On the real roots of generalized Thue–Morse polynomials

by

Christophe Doche (Talence)

In this article we investigate real roots of real polynomials. By results of M. Kac [8–10] we know that a polynomial of degree $n$ has on average $(2/\pi)\log n$ real zeros. See also results of Edelman and Kostlan [6] on the same subject. Some 10 years later Erdős and Offord [7] proved that the mean number of real roots of a random polynomial of degree $n$ with coefficients $\pm 1$ is again $(2/\pi)\log n$. This leads us to the following question: can we find sequences $(\alpha_i)_{i \in \mathbb{N}}$ with coefficients $\pm 1$ such that the corresponding polynomials $\sum_{i=0}^{n} \alpha_i X^i$ have $O(\log n)$ real roots, and are these sequences random in some sense?

We introduce generalized Thue–Morse sequences whose corresponding polynomials of large degree $n$ have at least $C\log n$ real roots, where $C$ is an explicit positive constant. Finally, we discuss the spectral measure of these sequences.

1. Introduction. Erdős and Offord [7] established that the average number of real zeros of the degree $n$ polynomial $\sum_{i=0}^{n} \pm X^i$ is equivalent to $(2/\pi)\log n$. A natural question could then be to find a sequence $(\alpha_i)_{i \in \mathbb{N}}$ of $\pm 1$ such that

$$
\frac{1}{N} \sum_{n=0}^{N-1} \varrho(g_n) \sim \frac{2}{\pi} \log N
$$

where $\varrho(g_n)$ is the number of real zeros of the polynomial

$$
g_n(X) = \sum_{i=0}^{n} \alpha_i X^i.
$$

In a previous article [5] we tested the Thue–Morse sequence $(\varepsilon_i)_{i \in \mathbb{N}}$, defined by $\varepsilon_i = (-1)^{\nu(i)}$ where $\nu(i)$ is the sum of the binary digits of $i$.  

2000 Mathematics Subject Classification: Primary 12D10.

Key words and phrases: real roots of polynomials, Thue–Morse sequence, spectral measure.
Unfortunately, we proved that (1) does not hold in this case. More precisely we have
\[
\frac{1}{N} \sum_{n=0}^{N-1} \varphi(f_n) \xrightarrow{N \to \infty} \frac{11}{4},
\]
with \( f_n(X) = \sum_{i=0}^{n} \varepsilon_i X^i \).

In this paper we show the existence of families \((\varepsilon_{w,i})_{i \in \mathbb{N}}\) of \((\pm 1)\)-sequences for which
\[
\liminf_{N \to \infty} \frac{1}{N \log N} \sum_{n=0}^{N-1} \varphi(f_{w,n}) > 0,
\]
where \( f_{w,n}(X) = \sum_{i=0}^{n} \varepsilon_{w,i} X^i \). These sequences have, to some extent, a similar structure to the Thue–Morse sequence. They can be obtained in a very similar way, by means of iterations of morphisms. We call them **generalized Thue–Morse sequences**.

Before explaining this we introduce two useful notations. Let \( w \) be a word of length \( \ell \) on the alphabet \{+, -\} and \( i \leq j \) be integers less than \( \ell \). Then \( \iota_{w,j} \) represents the factor of \( w \) beginning at letter \( i \) and finishing at letter \( j \) of \( w \). For instance \( \iota_{w,i} \) is simply the letter at position \( i \) of \( w \) and \( \iota_{w_{\ell-1}} = w \). We put also \( w[i] = 1 \) if \( \iota_{w,i} = + \) and \( w[i] = -1 \) if \( \iota_{w,i} = - \).

Now let \( \varphi \) be the morphism on the alphabet \{+, -\} defined by
\[
\varphi : \{+ \to +, - \to -.\}
\]
The first iterations of \( \varphi \) are
\[
\varphi(+) = +,-,
\varphi^2(+) = +-+-,
\varphi^3(+) = +-+-+-+-+.
\]
Let
\[
\mathcal{E} = \lim_{n \to \infty} \varphi^n(+)\]
be the Thue–Morse word. There exists an obvious link between \( \mathcal{E} \) and \((\varepsilon_i)_{i \in \mathbb{N}}\), i.e.
\[
\varepsilon_i = \mathcal{E}[i].
\]
In the next section we slightly modify the definition of \( \varphi \) to get a wider family of sequences.

2. **Generalized Thue–Morse sequences**. Let \( w \) be a word on \{+, -\} of length \( \ell \geq 2 \) beginning with +. We put
\[
\varphi_w : \{+ \to w, - \to \overline{w}\}
where $\overline{w}$, the “opposite” of $w$, is defined by

$$i \overline{w} = \begin{cases} - & \text{if } i w_i = +, \\ + & \text{if } i w_i = - . \end{cases}$$

The Thue–Morse $w$-word, $E_w$, is then

$$E_w = \lim_{n \to \infty} \varphi^n_w (+).$$

Denote by $(\varepsilon_{w,i})_{i \in \mathbb{N}}$ the corresponding Thue–Morse $w$-sequence with coefficients $\pm 1$, that is, the sequence satisfying

$$\varepsilon_{w,i} = E_w[i].$$

For example if $w = ++$ then

$$\sum_{i=0}^{\infty} \varepsilon_{w,i} X^i = \frac{1}{1 - X}.$$

For each word $w$ of length $\ell$ we consider $P_w$ its associated polynomial defined by

$$P_w(X) = \sum_{j=0}^{\ell - 1} w[j] X^j.$$

**Lemma 1.** Let $w$ be a word of length $\ell \geq 2$ on $\{+, -\}$ beginning with $+$ and $\varepsilon_w$ its associated generalized Thue–Morse sequence. Then

$$\sum_{i=0}^{\infty} \varepsilon_{w,i} X^i = \prod_{h=0}^{\infty} P_w(X^{\ell h}).$$

**Proof.** Let $v$ be some word of length $t$ on $\{+, -\}$ and

$$P_v(X) = \sum_{j=0}^{t - 1} v[j] X^j$$

its associated polynomial. It suffices to see that

$$P_{\varphi_v(v)}(X) = P_w(X) P_v(X^{\ell}).$$

Indeed,

$$j \ell \left( \varphi_v(v) \right)_{j \ell + t - 1} = \begin{cases} w & \text{if } v[j] = 1, \\ \overline{w} & \text{if } v[j] = -1, \end{cases}$$

for all $j$ in $[0, t - 1]$. So we get

$$P_{\varphi_v(v)}(X) = \sum_{j=0}^{t - 1} P_w(X) v[j] X^{j \ell}$$

and $P_{\varphi_v(v)}(X) = P_w(X) P_v(X^{\ell})$ as claimed. The lemma immediately follows. Note that the series converges in $[-1, 1]$. \contradiction
Using this we can give for $\varepsilon_{w,j}$ a more explicit meaning that generalizes the initial definition of the Thue–Morse sequence. Let $m_1, \ldots, m_q$ be the integers $j \in [0, \ell - 1]$ such that $j w_j = -$. Then from (2) it is clear that

$$
\varepsilon_{w,i} = (-1)^{\nu_{m_k}(i) + \ldots + \nu_{m_q}(i)}
$$

where $\nu_{m_k}(i)$ represents the number of $m_k$'s in the base expansion of $i$.

3. Real roots of generalized Thue–Morse polynomials. For the classical Thue–Morse polynomials the starting word is $w = +-$ so that its associated polynomial is $P_w(X) = 1 - X$. Therefore $\ell = 2$ and since the real roots of $P_w(X^{2^i})$ are $-1$ and $1$, we cannot use the convergence of $\sum_{i=0}^{\infty} \varepsilon_i X^i$ on $]-1, 1[$. Now the starting polynomial $P_w$ may vanish on $]0, 1[$ and the real roots of $P_w(X^{\ell^h})$ spread in this case along $[0, 1]$. Using the convergence of (2) on $]-1, 1[$ we show that $g(f_{w,n})$ is in $C \log n$. Let us make this more precise.

**Theorem 1.** Let $w$ be a word of length $\ell \geq 2$ on $\{+, -\}$ beginning with $+$ such that $P_w$ has only simple roots on $]-1, 1[$, say $t$ roots $\beta_1 < \ldots < \beta_t$ in $]0, 1[$ and $t'$ roots $\beta_{t+t'} < \ldots < \beta_{t+2} < \beta_{t+1}$ in $]-1, 0[$. Let $f_{w,n}$ be the generalized Thue–Morse polynomials associated with $\varepsilon_w$. Assume that $\beta_t < \beta_1^{1/t}$. Suppose in addition that $\beta_{t+t'}^{1/t'} < \beta_{t+1}$ if $\ell$ is odd. Then there exists $K > 0$ such that for all $\epsilon > 0$ there is an $N(\epsilon)$ such that for all $n \geq N(\epsilon)$ we have

$$
g(f_{w,n}) \geq \frac{2(1 - \epsilon)t \log n}{\log \ell} - K \quad \text{if } \ell \text{ is even},
$$

$$
g(f_{w,n}) \geq \frac{(1 - \epsilon)(t + t') \log n}{\log \ell} - K \quad \text{if } \ell \text{ is odd}.
$$

**Proof.** Put $\beta_{j,h} = \beta_j^{1/\ell^h}$ for all $h \in \mathbb{Z}$. The roots of

$$
f_{w,\ell^{t-1}}(X) = \prod_{h=0}^{t-1} P_w(X^{\ell^h})
$$

in $]0, 1[$ are therefore

$$
\beta_{1,0} \quad \ldots \quad \beta_{t,0},
$$

$$
\vdots
$$

$$
\beta_{1,h} \quad \ldots \quad \beta_{t,h},
$$

$$
\vdots
$$

$$
\beta_{1,k-1} \quad \ldots \quad \beta_{t,k-1}.
$$

Put also

$$
\delta_{j,h} = \begin{cases} 
\sqrt{\beta_{j,h} \beta_{j+1,h}} & \text{for } j \in [1, t - 1], h \in \mathbb{Z}, \\
\sqrt{\beta_{j,h} \beta_{j+1,h}} & \text{for } j = t, h \in \mathbb{Z}.
\end{cases}
$$
The next lemma plays an important part in the following.

**Lemma 2.** Let \( a \) be the multiplicity of 1 as a root of \( P_w \). Then there are two constants \( C_1 \) and \( C_2 \) such that for all large \( h, k \geq h \) and \( j \in [1, t] \),

\[
|f_{w, t^k - 1}(\delta_{j,h})| \geq C_1 C_2^{-h} e^{-ah(h+1)/2}.
\]

**Proof of Lemma 2.** First of all we determine \( C_1(j) \) and \( C_2(j) \) for any \( j \) in \([1, t]\). Since there are only finitely many \( j \)'s the lemma will follow immediately.

We remark that

\[
f_{w, t^k - 1}(\delta_{j,h}) = \prod_{s=0}^{k-1} P_w(\delta_{s,h}^t) = \prod_{i=h-k+1}^h P_w(\delta_{j,i}).
\]

Now

\[
P_w(\delta_{j,i}) = \frac{P_w(1)}{a!} (1 - \delta_{j,i})^a + o(1 - \delta_{j,i})^a
\]

when \( 1 - \delta_{j,i} \to 0 \), so there are \( C_5 \in \mathbb{R}_+ \) and \( i_0 > 0 \) such that for all \( i > i_0 \),

\[
|P_w(\delta_{j,i})| \geq C_5 |1 - \delta_{j,i}|^a \geq C_5 \left| \frac{\log \delta_{j,0}}{2h} \right|^u
\]

since \( |1 - e^x| \geq |x|/2 \) near 0. Thus

\[
\prod_{i=i_0}^h |P_w(\delta_{j,i})| \geq C_4 C_3^h e^{-ah(h+1)/2}.
\]

The factor

\[
\prod_{i=h-k+1}^{i_0-1} |P_w(\delta_{j,i})|
\]

leads us to study the behaviour of \( P_w \) near 0. Now \( P_w \) is locally either greater than 1 or less than 1. In the first case it is obvious that for a suitable constant \( C_1 \),

\[
\prod_{i=h-k+1}^h |P_w(\delta_{j,i})| \geq C_1 C_2^{-h} e^{-ah(h+1)/2}.
\]

In the second case \( \delta_{j,i} \geq \delta_{j,i}^2 \geq \ldots \geq \delta_{j,i}^{\ell-1} \) and these quantities are small in comparison with 1 for all large \(|i|\). Thus

\[
|P_w(\delta_{j,i}) - 1| \leq (\ell - 1) \delta_{j,i}.
\]

Since for \( i \leq 0 \) we have \( \delta_{j,i} = \delta_{j,0}^{\ell i} \), we obtain the convergence of

\[
\prod_{i=-\infty}^{i_0-1} |P_w(\delta_{j,i})|.
\]

The lemma is then proved. \( \blacksquare \)
Remark. When $P_w(1) \neq 0$ it is possible to replace $C_1 C_2^h \ell^{-wh(h+1)/2}$ in (3) by a positive constant independent of $h$.

Since $f_{w, \ell^k - 1}$ has $t-1$ simple roots between $\delta_{1,h}$ and $\delta_{t,h}$, it follows that $f_{w, \ell^k - 1}$ changes sign $t-1$ times, passing above and below the lines
\[ y = C_1 C_2^h \ell^{-wh(h+1)/2} \quad \text{and} \quad y = -C_1 C_2^h \ell^{-wh(h+1)/2}. \]

For example, consider Figure 1 which displays $f_{w, 728}$ built from the word $w = + -$. As $P_w(1) = -1$ the remark ensures that $f_{w, 728}$ and more generally $f_{w, 3^k - 1}$ winds itself round two absolute axes. Here they are $y = 0.067130$ and $y = -0.067130$ (bold lines on Figure 1).

![Figure 1. $f_{w, 728}(x)$ on $[0, 1]$](image)

Now if $n \geq \ell^k$ it is clear that on $]0,1[$,
\[ |f_{w,n}(x) - f_{w,\ell^k - 1}(x)| \leq \frac{x^{\ell^k}}{1-x}. \]

Let $\epsilon > 0$. Since $\delta_{t,[[1-\epsilon)k]} = \delta_{t,[[1-\epsilon)k]} - k \leq \delta_t^{\ell^k}$, we have
\[ \frac{\delta_{t,[[1-\epsilon)k]} - \ell^{ak(k+1)/2}}{1 - \delta_{t,[[1-\epsilon)k]} - C_1 C_2^{\ell^k}} \xrightarrow{k \to \infty} 0. \]

So there is $k(\epsilon)$ such that for all $k \geq k(\epsilon)$,
\[ \delta_{t,[[1-\epsilon)k]} - C_1 C_2^{\ell^k} \ell^{-wh(k+1)/2}(1 - \delta_{t,[[1-\epsilon)k]} - \ell^{ak(k+1)/2}(1-x) \leq 0. \]

Moreover the function $x^{\ell^k} - C_1 C_2^{\ell^k} \ell^{-wh(k+1)/2}(1-x)$ is increasing on $]0,1[$.
So on $]0, \delta_{\ell, [(1-\epsilon)k]}]$,
\[ \frac{x^\ell}{1-x} < C_1 C_2 \epsilon \ell^{-nk(k+1)/2}. \]
This inequality ensures that $f_{w,n}$ is subject to the same oscillations as $f_{w,\ell^{k-1}}$ provided Lemma 2 holds. Indeed, for $h$ and $k$ large,
\[ f_{w,n}(\delta_j, h) \geq f_{w,\ell^{k-1}}(\delta_j, h) - \frac{x^\ell}{1-x} > 0 \]
when $f_{w,\ell^{k-1}}(\delta_j, h) > 0$ and
\[ f_{w,n}(\delta_j, h) \leq f_{w,\ell^{k-1}}(\delta_j, h) + \frac{x^\ell}{1-x} < 0 \]
when $f_{w,\ell^{k-1}}(\delta_j, h) < 0$.

Let $g_1(f_{w,n})$ be the number of real roots of $f_{w,n}$ in $]0, 1[$. So
\[ g_1(f_{w,n}) \geq \frac{(1-\epsilon)\ell \log n}{\log \ell} - K_1, \]
for a suitable absolute constant $K_1$, as soon as $n \geq N(\epsilon) = \ell^k(\epsilon)$.

If $\ell$ is even then $f_{w,n}$ has at least $[(1-\epsilon)k]\ell$ roots in $]-1, 0[$ for large $k$ and $n \geq \ell^k$. Then
\[ g(f_{w,n}) \geq \frac{2(1-\epsilon)\ell \log n}{\log \ell} - K, \]
for all large $n$.

If $\ell$ is odd we apply what we have just seen to $P_{w}(-X)$. This polynomial has by hypothesis $\ell'$ roots in $]0, 1[$ so that $f_{w,n}(-X)$ has at least $[(1-\epsilon)k]\ell'$ roots in $]0, 1[$. Thus
\[ g(f_{w,n}) \geq \frac{(1-\epsilon)(\ell + \ell') \log n}{\log \ell} - K, \]
for all large $n$.

This ends the proof of Theorem 1. \(\blacksquare\)

Under the assumptions of Theorem 1 we have the following inequalities.

**Corollary.** If $\ell$ is even then
\[ \liminf_{N \to \infty} \frac{1}{N \log N} \sum_{n=0}^{N-1} g(f_{w,n}) \geq \frac{2\ell}{\log \ell}. \]
If $\ell$ is odd then
\[ \liminf_{N \to \infty} \frac{1}{N \log N} \sum_{n=0}^{N-1} g(f_{w,n}) \geq \frac{\ell + \ell'}{\log \ell}. \]
Proof. Assuming \( \ell \) to be even, we apply Theorem 1 to obtain
\[
\frac{1}{N \log N} \sum_{n=0}^{N-1} \varrho(f_{w,n}) \geq \frac{1}{N \log N} \cdot \frac{2(1-\epsilon)\ell}{\log \ell} \left( \sum_{n=N(\epsilon)}^{N-1} \log n - N K \right) \\
\geq \frac{2(1-\epsilon)\ell}{\log \ell} - \frac{K'}{\log N}.
\]
When \( N \) tends to infinity we get
\[
\liminf_{N \to \infty} \frac{1}{N \log N} \sum_{n=0}^{N-1} \varrho(f_{w,n}) \geq \frac{2(1-\epsilon)\ell}{\log \ell}.
\]
Since this is true for all \( \epsilon > 0 \) we deduce that
\[
\liminf_{N \to \infty} \frac{1}{N \log N} \sum_{n=0}^{N-1} \varrho(f_{w,n}) \geq \frac{2\ell}{\log \ell}.
\]
If \( \ell \) is odd the same argument works. ■

So generalized Thue–Morse sequences yield polynomials with many real roots; but what can we say on their random behaviour? In the next section we say a few words about this.

A good tool to evaluate the random nature of a sequence is to study its spectral measure defined for instance in [11] and in [1]. We recall basic definitions, and then we give just the important results without all intermediate steps.

Let \( \gamma_w(h) \) be the correlation function of \((\varepsilon_{w,n})_{n \in \mathbb{N}}\) defined by
\[
\gamma_w(h) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \varepsilon_{w,n} \varepsilon_{w,n+h}.
\]
This limit always exists for generalized Thue–Morse sequences. The spectral measure \( d\sigma_w \) is linked to \( \gamma_w(h) \) by the formula
\[
\gamma_w(h) = \int_0^1 e^{2\pi i x h} d\sigma_w(x).
\]
For all \( i \in [1, \ell-1] \), \( \gamma_w \) satisfies the recurrence relations
\[
\gamma_w(\ell k) = \gamma_w(k), \\
\gamma_w(\ell k + i) = \frac{a_{w,i}}{\ell} \gamma_w(k) + \frac{a_{w,\ell-i}}{\ell} \gamma_w(k + 1),
\]
where \( a_{w,i} = \sum_{j=0}^{\ell-1} w[j] w[i+j] \).
To establish that $d\sigma_w$ is continuous we know [1] that it is sufficient to prove that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \gamma_w(n)^2 = 0.$$ 

Let

$$I_w(h) = \lim_{N\to\infty} \frac{1}{N} \sum_{m<N} \gamma_w(m)\gamma_w(m+h),$$

and we are left to show $I_w(0) = 0$. Following the ideas of [1, Appendix I] we obtain the system

$$S: \begin{cases} e_{11}(w)I_w(0) + c_{12}(w)I_w(1) = 0, \\ e_{21}(w)I_w(0) + e_{22}(w)I_w(1) = 0, \end{cases}$$

where

$$e_{11}(w) = 1 - \frac{1}{\ell} - 2\sum_{j=1}^{\ell-1} \frac{a_{w,j}}{\ell^3},$$

$$e_{12}(w) = -2\sum_{j=1}^{\ell-1} \frac{a_{w,j}a_{w,\ell-j}}{\ell^3},$$

$$e_{21}(w) = \frac{2a_{w,1}}{\ell^2} + 2\sum_{j=1}^{\ell-2} \frac{a_{w,j}a_{w,j+1}}{\ell^3},$$

$$e_{22}(w) = -1 + \frac{2a_{w,\ell-1}}{\ell^2} + \sum_{j=1}^{\ell-2} \frac{a_{w,j+1}a_{w,\ell-j} + a_{w,j}a_{w,\ell-j-1}}{\ell^3}.$$
Since
\[ \sum_{j=1}^{\ell-2} \frac{a_{w,j+1}a_{w,\ell-j} + a_{w,j}a_{w,\ell-j-1}}{\ell^3} \leq \sum_{j=1}^{\ell-2} \frac{(\ell - j - 1)j + (\ell - j)(j + 1)}{\ell^3} \leq \frac{\ell^3 - 4\ell}{3\ell^3}, \]
we have
\[ |c_{22}(w)| \geq 1 - \frac{2|a_{w,\ell-1}|}{\ell^2} - \frac{\ell^3 - 4\ell}{3\ell^3} \geq \frac{2\ell^3 - 2\ell}{3\ell^3}. \]
We also notice that the last inequality is an equality if and only if
\[ a_{w,j} = \ell - j \quad \text{for} \quad j \in [\ell, \ell - 1], \]
\[ a_{w,\ell-1} = 1. \]
If (4) is not satisfied, then \( |c_{22}(w)| > (2\ell^3 - 2\ell)/(3\ell^3) \) and
\[ |\Delta(w)| \geq |c_{11}(w)c_{22}(w)| - |c_{12}(w)c_{21}(w)| > 0, \]
which implies that \( I_w(0) = 0 \). Now if (4) holds then \( w[j]w[j + 1] \) does not depend on \( j \in [\ell, \ell - 2] \). Therefore \( w[j] = 1 \) or \( w[j] = (-1)^j \) for all \( j \in [0, \ell - 1] \). When \( w[j] = (-1)^j \) the relation \( a_{w,\ell-1} = w[0]w[\ell - 1] = 1 \) shows that \( \ell \) is necessarily odd. Obviously if \( w \) is of type ++ or -- then \( \Delta(w) = 0 \), so that the result is proved.

This lemma ensures that the spectral measure of a generalized Thue-Morse sequence is continuous, except for trivial ++ or -- cases. However, although continuous, \( d\sigma_w \) is singular (see Theorem 6 of [4]). Therefore \( d\sigma_w \) is not absolutely continuous, which would be a true random behaviour. Nonetheless this ensures that \( (\varepsilon_{w,n})_{n\in\mathbb{N}} \) is pseudo-random in the sense of Bass [2] and Bertrandias [3].

Finally, when \( w \) is of type ++ its spectral measure is the Dirac mass \( \delta_0(x) \). If \( w \) is of type --, then \( d\sigma_w \) is \( \delta_{1/2}(x) \).

Acknowledgements. I am grateful to my supervisor, L. Habsieger, for his continuous help and to M. Mendès France for his helpful comments.

References

Real roots of Thue–Morse polynomials


Laboratoire d’Algorithmique Arithmétique
Université Bordeaux 1
351, cours de la Libération
F-33405 Talence Cedex, France
E-mail: cdoche@math.u-bordeaux.fr

Received on 8.10.1999
and in revised form on 18.1.2001