The Thue-Morse Transform

Benoît Cloitre

One World Combinatorics on Words Seminar
June 23, 2026

Plan

1. **Starting point**
2. The transform and the orbit of $a_0 = 0101\ldots$
3. The binary digit formula
4. Prouhet-Tarry-Escott identities
5. Composition laws
6. Morphic structure
7. Factor complexity
8. Perspectives

The Thue-Morse word

The Thue-Morse word begins

$$t = 0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,\ldots$$

The Thue-Morse word

The Thue-Morse word begins

$$\mathsf{t} = 0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,\ldots$$

The property we build on is its recurrence

$$\mathsf{t}(2n) = \mathsf{t}(n), \qquad \mathsf{t}(2n+1) = 1 - \mathsf{t}(n), \qquad \mathsf{t}(0) = 0.$$

The Thue-Morse word

The Thue-Morse word begins

$$\mathrm{t} = 0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\,\ldots$$

The property we build on is its recurrence

$$\mathrm{t}(2n) = \mathrm{t}(n), \qquad \mathrm{t}(2n+1) = 1 - \mathrm{t}(n), \qquad \mathrm{t}(0) = 0.$$

The driving question

The even and odd numbers partition $\mathbb{N}$. What happens if we keep the recurrence but replace that partition by another one?

A natural candidate: evil and odious numbers

The positions of the 0s and the 1s in t are the classical evil and odious numbers v and u. They partition $\mathbb{N}$. Keeping the recurrence, we define t' by

$$\mathrm{t}'(v(n)) = \mathrm{t}'(n), \qquad \mathrm{t}'(u(n)) = 1 - \mathrm{t}'(n), \qquad \mathrm{t}'(0) = 0.$$

A natural candidate: evil and odious numbers

The positions of the 0s and the 1s in t are the classical evil and odious numbers v and u. They partition $\mathbb{N}$. Keeping the recurrence, we define t' by

$$\mathrm{t}'(v(n)) = \mathrm{t}'(n), \qquad \mathrm{t}'(u(n)) = 1 - \mathrm{t}'(n), \qquad \mathrm{t}'(0) = 0.$$

We obtain a new word,

$$\mathrm{t}' = 0\,1\,0\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,1\,0\,\ldots$$

which is in the OEIS as `A341389`.

A natural candidate: evil and odious numbers

The positions of the 0s and the 1s in t are the classical evil and odious numbers v and u. They partition $\mathbb{N}$. Keeping the recurrence, we define t′ by

$$\mathrm{t}'(v(n)) = \mathrm{t}'(n), \qquad \mathrm{t}'(u(n)) = 1 - \mathrm{t}'(n), \qquad \mathrm{t}'(0) = 0.$$

We obtain a new word,

$$\mathrm{t}' = 0\,1\,0\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1\,0\,1\,0\,\ldots$$

which is in the OEIS as A341389.

However, John Layman had earlier recorded the positions of its 1s and 0s (A158705, A158704) and asked whether they form a Prouhet-Tarry-Escott identity. That question is the seed of this work.

Plan

1. Starting point
2. **The transform and the orbit of $a_0 = 0101\ldots$**
3. The binary digit formula
4. Prouhet-Tarry-Escott identities
5. Composition laws
6. Morphic structure
7. Factor complexity
8. Perspectives

The Thue-Morse transform $\mathcal{T}$

Let w start with 01, with infinitely many 0s and 1s. Let $v_w(n), u_w(n)$ be the positions of the n-th 0 and n-th 1.

Definition

$\mathcal{T}(w) = w'$ is defined by

$$w'(0) = 0, \qquad w'\big(v_w(n)\big) = w'(n), \qquad w'\big(u_w(n)\big) = 1 - w'(n).$$

The Thue-Morse transform $\mathcal{T}$

Let w start with 01, with infinitely many 0s and 1s. Let $v_w(n), u_w(n)$ be the positions of the n-th 0 and n-th 1.

Definition

$\mathcal{T}(w) = w'$ is defined by

$$w'(0) = 0, \qquad w'\big(v_w(n)\big) = w'(n), \qquad w'\big(u_w(n)\big) = 1 - w'(n).$$

$\mathcal{T}(w) = w'$ is well defined, since $\{v_w(n)\} \sqcup \{u_w(n)\} = \mathbb{N}$ and $v_w(n), u_w(n) > n$ for $n \geq 1$.

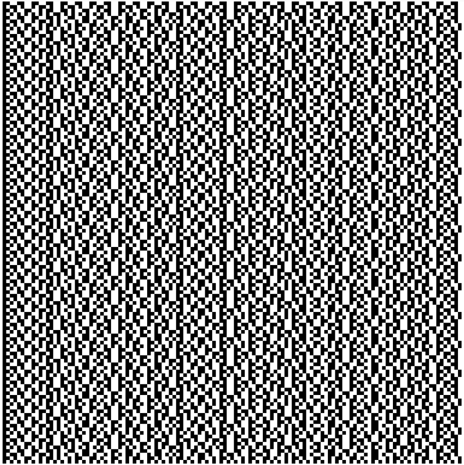
The orbit of the alternating word

Set $a_0(n) = n \bmod 2$ and $a_{m+1} = \mathcal{T}(a_m)$. Then

$$a_1 = \mathtt{t}, \quad a_2 = \mathtt{A341389}, \quad a_3 = \mathtt{A395958}, \quad a_4 = \mathtt{A395961}.$$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a_0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
a_1	0	1	1	0	1	0	0	1	1	0	0	1	0	1	1	0
a_2	0	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0
a_3	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
a_4	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

A 2D portrait of the orbit



White square when $a_m(n) = 0$, black when $a_m(n) = 1$, for $0 \leq m, n \leq 127$.

Plan

1. Starting point
2. The transform and the orbit of $a_0 = 0101\ldots$
3. **The binary digit formula**
4. Prouhet-Tarry-Escott identities
5. Composition laws
6. Morphic structure
7. Factor complexity
8. Perspectives

Main result: the binary digit formula

Write $n = \sum_{p \geq 0} b_p(n)\, 2^p$ and let & be the bitwise AND.

Theorem

For every $m \geq 1$ and every $n \geq 0$,

$$a_m(n) = \bigoplus_{\substack{p \geq 0 \\ p \,\&\, (m-1) = 0}} b_p(n).$$

Main result: the binary digit formula

Write $n = \sum_{p\geq 0} b_p(n)\, 2^p$ and let & be the bitwise AND.

Theorem

For every $m \geq 1$ and every $n \geq 0$,

$$a_m(n) = \bigoplus_{\substack{p\geq 0 \\ p\,\&\,(m-1)=0}} b_p(n).$$

- Concretely, $a_m(n)$ is the sum modulo two of the binary digits of n whose position avoids the 1-bits of $m-1$.
- Everything that follows (PTE, composition, complexity) is read off this one formula.
- For $m = 2^k$, this is the Thue–Morse digit-sum rule in base 2^{2^k}.

The formula at work

Goal. Compute $a_{10}(365)$, with $n = 101101101_2$ and $m - 1 = 1001_2$.

The formula at work

Goal. Compute $a_{10}(365)$, with $n = 101101101_2$ and $m - 1 = 1001_2$.

p	8	7	6	5	4	3	2	1	0
$m-1$ in base 2	1001	1001	1001	1001	1001	1001	1001	1001	1001
p in base 2	1000	111	110	101	100	11	10	1	0
$p \,\&\, (m-1)$	8	1	0	1	0	1	0	1	0
$b_p(n)$	1	0	1	1	0	1	1	0	1

The formula at work

Goal. Compute $a_{10}(365)$, with $n = 101101101_2$ and $m - 1 = 1001_2$.

p	8	7	6	5	4	3	2	1	0
$m - 1$ in base 2	1001	1001	1001	1001	1001	1001	1001	1001	1001
p in base 2	1000	111	110	101	100	11	10	1	0
$p \,\&\, (m-1)$	8	1	0	1	0	1	0	1	0
$b_p(n)$	1	0	1	1	0	1	1	0	1

Result. $p \,\&\, (m-1) = 0$ at positions $6, 4, 2, 0$, so

$$a_{10}(365) = b_6 \oplus b_4 \oplus b_2 \oplus b_0 = 1 \oplus 0 \oplus 1 \oplus 1 = 1.$$

Two proofs of the binary digit formula

Along the orbit each iterate satisfies

$$a_m(n) = a_m(\lfloor n/2 \rfloor) \oplus a_{m-1}(n), \qquad a_m(0) = 0,$$

because consecutive positions $\{2k, 2k+1\}$ carry one 0 and one 1. Unfolding it gives

$$a_m(n) = \bigoplus_{p \geq 0} \left(\binom{p+m-1}{m-1} \bmod 2 \right) b_p(n),$$

with the binomial arising two ways.

Two proofs of the binary digit formula

Along the orbit each iterate satisfies

$$a_m(n) = a_m(\lfloor n/2 \rfloor) \oplus a_{m-1}(n), \qquad a_m(0) = 0,$$

because consecutive positions $\{2k, 2k+1\}$ carry one 0 and one 1. Unfolding it gives

$$a_m(n) = \bigoplus_{p \ge 0} \left(\binom{p+m-1}{m-1} \bmod 2 \right) b_p(n),$$

with the binomial arising two ways.

Combinatorial (first proof).

Each $b_p(n)$ contributes once for each ordered sum $p = k_1 + \cdots + k_m$ of m nonnegative shifts, i.e. $\binom{p+m-1}{m-1}$ times.

Algebraic (second proof).

Write $a_m(n) = \bigoplus_p \gamma_m(p)\, b_p(n)$. Over $\mathbb{F}_2$, $\mathcal{T}$ multiplies the series $\sum_p \gamma_m(p) x^p$ by $(1-x)^{-1}$ (partial summation). From a_0 this gives $(1-x)^{-m} = \sum_p \binom{p+m-1}{m-1} x^p$.

Lucas' theorem closes the proof of the binary digit formula

Working modulo 2, Lucas' theorem closes the argument:

$$\binom{p+m-1}{m-1} \text{ is odd} \iff p+(m-1) \text{ has no carry in base 2} \iff p \,\&\, (m-1) = 0.$$

Lucas' theorem closes the proof of the binary digit formula

Working modulo 2, Lucas' theorem closes the argument:

$$\binom{p+m-1}{m-1} \text{ is odd } \iff p+(m-1) \text{ has no carry in base 2} \iff p \,\&\, (m-1) = 0.$$

Keeping only the odd coefficients recovers the binary digit formula,

$$a_m(n) = \bigoplus_{\substack{p \geq 0 \\ p \,\&\, (m-1)=0}} b_p(n).$$

Plan

Recalling the Prouhet-Tarry-Escott problem

Split $\{0, 1, \dots, L-1\}$ into two equal parts $A \sqcup B$. It solves the PTE problem of degree k if the power sums agree up to degree k:

$$\sum_{a \in A} a^j = \sum_{b \in B} b^j, \qquad j = 0, 1, \dots, k.$$

Recalling the Prouhet-Tarry-Escott problem

Split $\{0, 1, \ldots, L-1\}$ into two equal parts $A \sqcup B$. It solves the PTE problem of degree k if the power sums agree up to degree k:

$$\sum_{a \in A} a^j = \sum_{b \in B} b^j, \qquad j = 0, 1, \ldots, k.$$

Prouhet (1851) solved it by splitting $\{0, \ldots, 2^{k+1}-1\}$ by the parity of the binary digit sum. That parity sequence is exactly the Thue-Morse word $\mathrm{t} = a_1$, hence the name Prouhet-Thue-Morse sequence (Allouche, Shallit). For $k = 2$, the evil/odious split of $\{0, \ldots, 7\}$,

$$\{0, 3, 5, 6\} \quad \text{vs} \quad \{1, 2, 4, 7\},$$

has equal sums $(14 = 14)$, equal sums of squares $(70 = 70)$, but not cubes $(368 \neq 416)$.

From the binary digit formula to PTE

The binary digit formula turns every such evil/odious split into a generating function.

For $S \subset \{0, \ldots, M-1\}$ and $f_S(n) = \bigoplus_{p \in S} b_p(n)$,

$$\sum_{n=0}^{2^M-1} (-1)^{f_S(n)} x^n = \prod_{p \in S} (1 - x^{2^p}) \prod_{\substack{0 \le p < M \\ p \notin S}} (1 + x^{2^p}).$$

From the binary digit formula to PTE

The binary digit formula turns every such evil/odious split into a generating function.
For $S \subset \{0, \ldots, M-1\}$ and $f_S(n) = \bigoplus_{p \in S} b_p(n)$,

$$\sum_{n=0}^{2^M-1} (-1)^{f_S(n)} x^n = \prod_{p \in S} (1 - x^{2^p}) \prod_{\substack{0 \le p < M \\ p \notin S}} (1 + x^{2^p}).$$

The order of vanishing at $x = 1$ equals $|S|$, which forces equal power sums up to degree $|S| - 1$.

From the binary digit formula to PTE

The binary digit formula turns every such evil/odious split into a generating function. For $S \subset \{0, \ldots, M-1\}$ and $f_S(n) = \bigoplus_{p \in S} b_p(n)$,

$$\sum_{n=0}^{2^M-1} (-1)^{f_S(n)} x^n = \prod_{p \in S} (1 - x^{2^p}) \prod_{\substack{0 \le p < M \\ p \notin S}} (1 + x^{2^p}).$$

The order of vanishing at $x = 1$ equals $|S|$, which forces equal power sums up to degree $|S| - 1$.

Theorem

For a_m take $S_{m,M} = \{0 \le p < M : p \,\&\, (m-1) = 0\}$. The split of $[0, 2^M)$ into evil and odious positions of a_m solves the PTE problem of degree $|S_{m,M}| - 1$.

From the binary digit formula to PTE

The binary digit formula turns every such evil/odious split into a generating function. For $S \subset \{0, \ldots, M-1\}$ and $f_S(n) = \bigoplus_{p \in S} b_p(n)$,

$$\sum_{n=0}^{2^M-1} (-1)^{f_S(n)} x^n = \prod_{p \in S} (1 - x^{2^p}) \prod_{\substack{0 \le p < M \\ p \notin S}} (1 + x^{2^p}).$$

The order of vanishing at $x = 1$ equals $|S|$, which forces equal power sums up to degree $|S| - 1$.

Theorem

For a_m take $S_{m,M} = \{0 \le p < M : p \,\&\, (m-1) = 0\}$. The split of $[0, 2^M)$ into evil and odious positions of a_m solves the PTE problem of degree $|S_{m,M}| - 1$.

At $m = 2$, $S_{2,M}$ is the $\lceil M/2 \rceil$ even positions, so the split has degree $\lceil M/2 \rceil - 1$. This answers Layman's question for A158704, A158705.

Plan

Composition laws for u_m and v_m

The generalized evil and odious numbers $v_m(n), u_m(n)$ (positions of the n-th 0 and 1) read off a_m directly:

$$v_m(n) = 2n + a_m(2n), \qquad u_m(n) = 2n + a_m(2n+1).$$

Composition laws for u_m and v_m

The generalized evil and odious numbers $v_m(n), u_m(n)$ (positions of the n-th 0 and 1) read off a_m directly:

$$v_m(n) = 2n + a_m(2n), \qquad u_m(n) = 2n + a_m(2n+1).$$

The classical numbers v, u (case $m = 1$) satisfy

$$u(u(n)) = 2u(n), \quad v(v(n)) = 2v(n), \quad u(v(n)) = 2v(n) + 1, \quad v(u(n)) = 2u(n) + 1.$$

Composition laws for u_m and v_m

The generalized evil and odious numbers $v_m(n), u_m(n)$ (positions of the n-th 0 and 1) read off a_m directly:

$$v_m(n) = 2n + a_m(2n), \qquad u_m(n) = 2n + a_m(2n+1).$$

The classical numbers v, u (case $m = 1$) satisfy

$$u(u(n)) = 2u(n), \quad v(v(n)) = 2v(n), \quad u(v(n)) = 2v(n) + 1, \quad v(u(n)) = 2u(n) + 1.$$

Writing $w_m^0 = v_m$, $w_m^1 = u_m$ and $\varepsilon_m \equiv m \pmod 2$,

Theorem

For every $m \geq 1$ and all $\alpha, \beta \in \{0, 1\}$,

$$w_m^{\alpha}(w_m^{\beta}(n)) = 2\, w_m^{\beta}(n) + \big(c_m(n) \oplus \alpha \oplus \varepsilon_m \beta\big).$$

Composition laws for u_m and v_m

The generalized evil and odious numbers $v_m(n), u_m(n)$ (positions of the n-th 0 and 1) read off a_m directly:

$$v_m(n) = 2n + a_m(2n), \qquad u_m(n) = 2n + a_m(2n+1).$$

The classical numbers v, u (case $m = 1$) satisfy

$$u(u(n)) = 2u(n), \quad v(v(n)) = 2v(n), \quad u(v(n)) = 2v(n) + 1, \quad v(u(n)) = 2u(n) + 1.$$

Writing $w_m^0 = v_m$, $w_m^1 = u_m$ and $\varepsilon_m \equiv m \pmod 2$,

Theorem

For every $m \geq 1$ and all $\alpha, \beta \in \{0, 1\}$,

$$w_m^{\alpha}(w_m^{\beta}(n)) = 2\, w_m^{\beta}(n) + \big(c_m(n) \oplus \alpha \oplus \varepsilon_m \beta\big).$$

c_m is 2-automatic: $c_m(n) = a_m(4n)$ (m even), $a_m(4n) \oplus a_m(2n)$ (m odd); $c_2 = a_2$.

The eight compositions, by parity of m

m even

$$
\begin{aligned}
v_m\big(v_m(n)\big) &= 2v_m(n) + c_m(n),\\
u_m\big(v_m(n)\big) &= 2v_m(n) + 1 - c_m(n),\\
v_m\big(u_m(n)\big) &= 2u_m(n) + c_m(n),\\
u_m\big(u_m(n)\big) &= 2u_m(n) + 1 - c_m(n).
\end{aligned}
$$

m odd

$$
\begin{aligned}
v_m\big(v_m(n)\big) &= 2v_m(n) + c_m(n),\\
u_m\big(v_m(n)\big) &= 2v_m(n) + 1 - c_m(n),\\
v_m\big(u_m(n)\big) &= 2u_m(n) + 1 - c_m(n),\\
u_m\big(u_m(n)\big) &= 2u_m(n) + c_m(n).
\end{aligned}
$$

Plan

Each iterate is the fixed point of a uniform morphism

Set $k = \lceil \log_2 m \rceil$ and $B = 2^{2^k}$. The morphism μ_m on $\{0, 1\}$ is

$$\mu_m(0) = a_m(0)\, a_m(1) \cdots a_m(B-1), \qquad \mu_m(1) = \overline{\mu_m(0)}.$$

So $\mu_m(0)$ is simply the length-B prefix of a_m, and $\mu_m(1)$ its complement. Then a_m is the fixed point of μ_m starting with 0, hence 2-automatic.

Each iterate is the fixed point of a uniform morphism

Set $k = \lceil \log_2 m \rceil$ and $B = 2^{2^k}$. The morphism μ_m on $\{0, 1\}$ is

$$\mu_m(0) = a_m(0)\, a_m(1) \cdots a_m(B-1), \qquad \mu_m(1) = \overline{\mu_m(0)}.$$

So $\mu_m(0)$ is simply the length-B prefix of a_m, and $\mu_m(1)$ its complement. Then a_m is the fixed point of μ_m starting with 0, hence 2-automatic.

Example: the Layman morphism ($m = 2$, $B = 4$)

$\mu_2(0) = 0101$, $\mu_2(1) = 1010$, and

$$0 \;\to\; 0101 \;\to\; 0101\,1010\,0101\,1010 \;\to\; \cdots \;=\; a_2.$$

Each iterate is the fixed point of a uniform morphism

Set $k = \lceil \log_2 m \rceil$ and $B = 2^{2^k}$. The morphism μ_m on $\{0, 1\}$ is

$$\mu_m(0) = a_m(0)\, a_m(1) \cdots a_m(B-1), \qquad \mu_m(1) = \overline{\mu_m(0)}.$$

So $\mu_m(0)$ is simply the length-B prefix of a_m, and $\mu_m(1)$ its complement. Then a_m is the fixed point of μ_m starting with 0, hence 2-automatic.

Example: the Layman morphism ($m = 2$, $B = 4$)

$\mu_2(0) = 0101$, $\mu_2(1) = 1010$, and

$$0 \;\to\; 0101 \;\to\; 0101\,1010\,0101\,1010 \;\to\; \cdots \;=\; a_2.$$

Each μ_m is balanced in the sense of Černý (2013, 2017), who derived PTE identities from subword spectra. Iterating μ_m gives an alternative, morphism-based derivation of the same PTE phenomenon, on its natural block lengths.

Plan

1. Starting point
2. The transform and the orbit of $a_0 = 0101\ldots$
3. The binary digit formula
4. Prouhet-Tarry-Escott identities
5. Composition laws
6. Morphic structure
7. **Factor complexity**
8. Perspectives

A uniform recurrence for the factor complexity

Theorem

Let $k = \lceil \log_2 m \rceil$, $B = 2^{2^k}$. *For* $n \geq 2$,

$$p_m(Bn) = p_m(n) + (B-1)p_m(n+1),$$

and for $n \geq 2$ *and* $1 \leq r \leq B-1$,

$$p_m(Bn + r) = (B + 1 - r)p_m(n+1) + (r-1)p_m(n+2).$$

A uniform recurrence for the factor complexity

Theorem

Let $k = \lceil \log_2 m \rceil$, $B = 2^{2^k}$. *For* $n \geq 2$,

$$p_m(Bn) = p_m(n) + (B-1)p_m(n+1),$$

and for $n \geq 2$ *and* $1 \leq r \leq B-1$,

$$p_m(Bn+r) = (B+1-r)p_m(n+1) + (r-1)p_m(n+2).$$

At $m = 1$ ($B = 2$) this is exactly Brlek's recurrence for Thue-Morse,

$$p_1(2n) = p_1(n) + p_1(n+1), \qquad p_1(2n+1) = 2p_1(n+1).$$

The proof goes through the derivative word $\Delta_m(n) = a_m(n) \oplus a_m(n+1)$ and a synchronization property at delay $2B-1$.

Explicit formula at $m = 2^k$

For $m = 2^k$ and $B = 2^{2^k}$, one has $p_m(n) = 2n$ for $1 \leq n \leq B+1$, and for $i \geq 0$,

$$p_m(n) = \begin{cases} -2(B-1)B^i + 4(n-1), & B^{i+1}+1 < n \leq (2B-1)B^i + 1, \\ 2B^{i+1} + 2(n-1), & (2B-1)B^i + 1 < n \leq B^{i+2}+1. \end{cases}$$

The complexity is piecewise linear, the slope alternating between 4 and 2 and changing at the points B^{i+1} and $(2B-1)B^i$.

Plan

Woods-Robbins type products

The classical Woods-Robbins identity reads

$$\prod_{n\geq 0}\left(\frac{2n+1}{2n+2}\right)^{(-1)^{\mathtt{t}(n)}}=\frac{\sqrt{2}}{2},$$

where $\mathtt{t}=a_1$ is the Thue-Morse word.

Woods-Robbins type products

The classical Woods-Robbins identity reads

$$\prod_{n\geq 0}\left(\frac{2n+1}{2n+2}\right)^{(-1)^{\mathtt{t}(n)}} = \frac{\sqrt{2}}{2},$$

where $\mathtt{t} = a_1$ is the Thue-Morse word.

Generalization to a_{2^k} (with $B = 2^{2^k}$)

$$\prod_{n\geq 0}\left(\prod_{j=0}^{B/2-1}\frac{Bn+2j+1}{Bn+2j+2}\right)^{(-1)^{a_{2^k}(n)}} = 2^{-(2^{k-1})}.$$

Woods-Robbins type products

The classical Woods-Robbins identity reads

$$\prod_{n\geq 0}\left(\frac{2n+1}{2n+2}\right)^{(-1)^{\mathrm{t}(n)}} = \frac{\sqrt{2}}{2},$$

where $\mathrm{t} = a_1$ is the Thue-Morse word.

Generalization to a_{2^k} (with $B = 2^{2^k}$)

$$\prod_{n\geq 0}\left(\prod_{j=0}^{B/2-1}\frac{Bn+2j+1}{Bn+2j+2}\right)^{(-1)^{a_{2^k}(n)}} = 2^{-(2^{k-1})}.$$

For $k = 0$ $(B = 2)$ it reduces to the classical identity.

A q-ary transform

On $\{0, 1, \dots, q-1\}$, copy/complement is replaced by addition of the letter modulo q, with $\alpha_j(n)$ the position of the n-th letter j.

Definition ($\mathcal{T}_2 = \mathcal{T}$)

$$\mathcal{T}_q(w) = w' : \quad w'(0) = 0, \quad w'\big(\alpha_j(n)\big) = w'(n) + j \pmod{q}.$$

A q-ary transform

On $\{0, 1, \ldots, q-1\}$, copy/complement is replaced by addition of the letter modulo q, with $\alpha_j(n)$ the position of the n-th letter j.

Definition ($\mathcal{T}_2 = \mathcal{T}$)

$$\mathcal{T}_q(w) = w' : \quad w'(0) = 0, \quad w'\big(\alpha_j(n)\big) = w'(n) + j \pmod{q}.$$

$$a_0^{(q)}(n) = n \bmod q \xrightarrow{\mathcal{T}_q} \Big(\textstyle\sum_p \delta_p(n)\Big) \bmod q \quad \text{(base-}q\text{ digit sum, Prouhet 1851)},$$

where $\delta_p(n)$ are the base-q digits of n.

A q-ary transform

On $\{0,1,\ldots,q-1\}$, copy/complement is replaced by addition of the letter modulo q, with $\alpha_j(n)$ the position of the n-th letter j.

Definition ($\mathcal{T}_2 = \mathcal{T}$)

$$\mathcal{T}_q(w) = w' : \quad w'(0) = 0, \quad w'(\alpha_j(n)) = w'(n) + j \pmod q.$$

$$a_0^{(q)}(n) = n \bmod q \xrightarrow{\mathcal{T}_q} \Big(\textstyle\sum_p \delta_p(n)\Big) \bmod q \quad \text{(base-}q\text{ digit sum, Prouhet 1851)},$$

where $\delta_p(n)$ are the base-q digits of n.

Future work

The same program for $\mathcal{T}_q^{\,m}(a_0^{(q)})$: digit formula, PTE, composition, morphisms, complexity.

Other seeds

For the Fibonacci word f (A003849),

$$\mathcal{T}(f) = \mathrm{ftm} \quad (\texttt{A095076}),$$

the Fibonacci-Thue-Morse word, the parity of the Zeckendorf digit sum.

Other seeds

For the Fibonacci word f (A003849),

$$\mathcal{T}(f) = \mathrm{ftm} \quad (\text{A095076}),$$

the Fibonacci-Thue-Morse word, the parity of the Zeckendorf digit sum.

- ftm is Fibonacci-automatic.
- But $\mathcal{T}^2(f)$ appears not to be Fibonacci-automatic.

A dynamical viewpoint: toward X

The stable space

$X = \{w \text{ in the domain} : w(2k) \neq w(2k+1) \text{ for all } k\}$ (complementary pairs). $\mathcal{T}$ maps X into itself.

For any seed s in the domain,

$$F_m(N) = \#\{\, n < N : \mathcal{T}^m(s)(2n) \neq \mathcal{T}^m(s)(2n+1)\,\}.$$

A dynamical viewpoint: toward X

The stable space

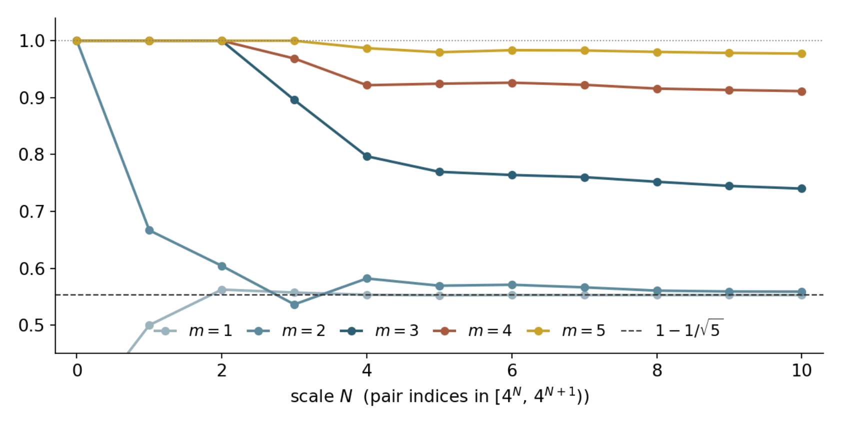
$X = \{w \text{ in the domain} : w(2k) \neq w(2k+1) \text{ for all } k\}$ (complementary pairs). $\mathcal{T}$ maps X into itself.

For any seed s in the domain,

$$F_m(N) = \#\{\, n < N : \mathcal{T}^m(s)(2n) \neq \mathcal{T}^m(s)(2n+1)\,\}.$$

- For $s = a_0$, every iterate lies in X, so $F_m(N) = N$.
- For the Fibonacci seed, $F_m(N)/N$ appears to converge, the orbit approaching X as m grows.

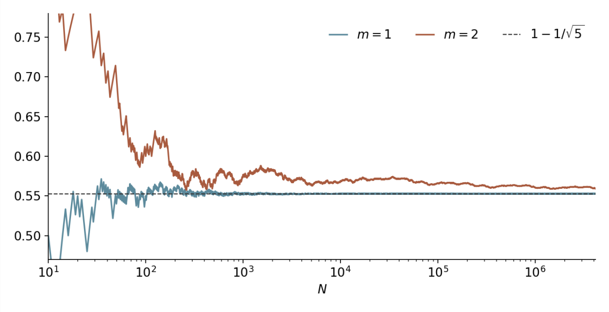
Approaching X across iterates



Local density of complementary pairs of $\mathcal{T}^m(f)$ on the slice $[4^N, 4^{N+1})$. The curves seem to level off, suggesting limiting densities.

Focus on $m = 1$ and $m = 2$

The graphs of $F_m(N)/N$ against N, for the Fibonacci word.



For $m = 1$ the ratio converges to $1 - 1/\sqrt{5}$. For $m = 2$ it seems to approach the same value, but slowly.

$m = 1$ is provable, $m = 2$ and $m = 3$ are open

$m = 1$. ftm is the coding $A, C \mapsto 0$, $B, D \mapsto 1$ of the fixed point of

$$A \to AD, \quad B \to BC, \quad C \to A, \quad D \to B,$$

primitive with dominant eigenvalue φ. The induced substitution on the aligned pairs $(2n, 2n+1)$ gives

$$F_1(N) = \left(1 - \tfrac{1}{\sqrt{5}}\right) N + R(N), \qquad R(N) = o(N),$$

so $F_1(N)/N \to 1 - 1/\sqrt{5}$. We conjecture $R(N) = O(\log N)$.

$m = 1$ is provable, $m = 2$ and $m = 3$ are open

$m = 1$. ftm is the coding $A, C \mapsto 0$, $B, D \mapsto 1$ of the fixed point of

$$A \to AD, \quad B \to BC, \quad C \to A, \quad D \to B,$$

primitive with dominant eigenvalue φ. The induced substitution on the aligned pairs $(2n, 2n+1)$ gives

$$F_1(N) = \left(1 - \tfrac{1}{\sqrt{5}}\right) N + R(N), \qquad R(N) = o(N),$$

so $F_1(N)/N \to 1 - 1/\sqrt{5}$. We conjecture $R(N) = O(\log N)$.

$m = 2$ **(open).** $\mathcal{T}^2(f) = \mathcal{T}(\mathrm{ftm})$ appears not to be Fibonacci-automatic. $F_2(N)/N$ seems to tend to the same $1 - 1/\sqrt{5}$, but with a remainder of order about $N^{0.84}$, well above $\sqrt{N}$. Convergence and the order of the remainder are open.

$m = 1$ is provable, $m = 2$ and $m = 3$ are open

$m = 1$. ftm is the coding $A, C \mapsto 0$, $B, D \mapsto 1$ of the fixed point of

$$A \to AD, \quad B \to BC, \quad C \to A, \quad D \to B,$$

primitive with dominant eigenvalue φ. The induced substitution on the aligned pairs $(2n, 2n+1)$ gives

$$F_1(N) = \left(1 - \tfrac{1}{\sqrt{5}}\right) N + R(N), \qquad R(N) = o(N),$$

so $F_1(N)/N \to 1 - 1/\sqrt{5}$. We conjecture $R(N) = O(\log N)$.

$m = 2$ **(open).** $\mathcal{T}^2(f) = \mathcal{T}(\mathrm{ftm})$ appears not to be Fibonacci-automatic. $F_2(N)/N$ seems to tend to the same $1 - 1/\sqrt{5}$, but with a remainder of order about $N^{0.84}$, well above $\sqrt{N}$. Convergence and the order of the remainder are open.

$m = 3$ **(open).** $F_3(N)/N$ seems to tend to a different limit, near 0.74, not $1 - 1/\sqrt{5}$. Its exact value is unknown.

Thank you.

Preprint `arXiv:2604.06243`