2)}\right|\right)^{1/2},$$

with an absolute implied constant.

★ Level of distribution and Gowers norm

Level of distribution of the Thue-Morse sequence: For all $\varepsilon > 0$ we have

$$\sum_{1 \le d \le x^{1-\varepsilon}} \max_{\substack{y,z \ge 0 \\ z-y \le x}} \max_{0 \le a < d} \left| \sum_{\substack{y \le n < z \\ n \equiv a \bmod d}} (-1)^{s_2(n)} \right| = \mathcal{O}(x^{1-\eta})$$

for some $\eta > 0$ depending on ε; the level of distribution equals **1.**

★ Level of distribution and Gowers norm

Level of distribution of the Thue-Morse sequence: For all $\varepsilon > 0$ we have

$$\sum_{1\le d\le x^{1-\varepsilon}} \max_{\substack{y,z\ge 0\\ z-y\le x}} \max_{0\le a<d} \left| \sum_{\substack{y\le n<z\\ n\equiv a \bmod d}} (-1)^{s_2(n)} \right| = \mathcal{O}(x^{1-\eta})$$

for some $\eta > 0$ depending on ε; the level of distribution equals **1.**

Gowers norm of the Thue-Morse sequence:

$$\sum_{0\le n,r_1,\ldots,r_m<2^\rho} \mathrm{e}\left(\alpha \sum_{\varepsilon\in\{0,1\}^m} s_\rho(n+\varepsilon\cdot r)\right) \ll 2^{(m+1)\rho\cdot(1-c\|\alpha\|^2)}$$

for all $\rho \ge 0$, where $\varepsilon \cdot r = \sum_{1\le i\le m} \varepsilon_i r_i$ and s_ρ is the *truncated sum-of-digits function* in base 2.

★ Open Problems

- Gelfond-Problem for general integer polynomials and ALL $q \geqslant 2$.
- Sarnak conjecture for general morphic sequences (fixed points of general finite substitutions)
- Precise behavior of $S(N) = \sum_{n \leqslant N} e(\alpha s_2(n))$
- Distribution of $s_Z(n^2) \bmod m$
- More general numeration systems
- ...

Thank you very much for your attention!