

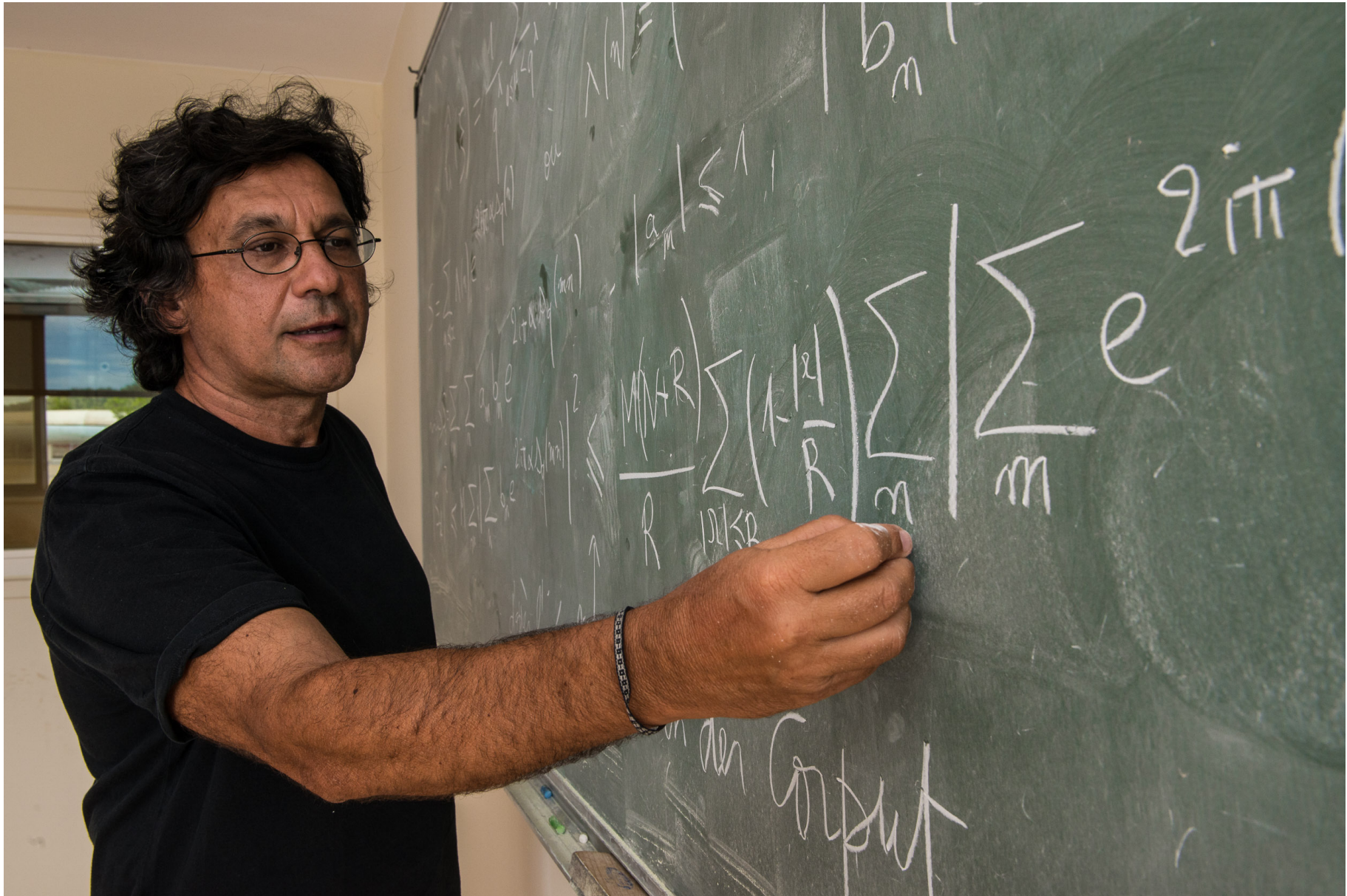
Conference in Number Theory and Applications
in Memory of Christian Mauduit (1959–2019)
Hammamet, Saturday, April 11th, 2026

The Thue-Morse word in base $3/2$

Julien Cassaigne

Institut de mathématiques de Marseille
UMR 7373 CNRS/AMU, Marseille, France
julien.cassaigne@math.cnrs.fr

Joint work with
Bastián Espinoza, Michel Rigo, Manon Stipulanti (Liège)
and Sasha Jaufeeraully (Marseille).



The Thue-Morse word in base $3/2$

- The classical Thue-Morse word
- Base $3/2$: two definitions
- Thue-Morse in base $3/2$
- Alternating fixed points
- First difference word
- Uniform recurrence
- Symmetry properties
- Complexity
- Frequencies

The classical Thue-Morse word

Write $n \in \mathbb{N}$ in binary: $n = \sum_{i=0}^k a_i 2^i$ with $a_i \in \{0, 1\}$.

Then let $t_n = \sum_{i=0}^k a_i \bmod 2 \in \{0, 1\} = A$.

The Thue-Morse word is

$$\mathbf{t} = t_0 t_1 t_2 \dots = 01101001100101101001 \dots \in A^{\mathbb{N}}$$

It is the fixed point of the substitution

$$\begin{aligned} \theta : A^* &\longrightarrow A^* \\ 0 &\longmapsto 01 \\ 1 &\longmapsto 10 \end{aligned}$$

The classical Thue-Morse word

$t = 01101001100101101001011001101001\dots$

The word t is uniformly recurrent.

Its language is invariant under complement and reversal.

Its factor complexity has linear growth $p(n) = \Theta(n)$.
More precisely, $3n \leq p(n+1) \leq \frac{10}{3}n$, bounds attained.

Both letters have frequency $f(0) = f(1) = \frac{1}{2}$ in t .
The frequencies of other words can be determined,
e.g. $f(00) = f(11) = \frac{1}{6}$ and $f(01) = f(10) = \frac{1}{3}$.

Base 3/2 (first definition)

[Propp 2017]

Let $n \in \mathbb{N}$, and write it as $n = \sum_{i=0}^k b_i \left(\frac{3}{2}\right)^i$ with $b_i \in \{0, 1, 2\}$.

There is a unique way to do this without leading zeros.

0	ε	9	2100	18	21200	27	212000
1	1	10	2101	19	21201	28	212001
2	2	11	2102	20	21202	29	212002
3	20	12	2120	21	21220	30	212020
4	21	13	2121	22	21221	31	212021
5	22	14	2122	23	21222	32	212022
6	210	15	21010	24	210110	33	212210
7	211	16	21011	25	210111	34	212211
8	212	17	21012	26	210112	35	212212

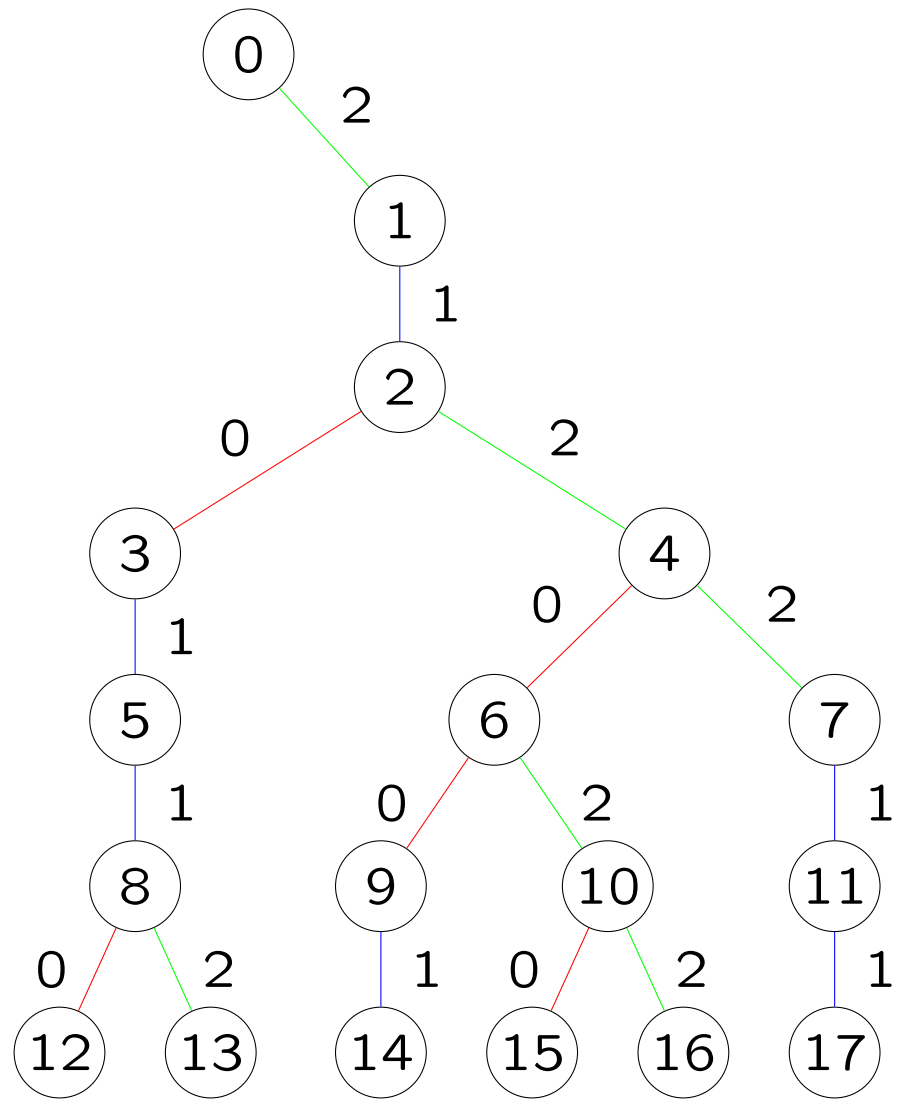
Base 3/2 (second definition)

[Akiyama, Frougny, Sakarovitch 2008]

Let $n \in \mathbb{N}$, and write it as $n = \sum_{i=0}^k \frac{c_i}{2} \left(\frac{3}{2}\right)^i$ with $\frac{c_i}{2} \in \{\frac{0}{2}, \frac{1}{2}, \frac{2}{2}\}$.

There is a unique way to do this without leading zeros.

0	ε	9	21200	18	2101100	27	21011000
1	2	10	21202	19	2101102	28	21011002
2	21	11	21221	20	2101121	29	21011021
3	210	12	210110	21	2120010	30	21011210
4	212	13	210112	22	2120012	31	21011212
5	2101	14	212001	23	2120201	32	21200101
6	2120	15	212020	24	2120220	33	21200120
7	2122	16	212022	25	2120222	34	21200122
8	21011	17	212211	26	2122111	35	21202011



Thue-Morse in base 3/2

[Dekking 2023]

With the first definition of base 3/2:

if $n = \sum_{i=0}^k b_i \left(\frac{3}{2}\right)^i$, let $x_n = \sum_{i=0}^k b_i \bmod 2$.

$$\mathbf{x} = x_0x_1x_2\dots = 01001010110110101010101101101\dots \in A^{\mathbb{N}}$$

With the second definition:

if $n = \sum_{i=0}^k \frac{c_i}{2} \left(\frac{3}{2}\right)^i$, let $y_n = \sum_{i=0}^k c_i \bmod 2$.

$$\mathbf{y} = y_0y_1y_2\dots = 001110111110110111110000110\dots \in A^{\mathbb{N}}$$

Thue-Morse in base 3/2

$x = 010010101101101010101101101101101010101101010\dots$
 $y = 001110111110110111110000110110001001110000111\dots$

They are linked: $y_n = x_{2n}$, so that $y = \delta_2(x)$
(δ_2 deletes every second symbol, not a substitution)
while $x = \varphi(y)$ with $\varphi(0) = 010$, $\varphi(1) = 101$ (a substitution).

They are fixed by **2-block substitutions**: $x = \kappa(x)$ and $y = \kappa'(y)$ where

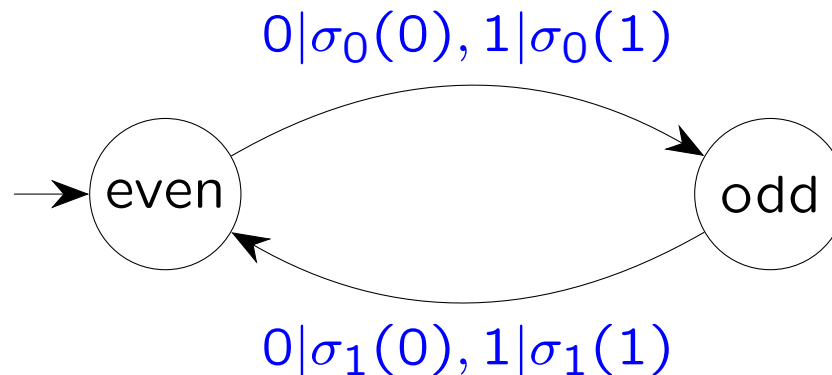
$\kappa :$	00	\mapsto	010		$\kappa' :$	00	\mapsto	001
	01	\mapsto	010	and		01	\mapsto	000
	10	\mapsto	101			10	\mapsto	110
	11	\mapsto	101			11	\mapsto	111

Alternating fixed points

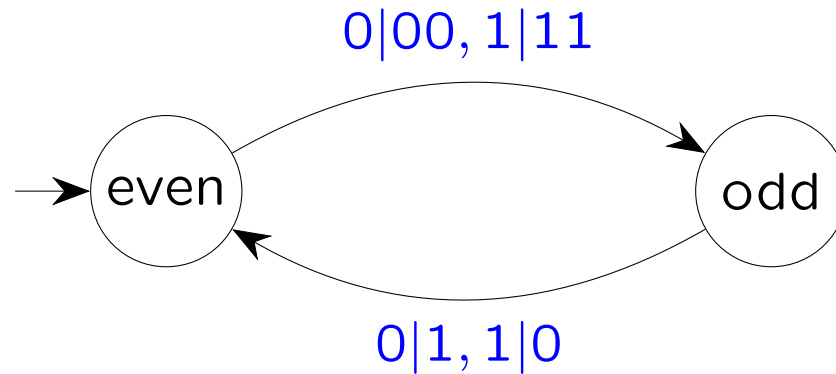
Let σ_0, σ_1 be two substitutions, and define

$$(\sigma_0, \sigma_1)(\mathbf{u}) = \sigma_0(u_0)\sigma_1(u_1)\sigma_0(u_2)\sigma_1(u_3)\dots\sigma_0(u_{2n})\sigma_1(u_{2n+1})\dots$$

\mathbf{u} is a fixed point of (σ_0, σ_1) if and only if it is a fixed point of the 2-block substitution $ab \mapsto \sigma_0(a)\sigma_1(b)$, or equivalently of the transducer:

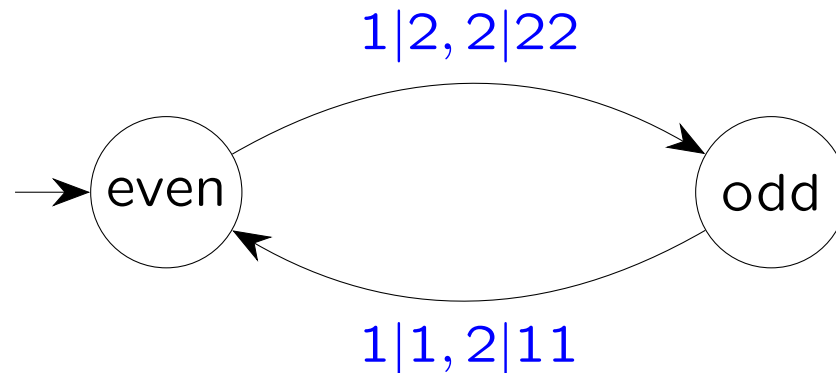


For instance, y is fixed by



Compare with the famous Oldenburger-Kolakoski word

$k = 2211212212211211221211212211211212212211 \dots$ fixed by



Substitution vs 2-block substitution

To prove e.g. uniform recurrence by induction:

For a fixed point \mathbf{u} of a substitution σ :

If some factor w occurs in \mathbf{u} , then it is a factor of $\sigma(v)$ for $|v| < |w|$ (provided w is not too short). Then every occurrence of v provides a new occurrence of w .

For a fixed point of a 2-block substitution κ :

If w occurs, then it is a factor of $\kappa(v)$ for $|v| < |w|$. But an occurrence of v provides a new occurrence of w only when it occurs at an even position.

First difference word

Let $z = \Delta y$, i.e., $z_n = y_{n+1} - y_n \pmod 2$.

(For the usual Thue-Morse word: Δt is the **period-doubling word**.)

$y = 001.110.111.110.110.111.110.000.110.110.001.001.110.000.111 \dots$

$z = 010.011.000.011.011.000.010.001.011.010.011.010.010.001.00 \dots$

As $y_{3n} = y_{3n+1}$, it follows that z is a (regular) **Toeplitz word**.

$z = (0??)^\omega \triangleleft \bar{z}$, where $u \triangleleft v$ means: fill the holes ($?$ symbols) of u with the letters of v .

Regular Toeplitz words have nice properties, as uniform recurrence.

Here: for all $k \geq 1$, the prefix of z of length $\lfloor \left(\frac{3}{2}\right)^k \rfloor$ occurs at all positions multiple of 3^k .

Uniform recurrence

Lemma. A binary word \mathbf{u} is uniformly recurrent if and only if $\Delta\mathbf{u}$ is uniformly recurrent.

Proof (\Leftarrow). Let w be a factor of \mathbf{u} . If \bar{w} does not occur, then as Δw occurs in $\Delta\mathbf{u}$ with bounded gaps, so does w in \mathbf{u} .

If \bar{w} occurs, consider v containing both w and \bar{w} . Then $\{v, \bar{v}\}$ occur with bounded gaps, and both contain w .

Theorem. The words \mathbf{y} and \mathbf{x} are uniformly recurrent.

Symmetry properties

Theorem.

Any factor of z (resp. y , x) occurs both at even and odd positions.

The languages of y and x are invariant under complement (z is not).

The languages of z , y , and x are invariant under reversal.

Complexity

The complexity of regular Toeplitz words has explicit polynomial growth.

Theorem. $p_z(n) = \Theta(n^r)$, with $r = \frac{\log(3)}{\log(3/2)} \simeq 2.7095$.

Lemma. $p_y(n + 1) = 2p_z(n)$.

Theorem. $p_y(n) = \Theta(n^r)$ and $p_x(n) = \Theta(n^r)$.

Note that, for the Oldenburger-Kolakoski word, it is conjectured (and conditionally proved) that $p_k(n) = \Theta(n^r)$ with the same exponent r .

Frequencies

Theorem. In \mathbf{y} and \mathbf{x} , letters 0 and 1 both occur with frequency $\frac{1}{2}$. More precisely, for any $c \in A$, $k \in \mathbb{N}$, and $j \in \mathbb{Z}$, the frequency of the letter c at a position congruent to j modulo 2^k is 2^{-k-1} .

Proof uses a new technique developed by Bastián Espinoza, based on harmonic analysis on \mathbb{Z}_2 .

Open. Compute explicitly the frequency $f(w)$ of a word w in \mathbf{x} or \mathbf{y} .

It seems that $f(00) = f(11) = \frac{1}{10}$ and $f(01) = f(10) = \frac{4}{10}$.

Bibliography

Julien Cassaigne, Bastiàn Espinoza, Michel Rigo, and Manon Stipulanti. Symbols frequencies in the Thue–Morse word in base $3/2$ and related conjectures. arXiv:2602.21895, 2026.

Frederik Michel Dekking. The Thue-Morse sequence in base $3/2$. *J. Integer Seq.* **26**, article 23.2.3, 2023.

