Some properties of a Rudin–Shapiro-like sequence

Narad Rampersad

(Joint work with Philip Lafrance and Randy Yee)

Department of Mathematics and Statistics
University of Winnipeg
Digital sequences

- there are several well-studied sequences whose $n$-th term is defined based on some property of the digits of $n$ when written in some chosen base

- e.g., the sum-of-digits function $s_k(n) =$ the sum of the digits of the base-$k$ representation of $n$

- e.g., $s_2(n)$ counts the number of 1’s in the binary representation of $n$

- $(t_n)_{n \geq 0} = ((-1)^{s_2(n)})_{n \geq 0}$ is the Thue–Morse sequence

$$+1 \ -1 \ -1 \ +1 \ -1 \ +1 \ +1 \ -1 \ \cdots$$
The Rudin–Shapiro sequence

- let $e_{2;11}(n)$ denote the number of occurrences of 11 in the binary representation of $n$
- e.g., $e_{2;11}(235) = 3$ since $235 = [11101011]_2$
- $(r_n)_{n \geq 0} = \left((-1)^{e_{2;11}(n)}\right)_{n \geq 0}$ is the Rudin–Shapiro sequence

\[ +1 \quad +1 \quad +1 \quad -1 \quad +1 \quad +1 \quad -1 \quad +1 \quad \cdots \]
Scattered subsequences

- instead of counting occurrences of a given block, we can count occurrences of a given pattern as a scattered subsequence in the digital representation of $n$
- e.g., $302 = [100101110]_2$ has 13 occurrences of 010 as scattered subsequences of its binary representation
- an occurrence of 10 as a scattered subsequence is called an inversion
- in general, an inversion in a word is an occurrence of $ba$ as a scattered subsequence, where $b > a$
Counting inversions

- write $\text{inv}_2(n)$ to denote the number of inversions in the binary representation of $n$
- define the sequence $(i_n)_{n \geq 0}$ by $i_n = (-1)^{\text{inv}_2(n)}$, so

$$(i_n)_{n \geq 0} = +1 \quad +1 \quad -1 \quad +1 \quad +1 \quad -1 \quad +1 \quad +1 \quad \cdots$$

- this sequence has many similarities with the Rudin–Shapiro sequence
Generalized Rudin–Shapiro sequences

- Allouche and Liardet (1991) studied generalizations of the Rudin–Shapiro sequence
- Fix $ab \neq 00$ and fix some positive integer $d$
- Define $(u_n)_{n \geq 0}$ such that $u_n$ equals the number of occurrences of $ab$ as a scattered subsequence of the binary representation of $n$, where $a$ and $b$ occur at distance $d + 1$ from each other.
- Taking $ab = 10$ counts the number of inversions where the inverted elements are separated by distance $d + 1$. 
Representation as an automatic sequence

\((i_n)_{n \geq 0}\) is a 2-automatic sequence
Operation of the automaton

The automaton calculates \( i_n \) as follows: the binary digits of \( n \) are processed from most significant to least significant, and when the last digit is read, the automaton halts in the state

\[
\begin{pmatrix}
(-1)^{s_2(n)} \\
(-1)^{\text{inv}_2(n)}
\end{pmatrix}.
\]

\( i_n \) is given by the lower component of the label of the state reached after reading the binary representation of \( n \) (the first component has the value \( t_n \)).
(i_n)_{n \geq 0}$ can be generated by iterating the morphism

$$A \rightarrow AB, \quad B \rightarrow CA, \quad C \rightarrow BD, \quad D \rightarrow DC,$$

to obtain the infinite sequence

$$ABCABDABCDABCDABCA \cdots$$

and then applying the recoding

$$A, B \rightarrow +1, \quad C, D \rightarrow -1.$$
Rudin–Shapiro morphism

cf. the Rudin–Shapiro sequence, which is obtained by iterating

\[ A \rightarrow AB, \quad B \rightarrow AC, \quad C \rightarrow DB, \quad D \rightarrow DC, \]

and then applying the same recoding as above.
Recurrence relations

\[ i_{2n} = i_n t_n \] \hspace{1cm} (1)
\[ i_{2n+1} = i_n \] \hspace{1cm} (2)

- let \( w \) denote the binary representation of \( n \)
- the number of 10’s in \( w0 \) equals the number of 10’s in \( w \) plus the number of 1’s in \( w \)

\[ i_{2n} = (-1)^{\text{inv}_2(n) + s_2(n)} = i_n t_n. \]

- \( i_{2n+1} = i_n \) is clear, since appending a 1 to \( w \) does not change the number of 10’s
Recurrence relations

\[ i_{4n} = i_n \]
\[ i_{4n+1} = i_{2n} \]
\[ i_{4n+2} = -i_{2n} \]
\[ i_{4n+3} = i_n. \]
Proving the relations

The Thue–Morse sequence satisfies

\[ t_{2n} = t_n \quad \text{and} \quad t_{2n+1} = -t_n. \]

Now we have

\[ i_{4n} = i_{2n} t_{2n} = i_{2n} t_n = i_n t_n t_n = i_n, \]

where we have applied (1) twice. Similarly, we get

\[ i_{4n+1} = i_{2(2n)+1} = i_{2n} \]

by applying (2).
Proving the relations

Next, we calculate

\[ i_{4n+2} = i_{2(2n+1)} = i_{2n+1}t_{2n+1} = i_n(-t_n) = -i_{2n}, \]

and finally,

\[ i_{4n+3} = i_{2(2n+1)+1} = i_{2n+1} = i_n. \]
The summatory function of \((t_{3n})_{n \geq 0}\)

Newman (1969) and Coquet (1983) studied the summatory function of the Thue–Morse sequence taken at multiples of 3. In particular,

\[
\sum_{0 \leq n < N} t_{3n} = N \log_4 3 G_0(\log_4 N) + \frac{1}{3} \eta(N),
\]

where \(G_0\) is a bounded, continuous, nowhere differentiable, periodic function with period 1, and

\[
\eta(N) = \begin{cases} 
0 & \text{if } N \text{ is even}, \\
(-1)^{s_2(3N-1)} & \text{if } N \text{ is odd}.
\end{cases}
\]
Brillhart, Erdős, and Morton (1983) and Dumont and Thomas (1989) studied the summatory function of the Rudin–Shapiro sequence. In this case,

$$\sum_{0 \leq n < N} r_n = \sqrt{N} G_1(\log_4 N)$$

where again $G_1$ is a bounded, continuous, nowhere differentiable, periodic function with period 1.
Summing the inversions sequence

Define the summatory function \( S(N) \) of \((i_n)_{n \geq 0}\) as

\[
S(N) = \sum_{0 \leq n \leq N} i_n.
\]

The first few values of \( S(N) \) are:

<table>
<thead>
<tr>
<th>(N)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S(N))</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
A plot of the function $S(N)$

The smooth curves are plots of $\sqrt{2}\sqrt{N}$ and $(\sqrt{3}/3)\sqrt{N}$. 
The growth of $S(N)$

**Theorem (Lafrance, R., Yee 2014)**

There exists a bounded, continuous, nowhere differentiable, periodic function $G$ with period 1 such that

$$S(N) = \sqrt{NG(\log_4 N)}.$$

This can be obtained by a criterion from the book of Allouche and Shallit derived from techniques of Tenenbaum (1997).
A plot of the periodic function $G$
Upper and lower limits of oscillation

Theorem (Lafrance, R., Yee 2014)

We have

$$\liminf_{n \to \infty} \frac{S(n)}{\sqrt{n}} = \frac{\sqrt{3}}{3} \quad \text{and} \quad \limsup_{n \to \infty} \frac{S(n)}{\sqrt{n}} = \sqrt{2}.$$
Combinatorial properties of \((i_n)_{n \geq 0}\)

**Theorem (Lafrance, R., Yee 2014)**

The sequence \((i_n)_{n \geq 0}\) contains

1. no 5-th powers,
2. cubes \(x^3\) exactly when \(|x| = 3\),
3. squares \(xx\) exactly when \(|x| \in \{1, 2\} \cup \{3 \cdot 2^k : k \geq 0\}\).
4. arbitrarily long palindromes.

These results were verified by Jeffrey Shallit and Hamoon Mousavi using their automated prover for properties of automatic sequences.
Automaton for period lengths of cubes in \((i_n)_{n \geq 0}\)
Automaton for period lengths of squares in $(i_n)_{n \geq 0}$
Automaton for lengths of palindromes in $(i_n)_{n \geq 0}$
The “square root” property

The Rudin–Shapiro sequence satisfies the following: 

There exists a constant $C$ such that for all $N \geq 0$

$$\sup_{\theta \in \mathbb{R}} \left| \sum_{0 \leq n < N} r_n e^{2\pi in\theta} \right| \leq C \sqrt{N}.$$ 

It would seem that (Allouche, personal communication) the sequence $(i_n)_{n \geq 0}$ does not have this property.
The End