Thue–Morse along the sequence of cubes

Lukas Spiegelhofer

Sep 19, 2023
ÖMG Tagung 2023, Universität Graz

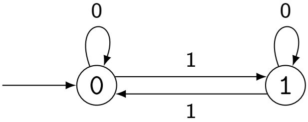
Section 1

Thue–Morse [tʉː mɔːrs]

The Thue–Morse sequence $\mathbf{t}$ is the fixed point of the substitution

$$0 \mapsto 01, \quad 1 \mapsto 10$$

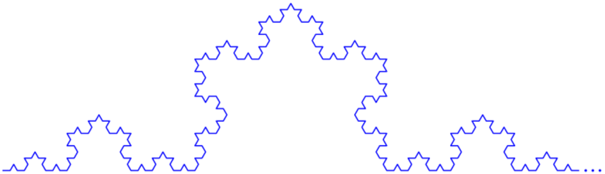
that starts with 0.
It is given by the binary sum-of-digits function s, reduced modulo 2.



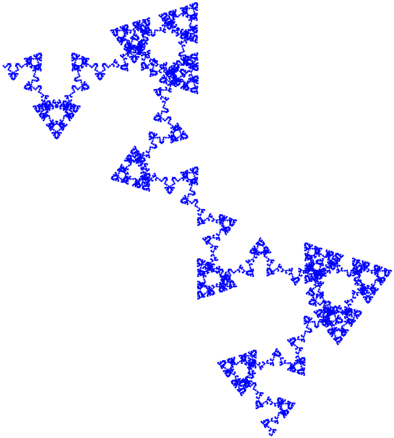
$$\mathbf{t} = 01101001100101101001011001101001 \cdots$$

Thue–Morse $\rightleftharpoons$ Koch

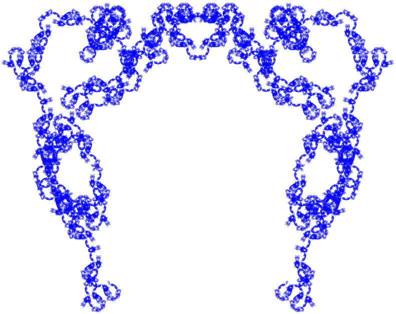
The sequence $n \mapsto (-1)^{s(n)}\, \mathrm{e}(-n/3)$ describes the orientation of the nth segment in the "unscaled Koch (snowflake) curve" (where $\mathrm{e}(x) = e^{2\pi i x}$):



The sum of digits along arithmetic progressions

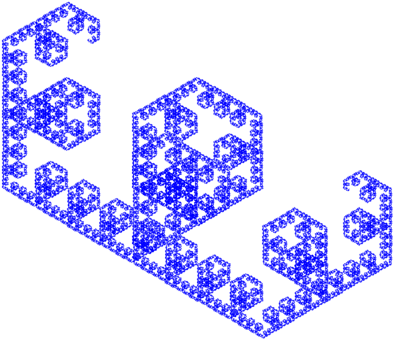


$\mathrm{e}\left(\frac{1}{2}s(3n) - n/5\right)$

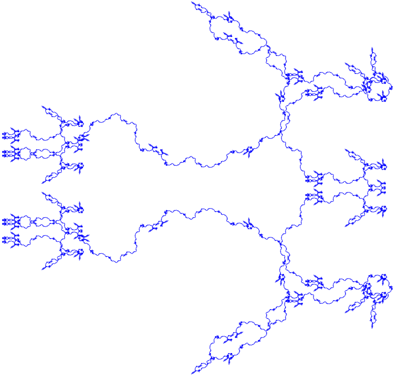


$\mathrm{e}\left(\frac{2}{5}s(3n) - n/5\right)$

The sum of digits along arithmetic progressions

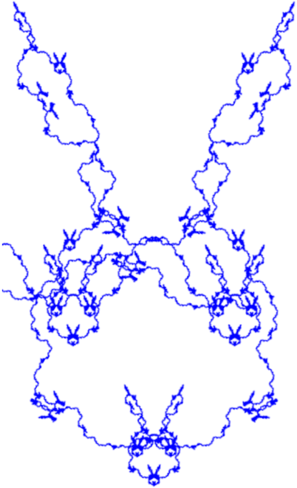


$\mathrm{e}(s(3n)/3)(-1)^n$



$\mathrm{e}(s(7n)/3)(-1)^n$

The sum of digits along arithmetic progressions

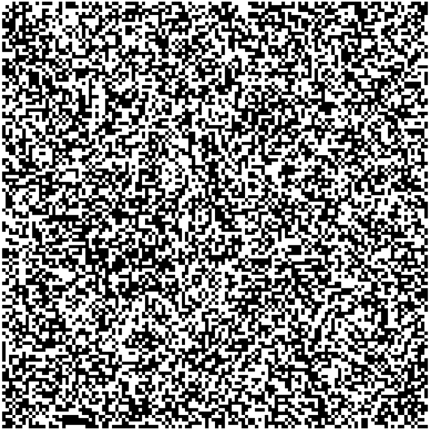


$\mathrm{e}(s(7n)/3)(-1)^n$, closeup

Every finite sequence $\omega \in \{0,1\}^L$ appears as an arithmetic subsequence of $\mathbf{t}$: the Thue–Morse word has full *arithmetical complexity* (Avgustinovich–Fon-Der-Flaass–Frid 2003, Müllner–Spiegelhofer 2017, Konieczny–Müllner 2023+).

Every finite sequence $\omega \in \{0,1\}^L$ appears as an arithmetic subsequence of $\mathbf{t}$: the Thue–Morse word has full *arithmetical complexity* (Avgustinovich–Fon-Der-Flaass–Frid 2003, Müllner–Spiegelhofer 2017, Konieczny–Müllner 2023+).

Short arithmetic subsequences of $\mathbf{t}$ even seems to behave randomly.

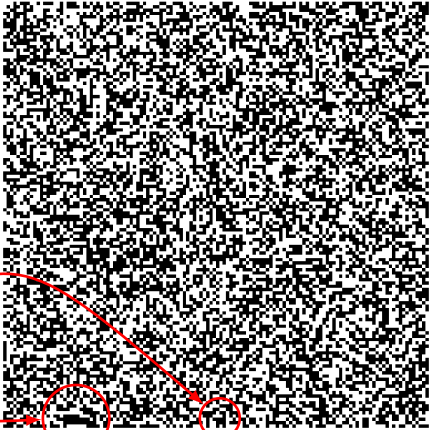


$N = 128 \times 128$ terms, common difference $N^R = 3^{21}$

Every finite sequence $\omega \in \{0,1\}^L$ appears as an arithmetic subsequence of $\mathbf{t}$: the Thue–Morse word has full *arithmetical complexity* (Avgustinovich–Fon-Der-Flaass–Frid 2003, Müllner–Spiegelhofer 2017, Konieczny–Müllner 2023+).

Short arithmetic subsequences of $\mathbf{t}$ even seems to behave randomly.

every pattern occurs



$N = 128 \times 128$ terms, common difference $N^R = 3^{21}$

Informal question

Let $A \gg N^R$, and assume that A contains many blocks of 1*s in binary. Is*

$$P : \{0, \ldots, N\} \rightarrow \{0, 1\}, \quad n \mapsto \mathbf{t}(nA + B)$$

a good pseudorandom number generator?

Gelfond's third problem

Let $S = s_q$ be the sum-of-digits function in base $q \geq 2$.

> *Finalement, signalons comme problème à résoudre l'estimation du nombre des valeurs du polynôme $P(t)$ ne prenant que des valeurs entières sur l'ensemble [...] des entiers rationels, pour lesquelles on a $S[P(n)] \equiv \ell \bmod m$.*

A. O. Gelfond, 1967/1968

Gelfond's third problem

Let $S = s_q$ be the sum-of-digits function in base $q \geq 2$.

> *Finalement, signalons comme problème à résoudre l'estimation du nombre des valeurs du polynôme $P(t)$ ne prenant que des valeurs entières sur l'ensemble [...] des entiers rationels, pour lesquelles on a $S[P(n)] \equiv \ell \bmod m$.*
>
> A. O. Gelfond, 1967/1968

That is, if P is a polynomial such that $P(\mathbb{N}) \subseteq \mathbb{N}$, we are interested in

$$A(q, P, m, \ell, x) := \#\{n < x : s_q(P(n)) \equiv \ell \bmod m\}.$$

Partial results

0110100110010110100101100110100110010110011010010 1

- Lower bounds for the numbers $A(q, P, m, \ell, x)$ are known (Dartyge–Tenenbaum 2006; Stoll 2012);
- For "sufficiently large bases" q coprime to the leading coefficient of P, and $\gcd(q-1, m) = 1$, the equivalence $A(q, P, m, \ell, x) \sim x/m$ has been proved (Drmota–Mauduit–Rivat 2011);
- The case $P(x) = x^2$ has been answered by Mauduit and Rivat (Acta Math., 2009).

Partial results

01101001100101101001011001101001100101100110100101

- Lower bounds for the numbers $A(q, P, m, \ell, x)$ are known (Dartyge–Tenenbaum 2006; Stoll 2012);
- For "sufficiently large bases" q coprime to the leading coefficient of P, and $\gcd(q-1, m) = 1$, the equivalence $A(q, P, m, \ell, x) \sim x/m$ has been proved (Drmota–Mauduit–Rivat 2011);
- The case $P(x) = x^2$ has been answered by Mauduit and Rivat (Acta Math., 2009).

Partial results

01 1 0 1 1 0 1

- Lower bounds for the numbers $A(q, P, m, \ell, x)$ are known (Dartyge–Tenenbaum 2006; Stoll 2012);
- For "sufficiently large bases" q coprime to the leading coefficient of P, and $\gcd(q-1, m) = 1$, the equivalence $A(q, P, m, \ell, x) \sim x/m$ has been proved (Drmota–Mauduit–Rivat 2011);
- The case $P(x) = x^2$ has been answered by Mauduit and Rivat (Acta Math., 2009).

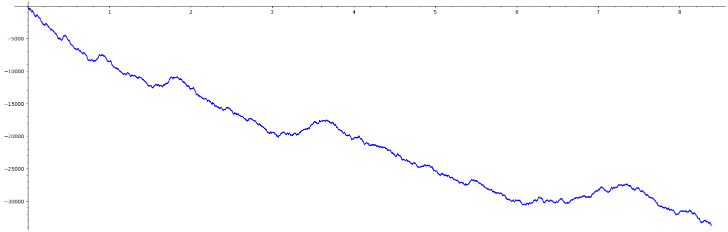
Generalizations

The Thue–Morse sequence along n^2 is normal (Drmota–Mauduit–Rivat): each finite sequence over $\{0, 1\}$ of length L appears with frequency 2^{-L} along $\mathbf{t}(n^2)$.

Generalizations

The Thue–Morse sequence along n^2 is normal (Drmota–Mauduit–Rivat): each finite sequence over $\{0,1\}$ of length L appears with frequency 2^{-L} along $\mathbf{t}(n^2)$.

Partial sums of $\mathbf{t}\big(n^2\big)$ for $x < 2^{23}$:



A *drift* appears to be present. How is this related to the fact that n^2 avoids $2 + 3\mathbb{Z}$?

Conjecture ( )

There exist real numbers c and η, and a 1-periodic, continuous, nowhere differentiable function Φ, such that

$$\sum_{n<x} \mathbf{t}\big(n^2\big) \sim c x^{\eta} \Phi\big(\log x / \log 2\big).$$

The main result

Theorem (S. 2023+)

There exist real numbers $c > 0$ and C such that for all $x \geq 1$,

$$\left| \#\left\{ n < x : \mathbf{t}\left(n^3\right) = 0 \right\} - \frac{x}{2} \right| \leq C x^{1-c}. \tag{1}$$

Section 2

Sketch of the proof

- We are interested in the sum

$$S_0 := \sum_{n<2^\nu} \mathrm{e}\left(\tfrac{1}{2}s(n^3)\right).$$

- We are interested in the sum

$$S_0 := \sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s(n^3)\Bigr).$$

- After an application of *van der Corput's inequality* it remains to handle the correlation

$$\sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s_2\bigl((n+r)^3\bigr) - \tfrac{1}{2} s_2\bigl(n^3\bigr)\Bigr).$$

- We are interested in the sum

$$S_0 := \sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s(n^3)\Bigr).$$

- After an application of *van der Corput's inequality* it remains to handle the correlation

$$\sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s_2\bigl((n+r)^3\bigr) - \tfrac{1}{2} s_2\bigl(n^3\bigr)\Bigr).$$

- But $(n+r)^3$ and n^3 *usually* have the same digits with indices above

$$\lambda := \nu(2+\varepsilon),$$

if r is *small* compared to 2^ν. These digits can therefore be discarded.

- We are interested in the sum

$$S_0 := \sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s(n^3)\Bigr).$$

- After an application of *van der Corput's inequality* it remains to handle the correlation

$$\sum_{n<2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s_2\bigl((n+r)^3\bigr) - \tfrac{1}{2} s_2\bigl(n^3\bigr)\Bigr).$$

- But $(n+r)^3$ and n^3 *usually* have the same digits with indices above

$$\lambda := \nu(2+\varepsilon),$$

if r is *small* compared to 2^ν. These digits can therefore be discarded.

- This is standard.

S. 2020

In the paper (Compos. Math. 2020) we apply van der Corput's inequality repeatedly in order to eliminate blocks of digits, piece by piece.

S. 2020

In the paper (Compos. Math. 2020) we apply van der Corput's inequality repeatedly in order to eliminate blocks of digits, piece by piece.

In this way, a statement on very sparse arithmetic subsequences of $\mathbf{t}$ could be derived. These progressions have length $\asymp N$, while their common difference is $\asymp N^R$, where $R > 0$ is arbitrary!

S. 2020

In the paper (Compos. Math. 2020) we apply van der Corput's inequality repeatedly in order to eliminate blocks of digits, piece by piece.

In this way, a statement on very sparse arithmetic subsequences of **t** could be derived. These progressions have length $\asymp N$, while their common difference is $\asymp N^R$, where $R > 0$ is arbitrary!

But: iterated van der Corput could so far not be used for removing sufficiently many digits of polynomial values, if $\deg P > 1$.

$$s_2\bigl(n^3\bigr) - s_2\bigl((n+r)^3\bigr) - s_2\bigl((n+s)^3\bigr) + s_2\bigl((n+r+s)^3\bigr)$$

A trivial decomposition

We write

$$n = 2^\rho n_1 + n_0,$$

where $3\rho \ge \lambda$ and $n_0 < 2^\rho$. The variable n_0 is treated as a parameter. Expanding $n^3 \bmod 2^\lambda$, we see that the cubic term in n_1 disappears.

A trivial decomposition

We write

$$n = 2^\rho n_1 + n_0,$$

where $3\rho \geq \lambda$ and $n_0 < 2^\rho$. The variable n_0 is treated as a parameter. Expanding $n^3 \bmod 2^\lambda$, we see that the cubic term in n_1 disappears.

On the *critical interval* $[2\rho, \lambda)$ of length $\kappa := \lambda - 2\rho$, the term n_1^2 is still relevant.

A trivial decomposition

We write

$$n = 2^\rho n_1 + n_0,$$

where $3\rho \geq \lambda$ and $n_0 < 2^\rho$. The variable n_0 is treated as a parameter. Expanding $n^3 \bmod 2^\lambda$, we see that the cubic term in n_1 disappears.

On the *critical interval* $[2\rho, \lambda)$ of length $\kappa := \lambda - 2\rho$, the term n_1^2 is still relevant.

? After removing this complication, a linear problem remains, which can be handled by an extension of the method in [S. 2020] ...

A trivial decomposition

We write

$$n = 2^\rho n_1 + n_0,$$

where $3\rho \geq \lambda$ and $n_0 < 2^\rho$. The variable n_0 is treated as a parameter. Expanding $n^3 \bmod 2^\lambda$, we see that the cubic term in n_1 disappears.

On the *critical interval* $[2\rho, \lambda)$ of length $\kappa := \lambda - 2\rho$, the term n_1^2 is still relevant.

? After removing this complication, a linear problem remains, which can be handled by an extension of the method in [S. 2020] . . .

! In the actual proof, the elimination of the digits in the critical interval $[2\rho, \lambda)$ comes first.

The critical interval of digits

For a subset $J \subseteq \mathbb{N}$, let s^J denote the *restricted binary sum-of-digits function*: only digits with indices in J are counted. We write

$$S_0 = \sum_{0 \le j < 2^\kappa} (-1)^{s_2(j)} \sum_{n < 2^\nu} \mathrm{e}\Bigl(\tfrac{1}{2} s^{\mathbb{N} \setminus [2\rho, \lambda)}\bigl(n^3\bigr)\Bigr) \left[\!\left[\frac{n^3}{2^\lambda} \in \left[\frac{j}{2^\kappa}, \frac{j+1}{2^\kappa}\right) + \mathbb{Z} \right]\!\right].$$

The critical interval of digits

For a subset $J \subseteq \mathbb{N}$, let s^J denote the *restricted binary sum-of-digits function*: only digits with indices in J are counted. We write

$$S_0 = \sum_{0 \le j < 2^\kappa} (-1)^{s_2(j)} \sum_{n < 2^\nu} \mathrm{e}\Big(\tfrac{1}{2} s^{\mathbb{N} \setminus [2\rho, \lambda)}\big(n^3\big)\Big) \left[\!\!\left[\frac{n^3}{2^\lambda} \in \left[\frac{j}{2^\kappa}, \frac{j+1}{2^\kappa}\right) + \mathbb{Z} \right]\!\!\right].$$

(1) An additional sum of length 2^κ is introduced;

The critical interval of digits

For a subset $J \subseteq \mathbb{N}$, let s^J denote the *restricted binary sum-of-digits function*: only digits with indices in J are counted. We write

$$S_0 = \sum_{0 \le j < 2^\kappa} (-1)^{s_2(j)} \sum_{n < 2^\nu} \mathrm{e}\Big(\tfrac{1}{2} s^{\mathbb{N} \setminus [2\rho, \lambda)}\big(n^3\big)\Big) \left[\!\left[\frac{n^3}{2^\lambda} \in \left[\frac{j}{2^\kappa}, \frac{j+1}{2^\kappa} \right) + \mathbb{Z} \right]\!\right].$$

(1) An additional sum of length 2^κ is introduced;

(2) The "prepared" set $\mathbb{N} \setminus [2\rho, \lambda)$ will lead to a linear sum-of-digits problem after cutting away the digits with indices $\geq \lambda$ (as above);

The critical interval of digits

For a subset $J \subseteq \mathbb{N}$, let s^J denote the *restricted binary sum-of-digits function*: only digits with indices in J are counted. We write

$$S_0 = \sum_{0 \le j < 2^\kappa} (-1)^{s_2(j)} \sum_{n < 2^\nu} \mathrm{e}\Big(\tfrac{1}{2} s^{\mathbb{N} \setminus [2\rho, \lambda)}(n^3)\Big) \left[\!\left[\frac{n^3}{2^\lambda} \in \left[\frac{j}{2^\kappa}, \frac{j+1}{2^\kappa}\right) + \mathbb{Z} \right]\!\right].$$

(1) An additional sum of length 2^κ is introduced;

(2) The "prepared" set $\mathbb{N} \setminus [2\rho, \lambda)$ will lead to a linear sum-of-digits problem after cutting away the digits with indices $\ge \lambda$ (as above);

(3) The rightmost factor is approximated by a trigonometric polynomial, evaluated at $n^3/2^\lambda$.

Even sketchier idea of the proof

- Writing $n = 2^\rho n_1 + n_0$ as before, the term n_1^3 does not appear in the argument of the trigonometric polynomial.

Even sketchier idea of the proof

- Writing $n = 2^\rho n_1 + n_0$ as before, the term n_1^3 does not appear in the argument of the trigonometric polynomial.
- Applying van der Corput's inequality, the degree of the argument (in n_1) is reduced by 1 once more.

Even sketchier idea of the proof

- Writing $n = 2^\rho n_1 + n_0$ as before, the term n_1^3 does not appear in the argument of the trigonometric polynomial.
- Applying van der Corput's inequality, the degree of the argument (in n_1) is reduced by 1 once more.
- The linear trigonometric polynomial in n_1 is *decoupled* from the sum over n, using suitable arithmetic subsequences and summation by parts.

Even sketchier idea of the proof

- Writing $n = 2^\rho n_1 + n_0$ as before, the term n_1^3 does not appear in the argument of the trigonometric polynomial.
- Applying van der Corput's inequality, the degree of the argument (in n_1) is reduced by 1 once more.
- The linear trigonometric polynomial in n_1 is *decoupled* from the sum over n, using suitable arithmetic subsequences and summation by parts.
- The sum over h together with the decoupled exponential term yields a geometric sum

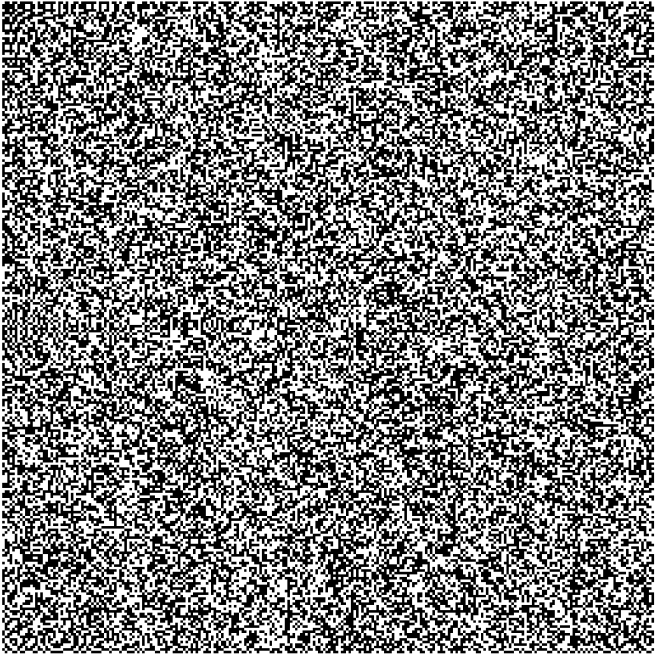
$$\sum_{0 \le h < H} \mathrm{e}(hx) \ll \min\Bigl(H, \|x\|^{-1}\Bigr),$$

where $\|x\|$ is the distance of x to the nearest integer.

This is only logarithmic in mean (over x)!

Essence of the proof

Summarizing, the additional sum introduced for digit detection in the critical interval only contributes a logarithm. A linear digital problem remains, for which there are methods available.



THANK YOU!

M. Drmota, C. Mauduit, and J. Rivat, *Normality along squares*, J. Eur. Math. Soc, 21 (2019), pp. 507–548.

A. O. Gel′fond, *Sur les nombres qui ont des propriétés additives et multiplicatives données*, Acta Arith., 13 (1967/1968), pp. 259–265.

C. Mauduit and J. Rivat, *La somme des chiffres des carrés*, Acta Math., 203 (2009), pp. 107–148.

L. Spiegelhofer, *The level of distribution of the Thue–Morse sequence*, Compos. Math., 156 (2020), pp. 2560–2587.

——, *Thue–Morse along the sequence of cubes*, 2023.
Preprint, http://arxiv.org/abs/2308.09498.

Supported by the FWF–ANR joint project ArithRand, and P36137 (FWF).

van der Corput's inequality

Lemma

Let I be a finite interval containing N integers and let a_n be a complex number for $n \in I$. For all integers $K \geq 1$ and $R \geq 1$ we have

$$\left|\sum_{n\in I} a_n\right|^2 \leq \frac{N+K(R-1)}{R} \sum_{|r|<R} \left(1-\frac{|r|}{R}\right) \sum_{\substack{n\in I\\ n+Kr\in I}} a_{n+Kr}\overline{a_n}.$$

van der Corput's inequality

Lemma

Let I be a finite interval containing N integers and let a_n be a complex number for $n \in I$. For all integers $K \geq 1$ and $R \geq 1$ we have

$$\left|\sum_{n\in I} a_n\right|^2 \leq \frac{N+K(R-1)}{R} \sum_{|r|<R} \left(1-\frac{|r|}{R}\right) \sum_{\substack{n\in I\\ n+Kr\in I}} a_{n+Kr}\overline{a_n}.$$

Instead of the original sum, we now have to estimate certain correlations (where KR will be small compared to N).