The level of distribution of the Thue–Morse sequence

Lukas Spiegelhofer

Vienna University of Technology

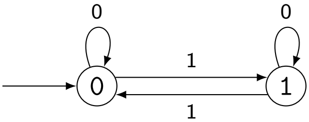
April 19, 2018
Kooperationsworkshop SFB 55, Linz

Section 1

Arithmetic subsequences of the Thue–Morse sequence

The Thue–Morse sequence $\mathbf{t}$

(1) is an automatic sequence:



(2) is the base-2 sum-of-digits function s, modulo 2.

(3) is the fixed point, starting with 0, of the substitution

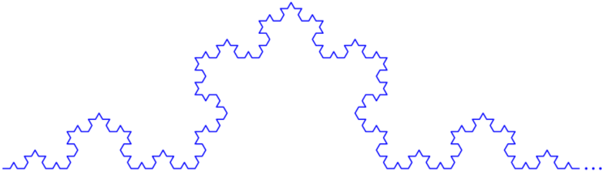
$$0 \mapsto 01$$
$$1 \mapsto 10.$$

(4) can be constructed starting with $\mathbf{t}^{(0)} = 0$ and setting $\mathbf{t}^{(k+1)} = \mathbf{t}^{(k)}\overline{\mathbf{t}^{(k)}}$ (Boolean complement).

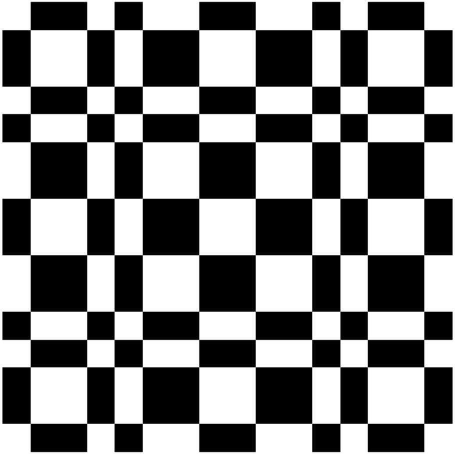
The Thue–Morse sequence, continued

A less well-known characterization uses the Koch snowflake curve.

(5) The sequence $n \mapsto (-1)^{s(n)}\,\mathrm{e}(-n/3)$ describes the orientation of the nth segment in the unscaled snowflake curve:



Thue–Morse, 16×16.



In particular, there are $\ll k$ many subwords of length k (as for any automatic sequence).

Thue–Morse along arithmetic progressions

011010011001011010010110011010011001011001101001011010

Thue–Morse along arithmetic progressions

011010011001011010010110011010011001011001101001011010

Thue–Morse along arithmetic progressions

0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0

Thue–Morse along arithmetic progressions

0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0

More 0s than 1s (Newman's phenomenon), but not in the main term.

Theorem (Gelfond 1968)

Let $d \geq 1$ and a be integers. There is an absolute $\lambda < 1$ such that

$$\left|\{1 \leq n \leq x : \mathbf{t}(n) = 0, n \equiv a \bmod d\}\right| = \frac{x}{2d} + \mathcal{O}\left(x^{\lambda}\right).$$

Thue–Morse along arithmetic progressions

0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0

More 0s than 1s (Newman's phenomenon), but not in the main term.

Theorem (Gelfond 1968)

Let $d \geq 1$ and a be integers. There is an absolute $\lambda < 1$ such that

$$\left|\{1 \leq n \leq x : \mathbf{t}(n) = 0, n \equiv a \bmod d\}\right| = \frac{x}{2d} + \mathcal{O}\left(x^{\lambda}\right).$$

That is, for all $d \geq 1$ and $a < d$ we have

$$\sum_{1 \leq m \leq M} (-1)^{s(md+a)} \ll M^{\lambda}.$$

The error term for sums over APs

- We cannot get a uniform constant C in the estimate

$$\sum_{1 \le m \le M} (-1)^{s(md+a)} \le CM^{\lambda}.$$

- This follows since there are arbitrarily long arithmetic progressions on which $\mathbf{t}$ is constant.

The error term for sums over APs

- We cannot get a uniform constant C in the estimate

$$\sum_{1\le m\le M} (-1)^{s(md+a)} \le CM^{\lambda}.$$

- This follows since there are arbitrarily long arithmetic progressions on which $\mathbf{t}$ is constant.
- Therefore we look at a certain average over d.

Theorem (Fouvry–Mauduit 1996)

$$\sum_{1\le d\le D} \max_{1\le y\le x} \max_{0\le a<d} \left| \sum_{\substack{0\le n<y\\ n\equiv a \bmod d}} (-1)^{s(n)} \right| \le Cx^{1-\eta}$$

for some $\eta > 0$ *and* C *and* $D = x^{0.5924}$.

The error term for sums over APs

- We cannot get a uniform constant C in the estimate

$$\sum_{1 \le m \le M} (-1)^{s(md+a)} \le CM^{\lambda}.$$

- This follows since there are arbitrarily long arithmetic progressions on which $\mathbf{t}$ is constant.
- Therefore we look at a certain average over d.

Theorem (Fouvry–Mauduit 1996)

$$\sum_{1 \le d \le D} \max_{1 \le y \le x} \max_{0 \le a < d} \left| \sum_{\substack{0 \le n < y \\ n \equiv a \bmod d}} (-1)^{s(n)} \right| \le C x^{1-\eta}$$

for some $\eta > 0$ *and* C *and* $D = x^{0.5924}$.

observed − expected

The error term for sums over APs

- We cannot get a uniform constant C in the estimate

$$\sum_{1\le m\le M} (-1)^{s(md+a)} \le CM^\lambda.$$

- This follows since there are arbitrarily long arithmetic progressions on which **t** is constant.
- Therefore we look at a certain average over d.

Theorem (Fouvry–Mauduit 1996)

selects the "worst" AP contained in $[0, x)$

$$\sum_{1\le d\le D} \max_{1\le y\le x} \max_{0\le a<d} \left| \sum_{\substack{0\le n<y\\ n\equiv a \bmod d}} (-1)^{s(n)} \right| \le Cx^{1-\eta}$$

observed − expected

for some $\eta > 0$ *and* C *and* $D = x^{0.5924}$.

The level of distribution of the Thue–Morse sequence

Theorem (Spiegelhofer 2018+)

The Thue–Morse sequence has level of distribution 1. *More precisely, let* $0 < \varepsilon < 1$. *There exist* $\eta > 0$ *and* C *such that*

$$\sum_{1 \le d \le D} \max_{0 \le y \le x} \max_{0 \le a < d} \left| \sum_{\substack{0 \le n < y \\ n \equiv a \bmod d}} (-1)^{s(n)} \right| \le C x^{1-\eta}$$

for $D = x^{1-\varepsilon}$.

The level of distribution of the Thue–Morse sequence

Theorem (Spiegelhofer 2018+)

The Thue–Morse sequence has level of distribution 1*. More precisely, let* $0 < \varepsilon < 1$*. There exist* $\eta > 0$ *and* C *such that*

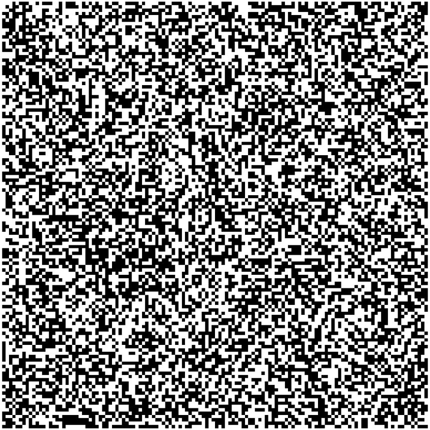
$$\sum_{1 \le d \le D} \max_{0 \le y \le x} \max_{0 \le a < d} \left| \sum_{\substack{0 \le n < y \\ n \equiv a \bmod d}} (-1)^{s(n)} \right| \le C x^{1-\eta}$$

for $D = x^{1-\varepsilon}$*.*

- This is a statement on *sparse* arithmetic progressions: the Thue–Morse sequence usually shows cancellation along N-term arithmetic progressions having common difference $\sim N^R$, where $R > 0$ is arbitrary.

Sparse arithmetic subsequences of TM

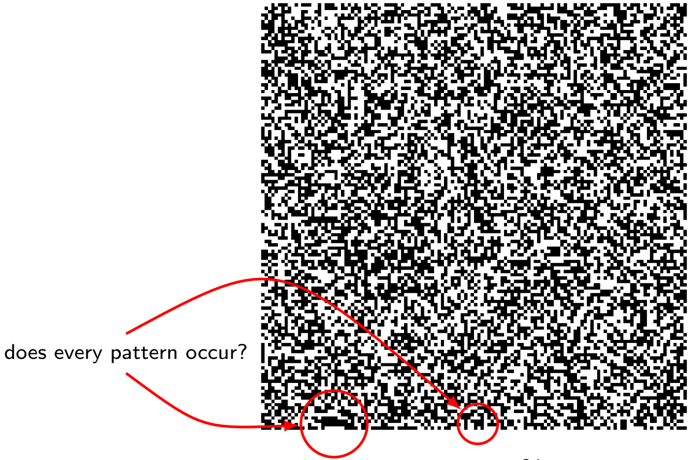
TM along short arithmetic subsequences (with respect to the common difference) even seems to behave randomly.



128×128 terms, common difference $= 3^{21} \rightsquigarrow$ chaos!

Sparse arithmetic subsequences of TM

TM along short arithmetic subsequences (with respect to the common difference) even seems to behave randomly.



128×128 terms, common difference $= 3^{21} \rightsquigarrow$ chaos!

Does every pattern occur?

Theorem (Müllner, Spiegelhofer 2017)

Every finite sequence over $\{0, 1\}$ *appears as an arithmetic subsequence of the Thue–Morse sequence.*

Does every pattern occur?

Theorem (Müllner, Spiegelhofer 2017)

Every finite sequence over $\{0, 1\}$ *appears as an arithmetic subsequence of the Thue–Morse sequence.*

This is a corollary of the following statement: for most $d \sim N^{2-\varepsilon}$, the number of times that $\omega \in \{0, 1\}^L$ appears as a subword of $(\mathbf{t}(nd + a))_{n<N}$ is close to the expected value.

Does every pattern occur?

Theorem (Müllner, Spiegelhofer 2017)

Every finite sequence over $\{0,1\}$ appears as an arithmetic subsequence of the Thue–Morse sequence.

This is a corollary of the following statement: for most $d \sim N^{2-\varepsilon}$, the number of times that $\omega \in \{0,1\}^L$ appears as a subword of $(\mathbf{t}(nd+a))_{n<N}$ is close to the expected value.

Handle occurrences of subwords in $n \mapsto \mathbf{t}(nd+a)$ for $d \sim N^R$.

Does every pattern occur?

Theorem (Müllner, Spiegelhofer 2017)

Every finite sequence over $\{0,1\}$ *appears as an arithmetic subsequence of the Thue–Morse sequence.*

This is a corollary of the following statement: for most $d \sim N^{2-\varepsilon}$, the number of times that $\omega \in \{0,1\}^L$ appears as a subword of $(\mathbf{t}(nd+a))_{n<N}$ is close to the expected value.

Handle occurrences of subwords in $n \mapsto \mathbf{t}(nd+a)$ for $d \sim N^R$.

?? Pseudorandom properties: is $\mathbf{t}(nd+a)$ a good PRNG?

Section 2

Sparse infinite subsequences of Thue–Morse

The Thue–Morse sequence along sparse subsequences

The sum of digits of $\lfloor n^c \rfloor$ is an approximation to the problem "the sum of digits of $p(n)$".

The Thue–Morse sequence along sparse subsequences

The sum of digits of $\lfloor n^c \rfloor$ is an approximation to the problem "the sum of digits of $p(n)$".

Theorem (Corollary of Mauduit–Rivat 2005)

Let $1 < c < 1.4$*. There exists a* $\eta > 0$ *such that*

$$\sum_{1 \le n \le x} (-1)^{s(\lfloor n^c \rfloor)} \ll x^{1-\eta}.$$

In particular, for $1 < c < 1.4$, the sequence $n \mapsto \mathbf{t}(\lfloor n^c \rfloor)$ is *simply normal*.

The Thue–Morse sequence along sparse subsequences

Theorem (Mauduit–Rivat 2009, Acta Math.)

The Thue–Morse sequence along the sequence of squares is simply normal.

The Thue–Morse sequence along sparse subsequences

Theorem (Mauduit–Rivat 2009, Acta Math.)

The Thue–Morse sequence along the sequence of squares is simply normal.

Theorem (Drmota–Mauduit–Rivat 2018)

The Thue–Morse sequence along the sequence of squares is normal: every block $B \in \{0, 1\}^k$ *appears as a subword with asymptotic frequency* $1/2^k$.

The Thue–Morse sequence along sparse subsequences

Theorem (Mauduit–Rivat 2009, Acta Math.)

The Thue–Morse sequence along the sequence of squares is simply normal.

Theorem (Drmota–Mauduit–Rivat 2018)

The Thue–Morse sequence along the sequence of squares is normal: every block $B \in \{0, 1\}^k$ *appears as a subword with asymptotic frequency* $1/2^k$*.*

- What about exponents $c \in [1.4, 2)$?

Theorem (Spiegelhofer 2014)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ *is simply normal for* $1 < c \leq 1.42$*.*

The Thue–Morse sequence along sparse subsequences

Theorem (Mauduit–Rivat 2009, Acta Math.)

The Thue–Morse sequence along the sequence of squares is simply normal.

Theorem (Drmota–Mauduit–Rivat 2018)

The Thue–Morse sequence along the sequence of squares is normal: every block $B \in \{0, 1\}^k$ appears as a subword with asymptotic frequency $1/2^k$.

- What about exponents $c \in [1.4, 2)$?

Theorem (Spiegelhofer 2014)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ is simply normal for $1 < c \leq 1.42$.

Theorem (Müllner–Spiegelhofer 2017)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ is normal for $1 < c < 1.5$.

Application: the Thue–Morse sequence along $\lfloor n^c \rfloor$

Theorem (Spiegelhofer 2018+)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ *is simply normal for* $1 < c < 2$.

Application: the Thue–Morse sequence along $\lfloor n^c \rfloor$

Theorem (Spiegelhofer 2018+)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ *is simply normal for* $1 < c < 2$.

The Thue–Morse sequence along $\lfloor n^c \rfloor$ is normal for $1 < c < 2$.

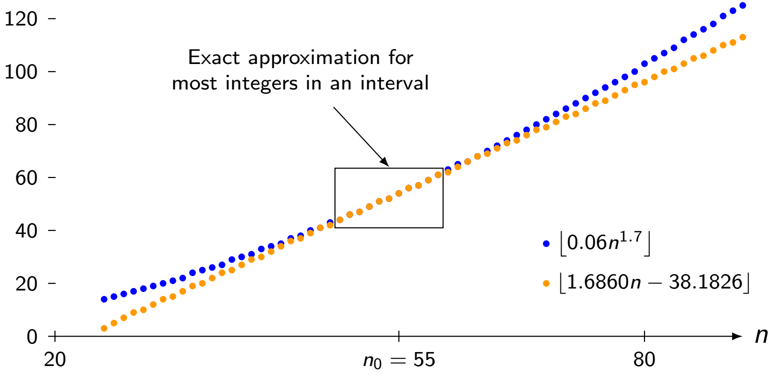
Application: the Thue–Morse sequence along $\lfloor n^c \rfloor$

Theorem (Spiegelhofer 2018+)

The Thue–Morse sequence along $\lfloor n^c \rfloor$ is simply normal for $1 < c < 2$.

☕ The Thue–Morse sequence along $\lfloor n^c \rfloor$ is normal for $1 < c < 2$.

Connection to our main theorem: linear approximation.



Piatetski-Shapiro via Beatty sequences

Proposition (Spiegelhofer 2014)

We write $f(x) = x^c$, where $1 < c < 2$ is a real number. There exists a constant C such that for all $N \geq 2$ and $K > 0$ we have

$$\left| \frac{1}{N} \sum_{N < n \leq 2N} (-1)^{s(\lfloor n^c \rfloor)} \right| \leq C \left(f''(N) K^2 + \frac{(\log N)^2}{K} + \frac{J(f'(N), K)}{f'(N) K} \right),$$

where

$$J(D, K) = \int_D^{2D} \max_{\beta \geq 0} \left| \sum_{0 \leq n < K} (-1)^{s(\lfloor n\alpha + \beta \rfloor)} \right| \mathrm{d}\alpha.$$

Piatetski-Shapiro via Beatty sequences

Proposition (Spiegelhofer 2014)

We write $f(x) = x^c$, where $1 < c < 2$ is a real number. There exists a constant C such that for all $N \geq 2$ and $K > 0$ we have

$$\left| \frac{1}{N} \sum_{N < n \leq 2N} (-1)^{s(\lfloor n^c \rfloor)} \right| \leq C \left(f''(N) K^2 + \frac{(\log N)^2}{K} + \frac{J(f'(N), K)}{f'(N) K} \right),$$

$D = f'(N)$ *is a large power of* K *for* $c \to 2$

where

$$J(D, K) = \int_D^{2D} \max_{\beta \geq 0} \left| \sum_{0 \leq n < K} (-1)^{s(\lfloor n\alpha + \beta \rfloor)} \right| \mathrm{d}\alpha.$$

Piatetski-Shapiro via Beatty sequences

Proposition (Spiegelhofer 2014)

We write $f(x) = x^c$, where $1 < c < 2$ is a real number. There exists a constant C such that for all $N \ge 2$ and $K > 0$ we have

$$\left| \frac{1}{N} \sum_{N < n \le 2N} (-1)^{s(\lfloor n^c \rfloor)} \right| \le C \left(f''(N) K^2 + \frac{(\log N)^2}{K} + \frac{J(f'(N), K)}{f'(N) K} \right),$$

$D = f'(N)$ *is a large power of* K *for* $c \to 2$

where

$$J(D, K) = \int_D^{2D} \max_{\beta \ge 0} \left| \sum_{0 \le n < K} (-1)^{s(\lfloor n\alpha + \beta \rfloor)} \right| \mathrm{d}\alpha.$$

- The problem is reduced to a Beatty sequence version of our main theorem!

Piatetski-Shapiro via Beatty sequences

Proposition (Spiegelhofer 2014)

We write $f(x) = x^c$, where $1 < c < 2$ is a real number. There exists a constant C such that for all $N \geq 2$ and $K > 0$ we have

$$\left| \frac{1}{N} \sum_{N<n\leq 2N} (-1)^{s(\lfloor n^c \rfloor)} \right| \leq C \left(f''(N)K^2 + \frac{(\log N)^2}{K} + \frac{J(f'(N), K)}{f'(N)K} \right),$$

$D = f'(N)$ *is a large power of* K *for* $c \to 2$

where

$$J(D,K) = \int_D^{2D} \max_{\beta \geq 0} \left| \sum_{0\leq n<K} (-1)^{s(\lfloor n\alpha+\beta \rfloor)} \right| \mathrm{d}\alpha.$$

- The problem is reduced to a Beatty sequence version of our main theorem!
- The proof of this new statement is analogous to the main theorem.

Thank you! [1]

[1] Supported by SFB 55 and the FWF–ANR joint project MuDeRa