J. Math. Anal. Appl. • • • (• • • •) • • •–• • •

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

On the abelian complexity of the Rudin–Shapiro sequence ☆

Xiaotao Lü [a], Jin Chen [b], Zhixiong Wen [a], Wen Wu [c,d,*]

[a] *School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, PR China*
[b] *College of Science, Huazhong Agricultural University, Wuhan 430070, PR China*
[c] *School of Mathematics, South China University of Technology, Guangzhou 510641, PR China*
[d] *Mathematics, P.O. Box 3000, 90014 University of Oulu, Finland*

A R T I C L E I N F O

Article history:
Received 22 June 2016
Available online xxxx
Submitted by B. Bongiorno

Keywords:
Rudin–Shapiro sequence
Abelian complexity
k-Regular sequence
Automatic sequence
Box dimension

A B S T R A C T

In this paper, we study the abelian complexity of the Rudin–Shapiro sequence and a related sequence. We show that these two sequences share the same complexity function $\rho(n)$, which satisfies certain recurrence relations. As a consequence, the abelian complexity function is 2-regular. Further, we prove that the box dimension of the graph of the asymptotic function $\lambda(x)$ is $3/2$, where $\lambda(x) = \lim_{k\to\infty} \rho(4^k x)/\sqrt{4^k x}$ and $\rho(x) = \rho(\lfloor x \rfloor)$ for every $x > 0$.

1. Introduction

The abelian complexity of infinite words has been examined by Coven and Hedlund in [6] as an alternative way to characterize periodic sequences and Sturmian sequences. Richomme, Saari, and Zamboni introduced this notion formally in [11], which initiated a general study of the abelian complexity of infinite words over finite alphabets. For example, the abelian complexity functions of some notable sequences, such as the Thue–Morse sequence and all Sturmian sequences, were studied in [11] and [6], respectively. There are also many other works devoted to this subject; see [3,9,7,10] and references therein. In the following, we shall give the definition of abelian complexity.

Let $\mathbf{w} = w(0)w(1)w(2)\cdots$ be an infinite sequence over a finite alphabet $\mathcal{A}$. Let $\mathcal{F}_{\mathbf{w}}(n)$ denote the set of all factors of $\mathbf{w}$ of length n, i.e.,

☆ This work was supported by the Fundamental Research Funds for the Central Universities (Project No. 2662015QD016) and NSFC (Nos. 11401188, 11371156, 11431007 and 11626110). The research of Wen Wu was also partially supported by the Academy of Finland, the Centre of Excellence in Analysis and Dynamic Research.

* Corresponding author at: School of Mathematics, South China University of Technology, Guangzhou 510641, PR China.
E-mail addresses: M201270021@hust.edu.cn (X. Lü), wind.golden@gmail.com (J. Chen), zhi-xiong.wen@mail.hust.edu.cn (Z. Wen), hust.wuwen@gmail.com (W. Wu).

http://dx.doi.org/10.1016/j.jmaa.2017.02.019

$$\mathcal{F}_{\mathbf{w}}(n) := \{w(i)w(i+1)\cdots w(i+n-1) : i \geq 0\}.$$

Two finite words u, v over the same alphabet $\mathcal{A}$ are *abelian equivalent* if $|u|_a = |v|_a$ for every letter $a \in \mathcal{A}$. Abelian equivalency induces an equivalence relation, denoted by $\sim_{ab}$. Now we are ready to state the definition of abelian complexity.

Definition 1. The *abelian complexity function* $\rho_{\mathbf{w}} : \mathbb{N} \to \mathbb{N}$ of $\mathbf{w}$ is defined by

$$\rho_{\mathbf{w}}(n) := \#(\mathcal{F}_{\mathbf{w}}(n)/\sim_{ab}).$$

The first part of this paper is devoted to the study of the regularity of the abelian complexity of the Rudin–Shapiro sequence $\mathbf{r} = r(0)r(1)r(2)\cdots$, whose generating function $R(z) := \sum_{n\geq 0} r(n)z^n$ satisfies the Mahler-type functional equation

$$R(z) + R(-z) = 2R(z^2).$$

Let $\mathbf{r}'$ denote the coefficient sequence of $R(-z)$. To state our results, we shall recall the definitions of k-regular and k-automatic sequences. For more details, see [2].

Definition 2. Let $k \geq 2$ be an integer. The *k-kernel* of an infinite sequence $\mathbf{w} = (w(n))_{n\geq 0}$ is the set of subsequences

$$\mathbf{K}_k(\mathbf{w}) := \{(w(k^e n + c))_{n\geq 0} \mid e \geq 0, 0 \leq c < k^e\}.$$

The sequence $\mathbf{w}$ is *k-automatic* if $\mathbf{K}_k(\mathbf{w})$ is finite. If the $\mathbb{Z}$-module generated by its k-kernel is finitely generated, then $\mathbf{w} = (w(n))_{n\geq 0}$ is *k-regular*.

Now we state our first result.

Theorem A. *The abelian complexity of the Rudin–Shapiro sequence* $\mathbf{r}$*, which is the same as the abelian complexity of* $\mathbf{r}'$*, is* 2*-regular.*

In the second part, inspired by the work of Brillhart, Erdős and Morton [4], we study the limit function

$$\lambda(x) := \lim_{k\to\infty} \frac{\rho(4^k x)}{\sqrt{4^k x}},$$

where $\rho(x) := \rho(\lfloor x \rfloor)$ for every $x > 0$. The function λ is continuous and non-differentiable almost everywhere; for details, see [5]. Further, $\lambda(x)$ is self-similar in the sense that $\lambda(x) = \lambda(4x)$ for every $x > 0$. The graph of $\lambda(x)$ on the interval $[1, 4]$, which is illustrated in Fig. 1, has potential to be a fractal curve; and it is.

To introduce our next result, we shall recall the definition of box dimension. Let $\delta > 0$. For every $m_1, m_2 \in \mathbb{Z}$, we call the following square

$$[m_1\delta, (m_1+1)\delta] \times [m_2\delta, (m_2+1)\delta]$$

a *δ-mesh* of $\mathbb{R}^2$. Let $F \subset \mathbb{R}^2$ be a non-empty bounded set in $\mathbb{R}^2$, and $N_\delta(F)$ be the number of δ-meshes that intersect F. The *upper* and *lower box dimensions* are defined by

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

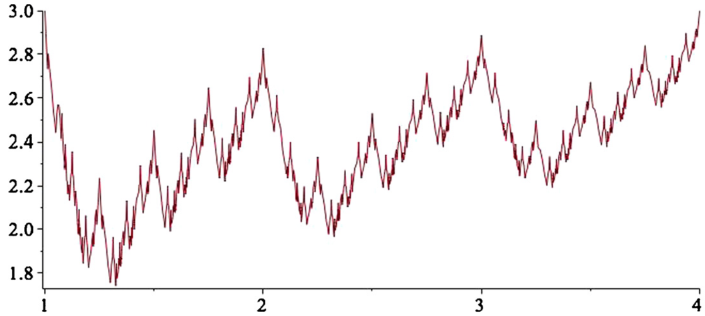

Fig. 1. The graph of $\lambda(x)$ for $x \in [1, 4]$.

$$\overline{\dim}_B F := \overline{\lim}_{\delta\to 0} \frac{\log N_\delta(F)}{-\log \delta} \text{ and } \underline{\dim}_B F := \underline{\lim}_{\delta\to 0} \frac{\log N_\delta(F)}{-\log \delta},$$

respectively. If $\overline{\dim}_B F = \underline{\dim}_B F$, then the common value is the *box dimension* of F. Let $\dim_B F$ denote the box dimension of F. For more details, see [8].

Theorem B. *The box dimension of the graph of* $\lambda(x)$ *on every subinterval of* $(0, +\infty)$ *is* $3/2$.

A variety of interesting fractals, both of theoretical and practical importance, occur as graphs of functions. Yue proved in [13] that the graph of one limit function studied in [4] also has box dimension 3/2. With probability 1, the graph of a one-dimensional Brownian sample function has Hausdorff dimension and box dimension 3/2; see [8, Theorem 16.4]. For every $b \geq 2$, the graph of the Weierstrass function $W(x) = \sum_{n=0}^{\infty} b^{-n/2} \cos(b^n x)$ has Hausdorff dimension and box dimension 3/2; see, for example, [8,12] and references therein. For the Hausdorff dimension of the graph of $\lambda(x)$, Theorem B poses a good candidate 3/2. It is natural to conjecture that the Hausdorff dimension of the graph of $\lambda(x)$ equals 3/2.

The outline of this paper is as follows. In Section 2, we state basic definitions and notation. In Section 3, we give recurrence relations for the abelian complexity functions of the sequences $\mathbf{r}$ and $\mathbf{r}'$. As a consequence, the abelian complexity function of the Rudin–Shapiro sequence is 2-regular, and the first difference of the abelian complexity function of the Rudin–Shapiro sequence is 2-automatic. In the last section, the box dimension of the graph of the function $\lambda(x)$ is studied.

2. Preliminary

In this section, we shall introduce some notation.

2.1. Finite and infinite words

An *alphabet* $\mathcal{A}$ is a finite and non-empty set (of symbols) whose elements are called *letters*. A (finite) *word* over the alphabet $\mathcal{A}$ is a concatenation of letters in $\mathcal{A}$. The concatenation of two words $u = u(0)u(1)\cdots u(m)$ and $v = v(0)v(1)\cdots v(n)$ is the word $uv = u(0)u(1)\cdots u(m)v(0)v(1)\cdots v(n)$. The set of all finite words over $\mathcal{A}$ including the *empty word* ε is denoted by $\mathcal{A}^*$. An infinite word $\mathbf{w}$ is an infinite sequence of letters in $\mathcal{A}$. The set of all infinite words over $\mathcal{A}$ is denoted by $\mathcal{A}^{\mathbb{N}}$.

The *length* of a finite word $w \in \mathcal{A}^*$, denoted by $|w|$, is the number of letters contained in w. We set $|\varepsilon| = 0$. For every word $u \in \mathcal{A}^*$ and every letter $a \in \mathcal{A}$, let $|u|_a$ denote the number of occurrences of a in u.

A finite word w is a *factor* of a finite (or an infinite) word v, written by $w \prec v$, if there exist a finite word x and a finite (or an infinite) word y such that $v = xwy$. When $x = \varepsilon$, w is called a *prefix* of v, denoted by $w \lhd v$; when $y = \varepsilon$, w is called a *suffix* of v, denoted by $w \rhd v$.

2.2. Digit sums

Now we assume that the alphabet $\mathcal{A}$ is composed of integers. Let $\mathbf{w} = w(0)w(1)w(2)\cdots \in \mathcal{A}^{\mathbb{N}}$ be an infinite word. For every $i \geq 0$ and $n \geq 1$, the sum of consecutive n letters in $\mathbf{w}$ starting from the position i is denoted by

$$\Sigma_{\mathbf{w}}(i, n) := \sum_{j=i}^{i+n-1} w(j).$$

The *maximal sum* and *minimal sum* of consecutive n ($n \geq 1$) letters in $\mathbf{w}$ are denoted by

$$M_{\mathbf{w}}(n) := \max_{i \geq 0} \Sigma_{\mathbf{w}}(i, n) \text{ and } m_{\mathbf{w}}(n) := \min_{i \geq 0} \Sigma_{\mathbf{w}}(i, n).$$

In addition, we always assume that $M_{\mathbf{w}}(0) = m_{\mathbf{w}}(0) = 0$.

Let DS denote the digit sum of a finite word $u = u(0)\cdots u(|u|-1) \in \mathcal{A}^*$, i.e.,

$$\mathrm{DS}(u) := \sum_{j=0}^{|u|-1} u(j).$$

Then

$$M_{\mathbf{w}}(n) = \max \left\{ \mathrm{DS}(v) : v \in \mathcal{F}_{\mathbf{w}}(n) \right\}$$

and

$$m_{\mathbf{w}}(n) = \min \left\{ \mathrm{DS}(v) : v \in \mathcal{F}_{\mathbf{w}}(n) \right\}.$$

The abelian complexity function of an infinite word $\mathbf{w}$ over $\{-1, 1\}$ is closely related to the digit sums of factors of $\mathbf{w}$.

Proposition 1. *Let* $\mathbf{w} \in \{-1, 1\}^{\mathbb{N}}$. *Then*

$$\rho_{\mathbf{w}}(n) = \frac{M_{\mathbf{w}}(n) - m_{\mathbf{w}}(n)}{2} + 1.$$

Proof. For a proof one can refer to [3, Proposition 2.2]. □

2.3. The Rudin–Shapiro sequence $\mathbf{r}$ *and a related sequence* $\mathbf{r}'$

The Rudin–Shapiro sequence

$$\mathbf{r} = r(0)r(1)\cdots r(n)\cdots \in \{-1, 1\}^{\mathbb{N}}$$

is given by the following recurrence relations:

$$r(0) = 1, \ r(2n) = r(n), \ r(2n+1) = (-1)^n r(n) \quad (n \geq 0). \tag{2.1}$$

The generating function $R(z) = \sum_{n \geq 0} r(n) z^n$ of the Rudin–Shapiro sequence satisfies the following Mahler-type functional equation

$$R(z) + R(-z) = 2R(z^2).$$

We also study the coefficient sequence of $R(-z)$, denoted by

$$\mathbf{r}' = r'(0)r'(1) \cdots \in \{-1, 1\}^{\mathbb{N}}.$$

Clearly, $r'(n) = (-1)^n r(n)$ for all $n \geq 0$. Thus

$$r'(0) = 1,\ r'(2n) = (-1)^n r'(n),\ r'(2n+1) = -r'(n) \quad (n \geq 0). \tag{2.2}$$

The Rudin–Shapiro sequence can also be generated by a substitution in the following way. Let $\sigma : \{a, b, c, d\} \to \{a, b, c, d\}^*$ and $\tau, \tau' : \{a, b, c, d\} \to \{-1, 1\}^*$, where

$$\begin{array}{rllll} \sigma: & a \mapsto ab, & b \mapsto ac, & c \mapsto db, & d \mapsto dc, \\ \tau: & a \mapsto 1, & b \mapsto 1, & c \mapsto -1, & d \mapsto -1, \\ \tau': & a \mapsto 1, & b \mapsto -1, & c \mapsto 1, & d \mapsto -1. \end{array}$$

Let $\mathbf{s} := \sigma^\infty(a)$ be the fixed point of σ beginning with a. Then

$$\mathbf{r} = \tau(\sigma^\infty(a)) \text{ and } \mathbf{r}' = \tau'(\sigma^\infty(a)).$$

Let $\mathcal{M}_{\mathbf{s}}(n)$ (and $\mathcal{M}'_{\mathbf{s}}(n)$) denote the set of all the factors of length n in $\mathbf{s}$ such that the sum of letters of such factor under the coding τ (and τ', respectively) attains the maximal value, i.e.,

$$\mathcal{M}_{\mathbf{s}}(n) := \{u \in \mathcal{F}_{\mathbf{s}}(n) \ :\ S(u) = M_{\mathbf{r}}(n)\},$$
$$\mathcal{M}'_{\mathbf{s}}(n) := \{u \in \mathcal{F}_{\mathbf{s}}(n) \ :\ S'(u) = M_{\mathbf{r}'}(n)\},$$

where $S := \mathrm{DS} \circ \tau$ and $S' := \mathrm{DS} \circ \tau'$.

3. The regularity of the abelian complexity of r and r′

In this section, we shall discuss the regularity of the abelian complexity functions of the Rudin–Shapiro sequence $\mathbf{r}$ and the sequence $\mathbf{r}'$. From now on, unless otherwise stated, we always set $\mathcal{A} = \{-1, 1\}$.

3.1. Statement of results

Theorem 1. *For every $n \geq 1$,*

$$M_{\mathbf{r}}(n) = M_{\mathbf{r}'}(n) =: M(n).$$

Moreover, $M(1) = 1$, $M(2) = 2$, $M(3) = 3$ and for $n \geq 1$,

$$M(4n) = 2M(n) + 2, \qquad M(4n+1) = 2M(n) + 1,$$
$$M(4n+2) = M(n) + M(n+1) + 1, \qquad M(4n+3) = 2M(n+1) + 1.$$

Corollary 1. *The sequence $(M(n))_{n\geq 0}$ is 2-regular.*

Proof. The result follows from Theorem 1, [2, Theorem 16.1.3 (e)] and [1, Theorem 2.9]. □

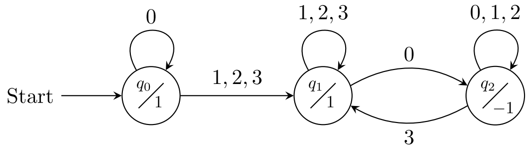

Fig. 2. The 4-automaton generating $(\Delta M(n))_{n\geq 0}$. (The most-significant digit is read first.)

For all $n \geq 0$, let

$$\Delta M(n) := M(n+1) - M(n).$$

The difference sequence $(\Delta M(n))_{n\geq 0}$ is characterized by the following result.

Corollary 2. $\Delta M(i) = 1$ *for* $0 \leq i \leq 3$, *and for* $n \geq 1$,

$$\begin{cases} \Delta M(4n) &= -\Delta M(4n+3) &= -1, \\ \Delta M(4n+1) &= \Delta M(4n+2) &= \Delta M(n). \end{cases} \tag{3.1}$$

Moreover, $(\Delta M(n))_{n\geq 0}$ *is a 2-automatic sequence.*

Proof. The difference sequence $(\Delta M(n))_{n\geq 0}$ can be generated by the automaton given in Fig. 2. □

Theorem 2. *For every* $n \geq 1$,

$$\rho_{\mathbf{r}}(n) = \rho_{\mathbf{r}'}(n) := \rho(n).$$

Moreover, $(\rho(n))_{n\geq 0}$ *is 2-regular.*

Proof. This result follows from Theorem 1 and Lemma 3 (proved below). □

3.2. Some lemmas

To prove Theorem 1, we need the following lemmas.

Lemma 1. *For every word* $w \in \{a, b, c, d\}^*$, *we have*

$$S(\sigma^2(w)) = 2S(w) \text{ and } S'(\sigma^2(w)) = 2S'(w).$$

Proof. Observing that both S and S' are morphisms from $(\{a, b, c, d\}^*, \cdot)$ to $(\mathbb{Z}, +)$, where '$\cdot$' is the concatenation of words, we only need to show the equalities in the lemma hold for every letter $x \in \{a, b, c, d\}$. By the definition of σ, we get

$$\sigma^2 : a \mapsto abac, b \mapsto abdb, c \mapsto dcac, d \mapsto dcdb.$$

Recall that $\tau : a \mapsto 1, b \mapsto 1, c \mapsto -1, d \mapsto -1$. Thus,

$$S(\sigma^2(a)) = S(abac) = \mathrm{DS} \circ \tau(abac) = \mathrm{DS}(111(-1)) = 2 = 2S(a).$$

One can verify the remaining cases in the same way. □

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

Lemma 2. *For every $n \geq 1$,*

$$M_{\mathbf{a}}(n) + m_{\mathbf{a}}(n) = 0,$$

where $\mathbf{a}$ *represents the Rudin–Shapiro sequence* $\mathbf{r}$ *or the sequence* $\mathbf{r}'$.

Proof. We only prove the case $\mathbf{a} = \mathbf{r}$. The result for $\mathbf{a} = \mathbf{r}'$ follows in the same way.

Let μ be the coding

$$\mu : a \mapsto d, b \mapsto c, c \mapsto b, d \mapsto a.$$

Then $\mu \circ \sigma = \sigma \circ \mu$ and $\mu \circ \mu = \mathrm{Id}$. We first prove the following two facts: for every $W \in \{a, b, c, d\}^n$ $(n \geq 1)$,

1. W is a factor of $\mathbf{s}$ if and only if $\mu(W)$ is a factor of $\mathbf{s}$;
2. $S(W) = M_{\mathbf{r}}(n)$ if and only if $S(\mu(W)) = m_{\mathbf{r}}(n)$.

For Fact 1, if W is a factor of $\mathbf{s}$, then W is a factor of $\sigma^k(a)$ for some k. Therefore, $\mu(W)$ is a factor of $\mu(\sigma^k(a)) = \sigma^k(d)$, which is a factor of $\sigma^{k+4}(a)$. Hence, $\mu(W)$ is also a factor of $\mathbf{s}$. The converse holds in the same argument by replacing W by $\mu(W)$. For Fact 2, suppose $S(W) = M_{\mathbf{r}}(n)$ and $S(\mu(W)) \neq m_{\mathbf{r}}(n)$. Without loss of generality, assume that $S(\mu(W)) > m_{\mathbf{r}}(n)$. This means that there exists a word $W' \in \mathcal{F}_{\mathbf{s}}(n)$ such that $|W'|_c + |W'|_d > |W|_c + |W|_d$. Therefore,

$$|\mu(W')|_a + |\mu(W')|_b = |W'|_c + |W'|_d > |\mu(W)|_c + |\mu(W)|_d = |W|_a + |W|_b.$$

It follows that $M_{\mathbf{r}}(n) = S(W) < S(\mu(W'))$, which is a contradiction. Using a similar argument, we can prove the converse.

Notice that $S(\mu(W)) = -S(W)$. Then, by Facts 1 and 2, the proof is completed. $\square$

Lemma 3. *For every $n \geq 1$,*

$$\rho_{\mathbf{a}}(n) = M_{\mathbf{a}}(n) + 1,$$

where $\mathbf{a}$ *represents the Rudin–Shapiro sequence* $\mathbf{r}$ *or the sequence* $\mathbf{r}'$.

Proof. The result follows from Proposition 1 and Lemma 2. $\square$

The following lemma characterizes $\Sigma_{\mathbf{r}}(\cdot, \cdot)$, which is useful in the study of $M_{\mathbf{r}}$.

Lemma 4. *For every $n \geq 1$ and $i \geq 0$, we have*

(1) $\Sigma_{\mathbf{r}}(4i, 4n) = 2\Sigma_{\mathbf{r}}(i, n)$,
(2) $\Sigma_{\mathbf{r}}(4i + 1, 4n) = \Sigma_{\mathbf{r}}(i, n) + \Sigma_{\mathbf{r}}(i + 1, n)$,
(3) $\Sigma_{\mathbf{r}}(4i + 2, 4n) = 2\Sigma_{\mathbf{r}}(i + 1, n)$,
(4) $\Sigma_{\mathbf{r}}(4i + 3, 4n) = 2\Sigma_{\mathbf{r}}(i + 1, n) - r(4i + 4n + 3) + r(4i + 3)$,
(5) $\Sigma_{\mathbf{r}}(4i, 4n + 1) = 2\Sigma_{\mathbf{r}}(i, n) + r(i + n)$,
(6) $\Sigma_{\mathbf{r}}(4i + 1, 4n + 1) = 2\Sigma_{\mathbf{r}}(i + 1, n) + r(i)$,
(7) $\Sigma_{\mathbf{r}}(4i + 2, 4n + 1) = 2\Sigma_{\mathbf{r}}(i + 1, n) + r(4i + 4n + 2)$,
(8) $\Sigma_{\mathbf{r}}(4i + 3, 4n + 1) = 2\Sigma_{\mathbf{r}}(i + 1, n) + r(4i + 3)$,
(9) $\Sigma_{\mathbf{r}}(4i, 4n + 2) = \Sigma_{\mathbf{r}}(i, n) + \Sigma_{\mathbf{r}}(i, n + 1) + r(i + n)$,

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

(10) $\Sigma_{\mathbf{r}}(4i+1,4n+2) = \Sigma_{\mathbf{r}}(i+1,n) + \Sigma_{\mathbf{r}}(i,n+1) + r(4i+4n+2)$,
(11) $\Sigma_{\mathbf{r}}(4i+2,4n+2) = \Sigma_{\mathbf{r}}(i+1,n) + \Sigma_{\mathbf{r}}(i+1,n+1) - r(i+n+1)$,
(12) $\Sigma_{\mathbf{r}}(4i+3,4n+2) = \Sigma_{\mathbf{r}}(i+1,n) + \Sigma_{\mathbf{r}}(i+1,n+1) + r(4i+3)$,
(13) $\Sigma_{\mathbf{r}}(4i,4n+3) = 2\Sigma_{\mathbf{r}}(i,n+1) - r(4i+4n+3)$,
(14) $\Sigma_{\mathbf{r}}(4i+1,4n+3) = 2\Sigma_{\mathbf{r}}(i,n+1) - r(i)$,
(15) $\Sigma_{\mathbf{r}}(4i+2,4n+3) = 2\Sigma_{\mathbf{r}}(i+1,n+1) - r(i+n+1)$,
(16) $\Sigma_{\mathbf{r}}(4i+3,4n+3) = 2\Sigma_{\mathbf{r}}(i+1,n+1) + r(4i+3)$.

Proof. By (2.1), we have for all $n \geq 0$,

$$r(4n) = r(4n+1) = r(n),\ r(4n+2) = -r(4n+3) = (-1)^n r(n).$$

Then by the previous equations and the definition of $\Sigma_{\mathbf{r}}$, these 16 equations can be verified directly. Here we give the proof of the first two equations as examples:

$$\begin{aligned}
\Sigma_{\mathbf{r}}(4i,4n) &= \sum_{j=4i}^{4i+4n-1} r(j) \\
&= \sum_{j=i}^{i+n-1} (r(4j) + r(4j+1) + r(4j+2) + r(4j+3)) \\
&= \sum_{j=i}^{i+n-1} (r(j) + r(j) + (-1)^j r(j) - (-1)^j r(j)) \\
&= 2\sum_{j=i}^{i+n-1} r(j) = 2\Sigma_{\mathbf{r}}(i,n)
\end{aligned}$$

and

$$\begin{aligned}
\Sigma_{\mathbf{r}}(4i+1,4n) &= \Sigma_{\mathbf{r}}(4i,4n) + r(4i+4n) - r(4i) = 2\Sigma_{\mathbf{r}}(i,n) + r(i+n) - r(i) \\
&= \Sigma_{\mathbf{r}}(i,n) + \Sigma_{\mathbf{r}}(i+1,n).
\end{aligned}$$

The remaining equations can be proved in the same way. □

Remark 1. Lemma 4 implies that the double sequence $(\Sigma_{\mathbf{r}})_{i\geq 0, n\geq 1}$ is a two-dimensional 2-regular sequence. For a definition of two-dimensional regular sequences, see [2].

The following lemma gives upper bounds of the maximal values of the sums of consecutive n terms of $\mathbf{r}$ and $\mathbf{r}'$.

Lemma 5. *For every* $n \geq 1$,

$$\begin{aligned}
M_{\mathbf{r}}(4n) &\leq 2M_{\mathbf{r}}(n) + 2, \\
M_{\mathbf{r}}(4n+1) &\leq 2M_{\mathbf{r}}(n) + 1, \\
M_{\mathbf{r}}(4n+2) &\leq M_{\mathbf{r}}(n) + M_{\mathbf{r}}(n+1) + 1, \\
M_{\mathbf{r}}(4n+3) &\leq 2M_{\mathbf{r}}(n+1) + 1.
\end{aligned}$$

Moreover, the above inequalities also hold for $M_{\mathbf{r}'}$.

Table 1
The initial values for Lemma 6.

n	1	2	3	4	5	6	7
W_n	a	ba	aba	$baba$	$babac$	$babdba$	$abdbaba$
$M_{\mathbf{r}}(n)$	1	2	3	4	3	4	5

Proof. For the first inequality, we shall use the first four equations of Lemma 4. By equations (1) to (3) of Lemma 4, we obtain that for $k = 0, 1, 2$,

$$\begin{aligned}\Sigma_{\mathbf{r}}(4i+k, 4n) &\le \max\{2\Sigma_{\mathbf{r}}(i,n), \Sigma_{\mathbf{r}}(i,n)+\Sigma_{\mathbf{r}}(i+1,n), 2\Sigma_{\mathbf{r}}(i+1,n)\}\\ &\le 2M_{\mathbf{r}}(n).\end{aligned}$$

When $k = 3$, by equation (4) of Lemma 4, we have

$$\begin{aligned}\Sigma_{\mathbf{r}}(4i+k, 4n) &= 2\Sigma_{\mathbf{r}}(i+1,n) - r(4i+4n+3) + r(4i+3)\\ &\le 2M_{\mathbf{r}}(n) + 2.\end{aligned}$$

Therefore, $M_{\mathbf{r}}(4n) \le 2M_{\mathbf{r}}(n) + 2$.

In a similar way, using the remaining 12 equations of Lemma 4, we can prove the remaining three inequalities for $M_{\mathbf{r}}$.

To prove the result for $M_{\mathbf{r}'}$, one can deduce a similar result to Lemma 4 for $\mathbf{r}'$, and apply a similar argument as above. We leave the details to the reader. □

3.3. Proof of Theorem 1

To prove Theorem 1, we only need to show that all equalities in Lemma 5 hold. For this, we shall construct two sequences of words, which attain the upper bounds in Lemma 5 for $\mathbf{r}$ and $\mathbf{r}'$, respectively. This will be done in Lemmas 6 and 7. Then, Theorem 1 follows directly from Lemmas 5, 6 and 7.

Now we shall give the sequence of words for $\mathbf{r}$. Let $(W_n)_{n\ge1}$ be the sequence of words defined by $W_1 = a$, $W_2 = ba$, $W_3 = aba$ and

$$\begin{cases} W_{4n} &= b\sigma^2(W_n)c^{-1},\\ W_{4n+1} &= b\sigma^2(W_n),\\ W_{4n+2} &= \begin{cases} b\sigma^2(W_{n+1})(bac)^{-1} & \text{if } \Delta M_{\mathbf{r}}(n) = 1,\\ cdb\sigma^2(W_n)c^{-1} & \text{if } \Delta M_{\mathbf{r}}(n) = -1,\end{cases}\\ W_{4n+3} &= \sigma^2(W_{n+1})c^{-1}.\end{cases} \tag{3.2}$$

Lemma 6. *Let $(W_n)_{n\ge1} \subset \{a,b,c,d\}^*$ be given by (3.2). Then, for every $n \ge 1$,*

(i) either $bW_n \prec \mathbf{s}$ or $dW_n \prec \mathbf{s}$ holds;
(ii) either $a \triangleright W_n$ or $c \triangleright W_n$ holds;
(iii) $W_n \in \mathcal{M}_{\mathbf{s}}(n)$.

Proof. We shall prove *(i)*, *(ii)* and *(iii)* simultaneously by induction.

Step 1. We shall show that the results hold for $n < 8$. Let $(W_n)_{n=1}^7$ be the words given in Table 1.

Notice that $(W_n)_{n=1}^7$ are factors of $\sigma^2(dba) = dcdbabdbabac$, which is a factor of $\mathbf{s}$, *(i)* and *(ii)* hold for $n < 8$. For $n = 1, 2, 3$, clearly $M_{\mathbf{r}}(n) = S(W_n)$, which implies $W_n \in \mathcal{M}_{\mathbf{s}}(n)$. Since $S(W_4) = 4 = 2M_{\mathbf{r}}(1) + 2$, $S(W_5) = 3 = 2M_{\mathbf{r}}(1) + 1$, $S(W_6) = 4 = M_{\mathbf{r}}(1) + M_{\mathbf{r}}(2) + 1$ and $S(W_7) = 5 = 2M_{\mathbf{r}}(2) + 1$, by Lemma 5, we have $S(W_n) = M_{\mathbf{r}}(n)$ and $W_n \in \mathcal{M}_{\mathbf{s}}(n)$ for $n = 4, 5, 6, 7$. Therefore, *(iii)* holds for $n < 8$.

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

Step 2. Assuming that *(i)*, *(ii)* and *(iii)* hold for $n < 4k$ ($k \geq 2$), we shall prove the results for $4k \leq n < 4(k+1)$. The proof in this step will be separated into the following two cases.
Case 1: $\Delta M_{\mathbf{r}}(k) = 1$. In this case, the induction hypotheses *(i)*, *(ii)* and *(iii)* yield the following facts:

(1a) $W_k \in \mathcal{M}_{\mathbf{s}}(k)$ and $W_{k+1} \in \mathcal{M}_{\mathbf{s}}(k+1)$;
(1b) $db\sigma^2(W_k)$ and $db\sigma^2(W_{k+1})$ are factors of $\mathbf{s}$;
(1c) either $a \triangleright W_k$ or $c \triangleright W_k$ holds, and $a \triangleright W_{k+1}$.

(In the last statement (1c), we can exclude the case $c \triangleright W_{k+1}$ since $\Delta M_{\mathbf{r}}(k) = 1$. In fact, if $W_{k+1} = Wc$, then

$$M_{\mathbf{r}}(k+1) = S(W_{k+1}) = S(W) + S(c) = S(W) - 1 \leq M_{\mathbf{r}}(k) - 1,$$

which contradicts the assumption $\Delta M_{\mathbf{r}}(k) = M_{\mathbf{r}}(k+1) - M_{\mathbf{r}}(k) = 1$.) Now, by (3.2) and (1b), dW_n is a factor of $\mathbf{s}$ for $4k \leq n \leq 4k+2$ and bW_{4k+3} is a factor of $\mathbf{s}$. This implies that (i) holds for $4k \leq n < 4(k+1)$, and

$$W_n \text{ is a factor of } \mathbf{s} \text{ for } 4k \leq n < 4(k+1). \tag{3.3}$$

By Fact (1c), we have $ac \triangleright \sigma^2(W_k)$ and $abac = \sigma^2(a) \triangleright \sigma^2(W_{k+1})$. Therefore, (3.2) gives

$$a \triangleright W_{4k},\ \ c \triangleright W_{4k+1},\ \ a \triangleright W_{4k+2} \text{ and } a \triangleright W_{4k+3}. \tag{3.4}$$

So *(ii)* holds. Now, by (3.2), (3.4), (1a) and Lemma 1, we have

$$\begin{cases} S(W_{4k}) &= S(b) + S(\sigma^2(W_k)) - S(c) = 2M_{\mathbf{r}}(k) + 2, \\ S(W_{4k+1}) &= S(b) + S(\sigma^2(W_k)) = 2M_{\mathbf{r}}(k) + 1, \\ S(W_{4k+2}) &= S(b) + S(\sigma^2(W_{k+1})) - S(bac) \\ &= 2M_{\mathbf{r}}(k+1) = M_{\mathbf{r}}(k) + M_{\mathbf{r}}(k+1) + 1, \\ S(W_{4k+3}) &= S(\sigma^2(W_{k+1})) - S(c) = 2M_{\mathbf{r}}(k+1) + 1. \end{cases} \tag{3.5}$$

By (3.3), (3.5) and Lemma 5, we have $W_n \in \mathcal{M}_{\mathbf{s}}(n)$ for $4k \leq n < 4(k+1)$, which is *(iii)*.
Case 2: $\Delta M_{\mathbf{r}}(k) = -1$. In this case, we first assert that dW_k is a factor of $\mathbf{s}$. By the induction hypothesis *(i)*, we only need to show that bW_k can not be a factor of $\mathbf{s}$. If this is not the case, then

$$M_{\mathbf{r}}(k+1) \geq S(bW_k) = 1 + S(W_k) = 1 + M_{\mathbf{r}}(k),$$

where the last equality follows from *(iii)*. Then, we have $\Delta M_{\mathbf{r}}(k) = M_{\mathbf{r}}(k+1) - M_{\mathbf{r}}(k) \geq 1$, which contradicts the assumption $\Delta M_{\mathbf{r}}(k) = M_{\mathbf{r}}(k+1) - M_{\mathbf{r}}(k) = 1$. Therefore, applying the induction hypotheses *(i)*, *(ii)* and *(iii)*, we have

(2a) $W_k \in \mathcal{M}_{\mathbf{s}}(k)$ and $W_{k+1} \in \mathcal{M}_{\mathbf{s}}(k+1)$;
(2b) $dcdb\sigma^2(W_k)$ and $b\sigma^2(W_{k+1})$ are factors of $\mathbf{s}$;
(2c) $ac \triangleright \sigma^2(W_k)$ and $ac \triangleright \sigma^2(W_{k+1})$.

By (3.2) and (2b), we have

$$dW_n \text{ is a factor of } \mathbf{s} \text{ for } 4k \leq n \leq 4k+2 \tag{3.6}$$

and bW_{4k+3} is a factor of $\mathbf{s}$. So *(i)* holds. This implies that

$$W_n \text{ is a factor of } \mathbf{s} \text{ for } 4k \le n < 4(k+1). \tag{3.7}$$

Combining (2c) and (3.2), we obtain that *(ii)* holds for $4k \le n < 4(k+1)$. Now, by (3.2), (2a), (2c) and Lemma 1, we have

$$\begin{cases} S(W_{4k}) &= S(b) + S(\sigma^2(W_k)) - S(c) = 2M_{\mathbf{r}}(k) + 2, \\ S(W_{4k+1}) &= S(b) + S(\sigma^2(W_k)) = 2M_{\mathbf{r}}(k) + 1, \\ S(W_{4k+2}) &= S(cbd) + S(\sigma^2(W_k)) - S(c) \\ &= 2M_{\mathbf{r}}(k) = M_{\mathbf{r}}(k) + M_{\mathbf{r}}(k+1) + 1, \\ S(W_{4k+3}) &= S(\sigma^2(W_{k+1})) - S(c) = 2M_{\mathbf{r}}(k+1) + 1. \end{cases} \tag{3.8}$$

By (3.7), (3.8) and Lemma 5, we have $W_n \in \mathcal{M}_{\mathbf{s}}(n)$ for $4k \le n < 4(k+1)$, which is *(iii)*. The proof is completed. □

For $\mathbf{r}'$, let $(\widetilde{W}_n)_{n\ge1}$ be the sequence of words defined by $\widetilde{W}_1 = c$, $\widetilde{W}_2 = ca$, $\widetilde{W}_3 = cac$ and

$$\begin{cases} \widetilde{W}_{4n} &= d^{-1}\sigma^2(\widetilde{W}_n)a, \\ \widetilde{W}_{4n+1} &= \sigma^2(\widetilde{W}_n)a, \\ \widetilde{W}_{4n+2} &= \begin{cases} (dca)^{-1}\sigma^2(\widetilde{W}_{n+1})a & \text{if } \Delta M_{\mathbf{r}'}(n) = 1, \\ d^{-1}\sigma^2(\widetilde{W}_n)abd & \text{if } \Delta M_{\mathbf{r}'}(n) = -1, \end{cases} \\ \widetilde{W}_{4n+3} &= d^{-1}\sigma^2(\widetilde{W}_{n+1}). \end{cases} \tag{3.9}$$

Lemma 7. *Let $(\widetilde{W}_n)_{n\ge1} \subset \{a,b,c,d\}^*$ be given by (3.9). Then, for every $n \ge 1$,*

(i) *either $\widetilde{W}_n a \prec \mathbf{s}$ or $\widetilde{W}_n b \prec \mathbf{s}$ holds;*
(ii) *either $c \lhd \widetilde{W}_n$ or $d \lhd \widetilde{W}_n$ holds;*
(iii) *$\widetilde{W}_n \in \mathcal{M}'_{\mathbf{s}}(n)$.*

Proof. The proof of this lemma is similar to the proof of Lemma 6. □

For every k-automatic sequence $\mathbf{w} = w(0)w(1)\cdots \in \{-1,1\}^{\mathbb{N}}$, the regularity of the maximal partial sums $(M_{\mathbf{w}}(n))_{n\ge1}$ and the minimal partial sums $(m_{\mathbf{w}}(n))_{n\ge1}$ imply the regularity of the abelian complexity $(\rho_{\mathbf{w}}(n))_{n\ge1}$. By proving the same result as Lemma 4, one can show that the double sequence $(\Sigma_{\mathbf{w}}(i,n))_{i\ge0,n\ge1}$ is a two-dimensional k-regular sequence. In fact, it is not hard to show that $(\Sigma_{\mathbf{w}}(i,n))_{i\ge0}$ is k-automatic for every fixed $n \ge 1$, and $(\Sigma_{\mathbf{w}}(i,n))_{n\ge1}$ is k-regular for every fixed $i \ge 0$. Moreover, Theorem 1 and Lemma 2 show that $(\max_{i\ge0}\Sigma_{\mathbf{w}}(i,n))_{n\ge1}$ and $(\min_{i\ge0}\Sigma_{\mathbf{w}}(i,n))_{n\ge1}$ are still k-regular when $\mathbf{w}$ is the Rudin–Shapiro sequence $\mathbf{r}$ or its related sequence $\mathbf{r}'$. This implies the regularity of the abelian complexity functions $(\rho_{\mathbf{r}}(n))_{n\ge0}$ and $(\rho_{\mathbf{r}'}(n))_{n\ge0}$. It is natural to ask that whether $(\max_{i\ge0}\Sigma_{\mathbf{w}}(i,n))_{n\ge1}$ and $(\min_{i\ge0}\Sigma_{\mathbf{w}}(i,n))_{n\ge1}$ are always k-regular for every k-automatic sequence $\mathbf{w}$ over $\{-1,1\}$.

4. Box dimension of $\lambda(x)$

Let $M(x) := M(\lfloor x \rfloor)$ $(x > 0)$ be the continuous version of the maximal digit sum function, and $\rho(x) = M(x) + 1$. Now we study the following limit function:

$$\lambda(x) := \lim_{k\to\infty} \frac{\rho(4^k x)}{\sqrt{4^k x}}. \tag{4.1}$$

From the above definition, provided the limit exists, it is easy to see that $\lambda(x)$ is self-similar in the sense that for every $x > 0$,

$$\lambda(4x) = \lambda(x).$$

The existence of the limit in (4.1) follows from the same argument in [4, Theorem 1]. For completeness, we give the details in the following Proposition 2.

Let

$$\sum_{j=0}^{\infty} x_j 4^{-j} \tag{4.2}$$

denote the 4-adic expansion of a real positive number $x > 0$, where $x_0 \in \mathbb{N}$ and $x_j \in \{0, 1, 2, 3\}$ for all $j \geq 1$. In the expansion (4.2), we always assume that there are infinitely many j such that $x_j \neq 3$. Let

$$a_j(x) := \begin{cases} -1, & \text{if } 4^j x < 1, \\ \Delta M(\lfloor 4^j x \rfloor - 1), & \text{otherwise,} \end{cases}$$

and

$$d(y) := \begin{cases} 1, & \text{if } y = 0 \text{ or } 2, \\ 0, & \text{if } y = 1, \\ 2, & \text{if } y = 3. \end{cases}$$

Proposition 2. *The limit (4.1) exists for all $x > 0$, and for every $x > 0$, it satisfies*

$$\lambda(x) = \frac{\rho(x) + a(x)}{\sqrt{x}}, \tag{4.3}$$

where $a(x) := \sum_{j=1}^{\infty} d(x_j) a_j(x) 2^{-j}$. Moreover, for every positive integer n,

$$\lambda(n) = (\rho(n) + 1)/\sqrt{n}.$$

Proof. By Theorem 1 and Corollary 2, we have

$$M(4n + i) = 2M(n) + 1 + d(i)\Delta M(4n + i - 1)$$

for all $n \geq 1$ and $i = 0, 1, 2, 3$. Let N be the smallest integer such that $4^N x \geq 1$. Then, for every $k \geq N$,

$$\begin{aligned} M(4^k x) &= M(\lfloor 4^k x \rfloor) = M(4\lfloor 4^{k-1} x \rfloor + x_k) \\ &= 2M(\lfloor 4^{k-1} x \rfloor) + 1 + d(x_k)\Delta M(\lfloor 4^k x \rfloor - 1) \\ &= 2M(4^{k-1} x) + 1 + d(x_k) a_k(x). \end{aligned}$$

For $1 \leq k < N$, $d(x_k) = d(0) = 1$ and $a_k(x) = -1$. Thus, we also have

$$\begin{aligned} M(4^k x) &= 0 = 1 + (-1) \\ &= 1 + d(x_k) a_k(x) \\ &= 2M(4^{k-1} x) + 1 + d(x_k) a_k(x). \end{aligned}$$

By induction, the above equation yields

$$M(4^k x) = 2^k M(x) + \sum_{j=1}^{k} d(x_j) a_j(x) 2^{k-j} + (2^k - 1).$$

Now, by Lemma 3

$$\frac{\rho(4^k x)}{\sqrt{4^k x}} = \frac{M(4^k x) + 1}{\sqrt{4^k x}} = \frac{\rho(x)}{\sqrt{x}} + \frac{1}{\sqrt{x}} \sum_{j=1}^{k} d(x_j) a_j(x) 2^{-j}.$$

Letting $k \to \infty$ and noticing that the series in (4.3) converges absolutely, we obtain (4.3).

When $x = n \in \mathbb{N}^+$, $x_0 = n$, $x_j = 0$ and $a_j = 4^j n - 1$ for all $j \geq 1$, the infinite sums in (4.3) turn out to be

$$\sum_{j=1}^{\infty} d(x_j) a_j(x) 2^{-j} = \sum_{j=1}^{\infty} \Delta M(4^j n - 1) 2^{-j} = 1,$$

where the last equality holds according to Corollary 2. Applying the above equation to (4.3), we complete the proof. □

4.1. Auxiliary lemmas

Now, we prove some auxiliary lemmas, which are used in the calculation of the box dimension of the graph of the function $\lambda(x)$. For every $k \geq 1$ and $0 \leq z < 4^k$, where $z \in \mathbb{N}$, let

$$I_k(z) := [z 4^{-k}, (z+1) 4^{-k}).$$

Then $[0,1) = \bigcup_{0 \leq z < 4^k} I_k(z)$. Let $\frac{z}{4^k} = \sum_{j=1}^{k} z_j 4^{-j}$ denote the 4-adic expansion of $z 4^{-k}$. If $y = \sum_{j=1}^{\infty} y_j 4^{-j} \in I_k(z)$, then $y_i = z_i$ for $i = 1, 2, \cdots, k$.

First, we determine the difference of values of $a(\cdot)$ at the end points of a 4-adic interval $I_k(z)$.

Lemma 8. *Let $k \geq 1$ and $z \in \mathbb{N}$ with $1 \leq z < 4^k$. Then*

$$a(z4^{-k}) - a((z+1)4^{-k}) = \begin{cases} -2^{-k}, & \text{if } z \leq 4^k - 2, \\ 1 - 2^{-k}, & \text{if } z = 4^k - 1. \end{cases}$$

Proof. When $z = 4^k - 1$, we have $z 4^{-k} = \sum_{j=1}^{k} 3 \cdot 4^{-j}$ and $(z+1)4^{-k} = 1$. So

$$a(z4^{-k}) - a((z+1)4^{-k}) = (2 - 2^{-k}) - 1 = 1 - 2^{-k}.$$

When $1 \leq z \leq 4^{-k} - 2$, $z4^{-k}$ and $(z+1)4^{-k}$ have the 4-adic expansions

$$z 4^{-k} = \sum_{j=1}^{k} z_j 4^{-j} \text{ and } (z+1) 4^{-k} = \sum_{j=1}^{k} z_j' 4^{-j}.$$

Implicitly, we assume that $z_j = z_j' = 0$ for $j > k$. Let $1 \leq h \leq k$ be the integer such that $z_h \neq 3$ and $z_j = 3$ for $j = h+1, \cdots, k$. Then

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

$$z'_j = \begin{cases} z_j, & \text{when } j < h, \\ z_j + 1, & \text{when } j = h, \\ 0, & \text{when } j > h. \end{cases}$$

Set $D_j := d(z_j)a_j(z4^{-k}) - d(z'_j)a_j((z+1)4^{-k})$. Then

$$a(z4^{-k}) - a((z+1)4^{-k}) = \sum_{j=1}^{\infty} D(j)2^{-j}.$$

Clearly, $D_j = 0$ when $j < h$ or $j > k$. Since $a_j(z4^{-k}) = a_j((z+1)4^{-k}) = 1$ for $h+2 \leq j \leq k$, we have for $h+2 \leq j \leq k$,

$$D_j = d(3) - d(0) = 1.$$

Set $u := 4^h \sum_{j=1}^{h} z_j 4^{-j}$. If $u \geq 1$, we have

$$\begin{aligned} D_h + 2^{-1}D_{h+1} &= (d(z_h)\Delta M(u-1) - d(z'_h)\Delta M(u)) \\ &\quad + 2^{-1}\left(d(3)\Delta M(4u+2) - d(0)\cdot\Delta M(4u+3)\right) \\ &= d(z_h)\Delta M(u-1) - d(z'_h)\Delta M(u) + \Delta M(u) - 2^{-1} \\ &= \begin{cases} d(0)\cdot 1 - d(1)\cdot(-1) + (-1) - 2^{-1}, & \text{if } z_h = 0, \\ d(1)\cdot(-1) - d(2)\cdot\Delta M(u) + \Delta M(u) - 2^{-1}, & \text{if } z_h = 1, \\ d(2)\cdot\Delta M(u) - d(3)\cdot\Delta M(u) + \Delta M(u) - 2^{-1}, & \text{if } z_h = 2, \end{cases} \\ &= -2^{-1}. \end{aligned}$$

If $u = 0$, then $z_h = 0$ and

$$\begin{aligned} D_h + 2^{-1}D_{h+1} &= d(0)\cdot(-1) - d(1)\Delta M(0) \\ &\quad + 2^{-1}\left(d(3)\Delta M(2) - d(0)\Delta M(3)\right) \\ &= -2^{-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} a(z4^{-k}) - a((z+1)4^{-k}) &= \sum_{j=1}^{\infty} D(j)2^{-j} \\ &= 2^{-h}D_h + 2^{-h-1}D_{h+1} + \sum_{j=h+2}^{k} D(j)2^{-j} \\ &= -2^{-h-1} + \left(2^{-h-1} - 2^{-k}\right) = -2^{-k}. \quad \square \end{aligned}$$

Lemma 9. *There exists $c > 0$, such that for every $x, y \in (0,1)$,*

$$|a(x) - a(y)| \leq c|x-y|^{1/2}.$$

Proof. Let $x, y \in (0, 1)$ and $x < y$. Let

$$\sum_{j=1}^{\infty} x_j 4^{-j} \text{ and } \sum_{j=1}^{\infty} y_j 4^{-j}$$

denote the 4-adic expansions of x and y. Set $D_j := d(x_j)a_j(x) - d(y_j)a_j(y)$. Then $|D_j| \leq 4$ for $j \geq 1$.

Let k be the integer such that $4^{-k-1} \leq y - x < 4^{-k}$. Then x and y can be covered by at most two (adjacent) 4-adic intervals of level k. Suppose $x, y \in I_k(z)$ for some $0 \leq z < 4^k$. Then $x_j = y_j$ for $i = 1, 2, \cdots, k$. Consequently, $D_j = 0$ for $1 \leq j \leq k$. So

$$\begin{aligned}
|a(x) - a(y)| &= \left| \sum_{j=k+1}^{\infty} D_j 2^{-j} \right| \\
&\leq 4 \sum_{j=k+1}^{\infty} 2^{-j} = 4 \cdot 2^{-k} \\
&\leq 8|x - y|^{1/2}.
\end{aligned}$$

On the other hand, suppose $x \in I_k(z)$ and $y \in I_k(z + 1)$, where $0 \leq z < 4^k - 1$. Let h be the largest integer such that $x, y \in I_h(z')$ for some $0 \leq z' < 4^h$. Clearly, $0 \leq h < k$. In this case, the 4-adic expansions of x and y satisfy

$$\begin{cases}
y_j = x_j, & \text{if } 1 \leq j \leq h, \\
y_j = x_j + 1, & \text{if } j = h + 1, \\
y_j = 0 \text{ and } x_j = 3, & \text{if } h + 2 \leq j \leq k.
\end{cases}$$

(We remark that $x_{h+1} \neq 3$ by the choice of h.) Hence, $D_j = 0$ for $1 \leq j \leq h$. Similar discussions as in Lemma 8 yield

$$D_{h+1} + 2^{-1} D_{h+2} = -2^{-1}.$$

Moreover, for $h + 3 \leq j \leq k$, $D_j = d(3) - d(0) = 1$. Therefore,

$$\begin{aligned}
|a(x) - a(y)| &= \left| \sum_{j=1}^{\infty} D_j 2^{-j} \right| = \left| \sum_{j=h+1}^{k} D_j 2^{-j} + \sum_{j=k+1}^{\infty} D_j 2^{-j} \right| \\
&\leq \left| 2^{-h-1}(D_{h+1} + 2^{-1} D_{h+2}) + \sum_{j=h+3}^{k} D_j 2^{-j} \right| + 4 \sum_{j=k+1}^{\infty} 2^{-j} \\
&= 5 \cdot 2^{-k} \leq 10|x - y|^{1/2}. \quad \square
\end{aligned}$$

4.2. Calculation of the box dimension

Theorem 3. *For every $0 < \alpha < \beta \leq 1$,*

$$\dim_B \{(x, \lambda(x)) : \alpha < x < \beta\} = \frac{3}{2}.$$

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019

Proof. For every $x, y \in (\alpha, \beta)$ and $x < y$, $\rho(x) = \rho(y) = \rho(0) = 1$,

$$\begin{aligned}|\lambda(x) - \lambda(y)| &= \left|\frac{\rho(x) + a(x)}{\sqrt{x}} - \frac{\rho(y) + a(y)}{\sqrt{y}}\right| \\ &= \left|\frac{a(x) + 1}{\sqrt{x}} - \frac{a(y) + 1}{\sqrt{y}}\right| \\ &= \left|\frac{a(x) - a(y)}{\sqrt{x}} + \frac{\sqrt{y} - \sqrt{x}}{\sqrt{xy}}(a(y) + 1)\right| \\ &\le \alpha^{-1/2}|a(x) - a(y)| + 3\alpha^{-1}\sqrt{y - x} \\ &\le (c\alpha^{-1/2} + 3\alpha^{-1})|x - y|^{1/2},\end{aligned}$$

where the last inequality holds by Lemma 9. Now by [8, Corollary 11.2 (a)],

$$\overline{\dim}_B\{(x, \lambda(x)) : \alpha < x < \beta\} \le \frac{3}{2}. \tag{4.4}$$

For every $k \ge 1$, let N_k be the number of 4^{-k}-mesh squares that intersect the graph of $\lambda(x)$ on (α, β). For every $k \ge 1$ and $\lfloor \alpha 4^k \rfloor < z \le \lfloor \beta 4^k \rfloor$, the number of 4^{-k}-mesh squares that intersect the graph of $\lambda(x)$ on $I_k(z)$ is larger than $\left|\lambda((z + 1)4^{-k}) - \lambda(z4^{-k})\right| / 4^{-k}$. Choose K_1 large enough such that for all $k > K_1$, $3 \cdot 2^k < \lfloor \alpha 4^k \rfloor$ $(< z)$. Then, by Lemma 8,

$$\begin{aligned}\left|\lambda((z + 1)4^{-k}) - \lambda(z4^{-k})\right| &= \left|\frac{1 + a((z + 1)4^{-k})}{\sqrt{(z + 1)4^{-k}}} - \frac{1 + a(z4^{-k})}{\sqrt{z4^{-k}}}\right| \\ &= \frac{1}{\sqrt{z4^{-k}}}\left|a((z + 1)4^{-k}) - a(z4^{-k})\right. \\ &\qquad \left. + \frac{\sqrt{z4^{-k}} - \sqrt{(z + 1)4^{-k}}}{\sqrt{(z + 1)4^{-k}}}(1 + a((z + 1)4^{-k}))\right| \\ &\ge \frac{1}{\sqrt{\beta}}\left(2^{-k} - \frac{\left|1 + a((z + 1)4^{-k})\right|}{z + 1 + \sqrt{z^2 + z}}\right) \\ &\ge 2^{-k} \cdot \frac{1}{\sqrt{\beta}}\left(1 - \frac{3 \cdot 2^k}{z + 1 + \sqrt{z^2 + z}}\right) > \frac{1}{2\sqrt{\beta}} \cdot 2^{-k}.\end{aligned}$$

Choose K_2 large enough such that for all $k > K_2$, $\lfloor \beta 4^k \rfloor - \lfloor \alpha 4^k \rfloor - 1 > 4^k(\beta - \alpha)/2$. Hence, for every $k > \max\{K_1, K_2\}$,

$$\begin{aligned}N_k &\ge \sum_{\lfloor \alpha 4^k \rfloor < z < \lfloor \beta 4^k \rfloor} \frac{\left|\lambda((z + 1)4^{-k}) - \lambda(z4^{-k})\right|}{4^{-k}} \\ &\ge \frac{1}{2\sqrt{\beta}} \sum_{\lfloor \alpha 4^k \rfloor < z < \lfloor \beta 4^k \rfloor} \frac{2^{-k}}{4^{-k}} = \frac{\lfloor \beta 4^k \rfloor - \lfloor \alpha 4^k \rfloor - 1}{2\sqrt{\beta}} \cdot 2^k \\ &> \frac{\beta - \alpha}{4\sqrt{\beta}} \cdot 2^{3k}.\end{aligned}$$

Therefore,

$$\begin{aligned}\underline{\dim}_B\{(x,\lambda(x)) : \alpha < x < \beta\} &= \liminf_{k\to\infty} \frac{\log N_k}{-\log 4^{-k}} \\ &\geq \liminf_{k\to\infty} \frac{\log\left(2^{3k}(\beta-\alpha)/4\sqrt{\beta}\right)}{-\log 4^{-k}} = \frac{3}{2}. \end{aligned} \tag{4.5}$$

The result follows from (4.4) and (4.5). □

Corollary 3. *For every* $0 < \alpha < \beta$,

$$\dim_B\{(x,\lambda(x)) : \alpha < x < \beta\} = \frac{3}{2}.$$

Proof. Let K be an integer such that $\beta/4^K \leq 1$. Since $\lambda(4x) = \lambda(x)$ for $x > 0$, the following mapping

$$f : (x,\lambda(x)) \mapsto (4^K x, \lambda(4^K x))$$

is a bi-Lipschitz mapping in $\mathbb{R}^2$, and

$$\begin{aligned} f\left(\{(x,\lambda(x)) : 4^{-K}\alpha < x < 4^{-K}\beta\}\right) &= \left\{\left(4^K x, \lambda(4^K x)\right) : 4^{-K}\alpha < x < 4^{-K}\beta\right\} \\ &= \{(y,\lambda(y)) : \alpha < y < \beta\}. \end{aligned}$$

The result follows from Theorem 3 and the above equation. □

References

[1] J.P. Allouche, J. Shallit, The ring of k-regular sequences, Theoret. Comput. Sci. 98 (2) (1992) 163–197.
[2] J.P. Allouche, J. Shallit, Automatic Sequences: Theory, Applications, Generalizations, Cambridge University Press, 2003.
[3] L. Balková, K. Břinda, O. Turek, Abelian complexity of infinite words associated with quadratic Parry numbers, Theoret. Comput. Sci. 412 (45) (2011) 6252–6260.
[4] J. Brillhart, P. Erdős, P. Morton, On sums of Rudin–Shapiro coefficients II, Pacific J. Math. 107 (1) (1983) 39–69.
[5] J. Chen, X.-T. Lü, Z.-X. Wen, W. Wu, On the quasi-linear functions, preprint.
[6] E.M. Coven, G.A. Hedlund, Sequences with minimal block growth, Math. Syst. Theory 7 (1973) 138–153.
[7] J. Currie, N. Rampersad, Recurrent words with constant Abelian complexity, Adv. in Appl. Math. 47 (1) (2011) 116–124.
[8] K.J. Falconer, Fractal Geometry: Mathematical Foundations and Applications, second edition, John Wiley & Sons, 2004.
[9] B. Madill, N. Rampersad, The abelian complexity of the paperfolding word, Discrete Math. 313 (7) (2013) 831–838.
[10] G. Richomme, K. Saari, L.Q. Zamboni, Balance and abelian complexity of the Tribonacci word, Adv. in Appl. Math. 45 (2) (2010) 212–231.
[11] G. Richomme, K. Saari, L.Q. Zamboni, Abelian complexity of minimal subshifts, J. Lond. Math. Soc. 83 (1) (2011) 79–95.
[12] W.-X. Shen, Hausdorff dimension of the graphs of the classical Weierstrass functions, arXiv:1505.03986.
[13] H. Yue, The Bouligand dimension of the Rudin–Shapiro function (in Chinese), Acta Math. Appl. Sin. 18 (1) (1995) 27–36.

Please cite this article in press as: X. Lü et al., On the abelian complexity of the Rudin–Shapiro sequence, J. Math. Anal. Appl. (2017), http://dx.doi.org/10.1016/j.jmaa.2017.02.019