More on the dynamics of the symbolic square root map

Jarkko Peltomäki[*,1,2] and Markus Whiteland[2]

[1]Turku Centre for Computer Science TUCS, Turku, Finland
[2]University of Turku, Department of Mathematics and Statistics, Turku, Finland

Abstract

In our earlier paper [A square root map on Sturmian words, Electron. J. Combin. 24.1 (2017)], we introduced a symbolic square root map. Every optimal squareful infinite word s contains exactly six minimal squares and can be written as a product of these squares: $s = X_1^2 X_2^2 \cdots$. The square root $\sqrt{s}$ of s is the infinite word $X_1 X_2 \cdots$ obtained by deleting half of each square. We proved that the square root map preserves the languages of Sturmian words (which are optimal squareful words). The dynamics of the square root map on a Sturmian subshift are well understood. In our earlier work, we introduced another type of subshift of optimal squareful words which together with the square root map form a dynamical system. In this paper, we study these dynamical systems in more detail and compare their properties to the Sturmian case. The main results are characterizations of periodic points and the limit set. The results show that while there is some similarity it is possible for the square root map to exhibit quite different behavior compared to the Sturmian case.

Keywords: Sturmian word, optimal squareful word, symbolic square root map

1 Introduction

Kalle Saari showed in [9, 10] that every Sturmian word contains exactly six minimal squares (that is, squares having no proper square prefixes) and that each position of a Sturmian word begins with a minimal square. Thus a Sturmian word $\mathbf{s}$ can be expressed as a product of minimal squares: $\mathbf{s} = X_1^2 X_2^2 \cdots$. In our earlier work [5], see also [4], we defined the square root $\sqrt{\mathbf{s}}$ of the word $\mathbf{s}$ to be the infinite word $X_1 X_2 \cdots$ obtained by deleting half of each square X_i^2. We proved that the words $\mathbf{s}$ and $\sqrt{\mathbf{s}}$ have the same language, that is, the square root map preserves the languages of Sturmian words. More precisely, we showed that if $\mathbf{s}$ has slope α and intercept ρ, then $\sqrt{\mathbf{s}}$ has intercept $\psi(\rho)$, where $\psi(\rho) = \frac{1}{2}(\rho + 1 - \alpha)$. The simple form of the function ψ immediately describes the dynamics of the square root map in the subshift Ω_α of Sturmian words of slope α: all words in Ω_α are attracted to the set $\{01\mathbf{c}_\alpha, 10\mathbf{c}_\alpha\}$ of words of intercept $1 - \alpha$; here $\mathbf{c}_\alpha$ is the standard Sturmian word of slope α.

The square root map makes sense for any word expressible as a product of squares. Saari defines in [10] an intriguing class of such infinite words which he calls optimal squareful words. Optimal squareful words are aperiodic infinite words containing the least number of minimal squares such that every position begins with a square. It turns out that such a word must be binary, and it must contain exactly six minimal squares; less than six minimal squares forces the word to be ultimately periodic. Moreover, the six minimal squares must be the minimal squares

---

of some Sturmian language; the set of optimal squareful words is however larger than the set of Sturmian words. The six minimal squares of an optimal squareful word take the following form for some integers $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{a} \geq 1$ and $\mathfrak{b} \geq 0$:

$$0^2, \qquad (10^{\mathfrak{a}})^2,$$
$$(010^{\mathfrak{a}-1})^2, \quad (10^{\mathfrak{a}+1}(10^{\mathfrak{a}})^{\mathfrak{b}})^2,$$
$$(010^{\mathfrak{a}})^2, \quad (10^{\mathfrak{a}+1}(10^{\mathfrak{a}})^{\mathfrak{b}+1})^2.$$

It is natural to ask if there are non-Sturmian optimal squareful words whose languages the square root map preserves. In [5], we proved by an explicit construction that such words indeed exist. The construction is as follows. The substitution

$$\tau : \begin{array}{l} S \mapsto LSS \\ L \mapsto SSS \end{array}$$

produces two infinite words $\Gamma_1^\circ = SSSLSSLSS \cdots$ and $\Gamma_2^\circ = LSSLSSLSS \cdots$ having the same language $\mathcal{L}$. Let $\widetilde{s}$ be a (long enough) reversed standard word in some Sturmian language and $L(\widetilde{s})$ be the word obtained from $\widetilde{s}$ by exchanging its first two letters. By substituting the language $\mathcal{L}$ by the substitution σ mapping the letters S and L respectively to $\widetilde{s}$ and $L(\widetilde{s})$, we obtain a subshift Ω consisting of optimal squareful words. We proved that the words Γ_1 and Γ_2, the σ-images of Γ_1° and Γ_2°, are fixed by the square root map and, more generally, either $\sqrt{\mathbf{w}} \in \Omega$ or $\sqrt{\mathbf{w}}$ is periodic for all $\mathbf{w} \in \Omega$.

The aim of this paper is to study the dynamics of the square root map in the subshift Ω in the slightly generalized case where $\tau(S) = LS^{2\mathfrak{c}}$ and $\tau(L) = S^{2\mathfrak{c}+1}$ for some positive integer $\mathfrak{c}$ and to see in which ways the dynamics differ from the Sturmian case. Our main results are the characterization of periodic and asymptotically periodic points and the limit set. We show that asymptotically periodic points must be ultimately periodic points and that periodic points must be fixed points; there are only two fixed points: Γ_1 and Γ_2. We prove that any word in Ω that is not an infinite product of the words $\sigma(S)$ and $\sigma(L)$ must eventually be mapped to a periodic word, thus having a finite orbit, while products of $\sigma(S)$ and $\sigma(L)$ are always mapped to aperiodic words. It follows that the limit set contains exactly the words that are products of $\sigma(S)$ and $\sigma(L)$. Additionally, we show that the limit set can be expressed as a disjoint union of infinitely many invariant subsets. Moreover, we study the injectivity of the square root map on Ω: only certain left extensions of the words Γ_1 and Γ_2 may have more than one preimage.

Let us make a brief comparison with the Sturmian case to see that the obtained results indicate that the square root map behaves somewhat differently on Ω. The mapping ψ, defined above, is injective, so in the Sturmian case all words have at most one preimage. As ψ maps points strictly towards the point $1 - \alpha$ on the circle, all points are asymptotically periodic (see Definition 4.2) and all periodic points are fixed points. The fixed points are the two words $01\mathbf{c}_\alpha$ and $10\mathbf{c}_\alpha$ mentioned above, and the limit set consists only of these two fixed points.

This paper is an extended version of the conference paper [6] presented at the conference WORDS 2017 in Montréal, September 2017. The conference paper omitted proofs, which are now given in this full paper. Further analysis on limit sets, namely the study of invariant sets, is presented in Section 3. Section 5 contains completely new material. Additionally Sections 6 and 7 contain novel discussion on the topic.

The paper is organized as follows. The following section gives needed preliminary results on Sturmian words and standard words, and it describes the construction of the subshift Ω in full detail. In Section 3, we proceed to characterize the limit set and to study injectivity. Section 4 contains results on periodic points. Next, in Section 5, we take a closer look at solutions to a specific word equation that are important in our constructions. We conclude the paper by additional remarks in Section 6 and open problems in Section 7.

2 Notation and Preliminary Results

For basic word-combinatorial concepts, we refer the reader to the book [2] or to the corresponding section of our previous paper [5]. In this paper, we consider finite or infinite binary words, which we take to be over the alphabets $\{0, 1\}$ or $\{S, L\}$. We denote the length of the word w by $|w|$. The empty word has length 0 and is denoted by ε. We refer to the kth letter of w by $w[k]$, and we index letters from 0. If $w = u^2$, then we call w a *square* with *square root* u. A square is *minimal* if it does not have a square as a proper prefix. By $L(w)$ we mean the word obtained from w by exchanging its first two letters (we will not apply L to too short words). Let C be the cyclic shift operator defined by the formula $C(a_0 a_1 \cdots a_{n-1}) = a_1 \cdots a_{n-1} a_0$ for letters a_i. The words w, $C(w)$, $C^2(w)$, ..., $C^{|w|-1}(w)$ are the *conjugates of w*. If u is a conjugate of w, then we say that u is conjugate to w. We write $u \lhd v$ if the word u is lexicographically less than v. For binary words over $\{0, 1\}$, we set $0 \lhd 1$. If $w = uv$, then by $w \cdot v^{-1}$ we mean the word u. *Substitutions*, sometimes called morphisms, are maps $f \colon A^* \to B^*$ such that $f(uv) = f(u)f(v)$ for all $u, v \in A^*$.

An infinite word is *ultimately periodic* if it is of the form $uvvv \cdots$; otherwise it is *aperiodic*. We distinguish finite words from infinite words by writing the symbols referring to infinite words in boldface. A *subshift* Ω is a set of infinite words whose language is included in some extendable and factor-closed language $\mathcal{L}(\Omega)$, which is called the language of the subshift. If $\mathcal{L}(\Omega)$ is the language of some infinite word $\mathbf{w}$, then we say that the corresponding subshift is generated by $\mathbf{w}$. Subshifts are clearly shift-invariant; the shift operator on infinite words is denoted by T. If every word in a subshift is aperiodic, then we call the subshift *aperiodic*. A subshift is *minimal* if it does not contain nonempty subshifts as proper subsets.

2.1 Sturmian Words and Standard Words

Several proofs in [5] regarding Sturmian words and the square root map require knowledge on continued fractions. In this paper, only some familiarity with continued fractions is required. We only recall that every irrational real number α has a unique infinite continued fraction expansion:

$$\alpha = [a_0; a_1, a_2, a_3, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ldots}}} \tag{1}$$

with $a_0 \in \mathbb{Z}$ and $a_k \in \mathbb{Z}_+$ for $k \geq 1$. The numbers a_i are called the *partial quotients* of α. The rational numbers $[a_0; a_1, a_2, a_3, \ldots, a_k]$, denoted by p_k/q_k, are called *convergents* of α. The *semiconvergents* (or intermediate fractions) $p_{k,\ell}/q_{k,\ell}$ of α are defined as the fractions

$$\frac{\ell p_{k-1} + p_{k-2}}{\ell q_{k-1} + q_{k-2}}$$

for $1 \leq \ell < a_k$ and $k \geq 2$ (if they exist). An introduction to continued fractions in relation to Sturmian words can be found in [4, Chapter 4].

We view here Sturmian words as the infinite words obtained as codings of orbits of points in an irrational circle rotation with two intervals. For alternative definitions and further details, see [7, 2]. We identify the unit interval $[0, 1)$ with the unit circle $\mathbb{T}$. Let α in $(0, 1)$ be irrational. The map $R \colon \mathbb{T} \to \mathbb{T}$, $\rho \mapsto \{\rho + \alpha\}$, where $\{\rho\}$ stands for the fractional part of the number ρ, defines a rotation on $\mathbb{T}$. Divide the circle $\mathbb{T}$ into two intervals I_0 and I_1 defined by the points 0 and $1 - \alpha$. Then define the coding function ν by setting $\nu(\rho) = 0$ if $\rho \in I_0$ and $\nu(\rho) = 1$ if $\rho \in I_1$. The coding of the orbit of a point ρ is the infinite word $\mathbf{s}_{\rho,\alpha}$ obtained by setting its nth, $n \geq 0$, letter to equal $\nu(R^n(\rho))$. This word $\mathbf{s}_{\rho,\alpha}$ is defined to be the Sturmian word of slope α and intercept ρ.

To make the definition proper, we need to define how ν behaves in the endpoints 0 and $1 - \alpha$. We have two options: either take $I_0 = [0, 1 - \alpha)$ and $I_1 = [1 - \alpha, 1)$ or $I_0 = (0, 1 - \alpha]$ and $I_1 = (1 - \alpha, 1]$. The difference is seen in the codings of the orbits of the points $\{-n\alpha\}$. This choice is largely irrelevant in this paper with the exception of the definition of the mapping ψ in the next subsection. The only difference between Sturmian words of slope $[0; 1, a_2, a_3, \ldots]$ and Sturmian words of slope $[0; a_2 + 1, a_3, \ldots]$ is that the roles of the letters 0 and 1 are reversed. We make the typical assumption that $a_1 \geq 2$ in (1). This assumption implies that 00 occurs in Sturmian words of slope α.

Since the sequence $(\{n\alpha\})_{n \geq 0}$ is dense in $[0, 1)$—as is well-known—Sturmian words of slope α have a common language denoted by $\mathcal{L}(\alpha)$. The Sturmian words of slope α form the Sturmian subshift Ω_α, which is minimal and aperiodic. Let w denote a word $a_0 a_1 \cdots a_{n-1}$ of length n in $\mathcal{L}(\alpha)$. Then there exists a unique subinterval $[w]$ of $\mathbb{T}$ such that $\mathbf{s}_{\rho,\alpha}$ begins with w if and only if $\rho \in [w]$. Clearly $[w] = I_{a_0} \cap R^{-1}(I_{a_1}) \cap \ldots \cap R^{-(n-1)}(I_{a_{n-1}})$. The points 0, $\{-\alpha\}$, $\{-2\alpha\}$, $\ldots$, $\{-n\alpha\}$ partition the circle into $n + 1$ subintervals which are in one-to-one correspondence with the words of $\mathcal{L}(\alpha)$ of length n. Arranging these $n + 1$ points into increasing order gives an ordering of the level n intervals: $I_0(n)$, $I_1(n)$, $\ldots$, $I_n(n)$. According to the following proposition, see [1, Proposition 3.2], this ordering of the intervals arranges the associated factors into lexicographic order.

Proposition 2.1. *[1, Proposition 3.2] Let n, i, and j be integers such that $0 \leq i, j \leq n$. Let $u, v \in \mathcal{L}(\alpha)$ be the factors of length n such that $[u] = I_i(n)$ and $[v] = I_j(n)$. Then $u \lhd v$ if and only if $i < j$.*

Let (d_k) be a sequence of positive integers. Corresponding to (d_k), we define a sequence (s_k) of *standard words* by the recurrence

$$s_k = s_{k-1}^{d_k} s_{k-2}$$

with initial values $s_{-1} = 1$, $s_0 = 0$. The sequence (s_k) converges to an infinite word $\mathbf{c}_\alpha$, which is a Sturmian word of intercept α and slope α, where α is an irrational with continued fraction expansion $[0; d_1 + 1, d_2, d_3, \ldots]$. Thus standard words related to the sequence (d_k) are called standard words of slope α. If $d_k = 1$ for all $k \geq 1$, then the associated standard words are called *Fibonacci words*. The standard words are the basic building blocks of Sturmian words, and they have rich and surprising properties. First of all, all standard words are primitive and the final two letters of a (long enough) standard word are different. The words $s_k s_{k-1}$ and $s_{k-1} s_k$ differ only by their final two letters for all $k \geq 1$. Actually, in connection to the square root map, it is more natural to consider reversed standard words obtained by writing standard words from right to left. If s is a standard word in $\mathcal{L}(\alpha)$, then also the reversed standard word $\widetilde{s}$ is in $\mathcal{L}(\alpha)$ because $\mathcal{L}(\alpha)$ is closed under reversal. Thus the preceding property that two consecutive standard words almost commute can be rephrased in our notation as the fact that $\widetilde{s}_k \widetilde{s}_{k-1} = L(\widetilde{s}_{k-1} \widetilde{s}_k)$ for all $k \geq 1$. It is a special property of a reversed standard word $\widetilde{s}$ that $\widetilde{s}$ and $L(\widetilde{s})$ are conjugates, see [5, Proposition 6]. For the above properties and more, see [2, Chapter 2.2].

2.2 Optimal Squareful Words and the Square Root Map

An infinite word is *squareful* if its every position begins with a square. An infinite word is *optimal squareful* if it is aperiodic and squareful and it contains the least possible number of distinct minimal squares. In [10], Kalle Saari proves that optimal squareful words contain six distinct minimal squares; a squareful word containing at most five minimal squares is necessarily ultimately periodic. Moreover, Saari shows that optimal squareful words are binary and that the six minimal squares are of very restricted form. The square roots of the six minimal squares of an optimal

squareful word containing 00 are

$$S_1 = 0, \qquad S_4 = 10^{\mathfrak{a}},$$
$$S_2 = 010^{\mathfrak{a}-1}, \quad S_5 = 10^{\mathfrak{a}+1}(10^{\mathfrak{a}})^{\mathfrak{b}}, \tag{2}$$
$$S_3 = 010^{\mathfrak{a}}, \qquad S_6 = 10^{\mathfrak{a}+1}(10^{\mathfrak{a}})^{\mathfrak{b}+1}.$$

for some integers $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{a} \geq 1$ and $\mathfrak{b} \geq 0$. We call an optimal squareful word containing the *minimal square roots* of (2) an *optimal squareful word with parameters* $\mathfrak{a}$ *and* $\mathfrak{b}$. Throughout this paper, we reserve this meaning for the fraktur letters $\mathfrak{a}$ and $\mathfrak{b}$. Furthermore, we agree that the symbols S_i always refer to the minimal square roots of (2).

Let $\mathbf{s}$ be an optimal squareful word and write it as a product of minimal squares: $\mathbf{s} = X_1^2 X_2^2 \cdots$ (such a product is unique). The *square root* $\sqrt{\mathbf{s}}$ of $\mathbf{s}$ is the word $X_1 X_2 \cdots$ obtained by deleting half of each minimal square X_i^2. We reserve the notation $\sqrt[n]{\mathbf{s}}$ for the nth square root of $\mathbf{s}$. We chose this notation for its simplicity; the nth square root of a number x would typically be denoted by $\sqrt[2^n]{x}$. We often consider square roots of finite words. We let $\Pi(\mathfrak{a}, \mathfrak{b})$ be the language of all nonempty words w such that w is a factor of some optimal squareful word with parameters $\mathfrak{a}$ and $\mathfrak{b}$ and w is factorizable as a product of minimal squares (2). Let $w \in \Pi(\mathfrak{a}, \mathfrak{b})$, that is, $w = X_1^2 \cdots X_n^2$ for minimal square roots X_i. Then we can define the square root $\sqrt{w}$ of w by setting $\sqrt{w} = X_1 \cdots X_n$. The square root map (on infinite words) is continuous with respect to the usual topology on infinite words (see [2, Section 1.2.2.]). The following lemma, used later, sharpens this observation.

Lemma 2.2. *Let $\mathbf{u}$ and $\mathbf{v}$ be two optimal squareful words with the same parameters $\mathfrak{a}$ and $\mathfrak{b}$. If $\mathbf{u}$ and $\mathbf{v}$ have a common prefix of length ℓ, then $\sqrt{\mathbf{u}}$ and $\sqrt{\mathbf{v}}$ have a common prefix of length $\lceil \ell/2 \rceil$.*

Proof. Say $\mathbf{u}$ and $\mathbf{v}$ have a nonempty common prefix w. We may suppose that $w \notin \Pi(\mathfrak{a}, \mathfrak{b})$ as otherwise the claim is clear. Let z be the longest prefix of w that is in $\Pi(\mathfrak{a}, \mathfrak{b}) \cup \{\varepsilon\}$, and let X^2 and Y^2 respectively be the minimal square prefixes of the words $T^{|z|}(\mathbf{u})$ and $T^{|z|}(\mathbf{v})$. Hence $\sqrt{\mathbf{u}}$ begins with $\sqrt{z}X$ and $\sqrt{\mathbf{v}}$ begins with $\sqrt{z}Y$. Since X and Y begin with the same letter, it is easy to see that either X is a prefix of Y or Y is a prefix of X. By symmetry, we suppose that X is a prefix of Y. It follows that $\sqrt{\mathbf{u}}$ and $\sqrt{\mathbf{v}}$ have a common prefix of length $|zX^2|/2$. By the maximality of z, we have $|zX^2| > |w|$ proving that $\sqrt{\mathbf{u}}$ and $\sqrt{\mathbf{v}}$ have a common prefix of length $\lceil |w|/2 \rceil$. $\qquad \square$

Sturmian words form a proper subset of optimal squareful words. If $\mathbf{s}$ is a Sturmian word of slope α having continued fraction expansion as in (1), then it is an optimal squareful word with parameters $\mathfrak{a} = a_1 - 1$ and $\mathfrak{b} = a_2 - 1$. The square root map is especially interesting for Sturmian words because it preserves their languages. Define a function $\psi \colon \mathbb{T} \to \mathbb{T}$ as follows. For $\rho \in (0, 1)$, we set

$$\psi(\rho) = \frac{1}{2}(\rho + 1 - \alpha),$$

and we set

$$\psi(0) = \begin{cases} \frac{1}{2}(1 - \alpha), & \text{if } 0 \in I_0, \\ 1 - \frac{\alpha}{2}, & \text{if } 0 \notin I_0. \end{cases}$$

The mapping ψ moves a point ρ on $\mathbb{T}$ towards the point $1 - \alpha$ by halving the distance between the points ρ and $1 - \alpha$. The distance to $1 - \alpha$ is measured in the interval I_0 or I_1 depending on which of these intervals the point ρ belongs to. In [5], we proved the following result relating the intercepts of a Sturmian word and its square root.

Theorem 2.3. *[5, Theorem 9] Let $\mathbf{s}_{\rho,\alpha}$ be a Sturmian word of slope α. Then $\sqrt{\mathbf{s}_{\rho,\alpha}} = \mathbf{s}_{\psi(\rho),\alpha}$.*

Remark 2.4. Later in the proof of Theorem 3.2, we need a version of Theorem 2.3 for rational slopes. Indeed, Theorem 2.3 is true also if α is rational provided that the continued fraction expansion of α has enough partial quotients. More precisely, the proof of Theorem 2.3 in [5] considers only certain properties of the denominators of the (semi)convergents $q_{2,1}$ and $q_{3,1}$ of α. Thus if the continued fraction expansion of α has at least three partial quotients, then Theorem 2.3 holds. Moreover, this condition on the continued fraction expansion guarantees that the codings of rational rotations of slope α contain the six minimal squares of (2). The condition is equivalent to requiring that the parameter S, defined later, satisfies $|S| > |S_6|$.

Specific solutions to the word equation

$$X_1^2 X_2^2 \cdots X_n^2 = (X_1 X_2 \cdots X_n)^2 \tag{3}$$

in the Sturmian language $\mathcal{L}(\alpha)$ play an important role. We are interested only in the solutions of (3) where all words X_i are minimal square roots (2). Thus we give the following definition.

Definition 2.5. A nonempty word w is a *solution to* (3) if w can be written as a product of minimal square roots $w = X_1 X_2 \cdots X_n$ which satisfy the word equation (3). The solution is *primitive* if w is primitive. The word w is a *solution to* (3) *in a language* $\mathcal{L}$ if w is a solution to (3) and $w^2 \in \mathcal{L}$.

Consider for example the word $S_2 S_1 S_4$ for $\mathfrak{a} = 1$. We have

$$(S_2 S_1 S_4)^2 = (01 \cdot 0 \cdot 10)^2 = 01010 \cdot 01010 = (01)^2 \cdot 0^2 \cdot (10)^2 = S_2^2 S_1^2 S_4^2,$$

so the word $S_2 S_1 S_4$ is a solution to (3).

In [5, Theorem 18], the following result was proved.

Theorem 2.6. *[5, Theorem 18] If $\widetilde{s}$ is a reversed standard word, then the words $\widetilde{s}$ and $L(\widetilde{s})$ are primitive solutions to (3) in the language of some Sturmian word.*

Solutions to (3) are important as they can be used to build fixed points of the square root map. If (u_k) is a sequence of solutions to (3) with the property that u_k^2 is a proper prefix of u_{k+1} for $k \geq 1$, then the infinite word $\mathbf{w}$ obtained as the limit $\lim_{k \to \infty} u_k$ has arbitrarily long prefixes $X_1^2 \cdots X_n^2$ with the property that $X_1 \cdots X_n$ is a prefix of $\mathbf{w}$. In other words, the word $\mathbf{w}$ is a fixed point of the square root map. All known constructions of fixed points rely on this method. For example, the two Sturmian words $01\mathbf{c}_\alpha$ and $10\mathbf{c}_\alpha$ of slope α and intercept $1 - \alpha$ both have arbitrarily long squares u^2 as prefixes, where $u = L(\widetilde{s})$ for a reversed standard word $\widetilde{s}$ [5, Proposition 27]. In the next subsection, we see that the dynamical system studied in this paper is also fundamentally linked to fixed points obtained from solutions of (3).

The following lemma [5, Lemma 21] is of technical nature, but it conveys an important message: under the assumptions of the lemma, swapping two adjacent and distinct letters that do not occur as a prefix of a minimal square affects a product of minimal squares only locally and does not change its square root. This establishes the often-used fact that $\widetilde{s}\widetilde{s}$ and $\widetilde{s}L(\widetilde{s})$ are both in $\Pi(\mathfrak{a}, \mathfrak{b})$ and have the same square root for a reversed standard word $\widetilde{s}$. For example, if $\widetilde{s} = 1001001010010$, then

$$\widetilde{s}\widetilde{s} = 1001001010 \cdot 0101 \cdot 00 \cdot 1001010010 \quad \text{and}$$
$$\widetilde{s}L(\widetilde{s}) = 1001001010 \cdot 010010 \cdot 1001010010,$$

so the change is indeed local and does not affect the square root. Notice that every long enough reversed standard word has S_6 as a proper suffix.

Lemma 2.7. *[5, Lemma 21] Let u and v be words such that*

6

- *u is a nonempty suffix of S_6,*

- *$|v| \geq |S_5 S_6|$,*

- *v begins with xy for distinct letters x and y,*

- *uv and $L(v)$ are factors of some optimal squareful words with the same parameters.*

Suppose there exists a minimal square X^2 such that $|X^2| > |u|$ and X^2 is a prefix of uv or $uL(v)$. Then there exist minimal squares $Y_1^2, \ldots, Y_n^2$ such that X^2 and $Y_1^2 \cdots Y_n^2$ are prefixes of uv and $uL(v)$ of the same length and $X = Y_1 \cdots Y_n$.

2.3 The Subshift Ω

In this subsection, we define the main object of study of this paper. The main results presented in this subsection were obtained in [5] in the case $\mathfrak{c} = 1$; we discuss their generalization in detail.

Let $\mathfrak{c}$ be a fixed positive integer. Repeated application of the substitution

$$\tau: \begin{matrix} S \mapsto LS^{2\mathfrak{c}} \\ L \mapsto S^{2\mathfrak{c}+1} \end{matrix}$$

to the letter S produces two infinite words

$$\Gamma_1^\circ = SS^{2\mathfrak{c}}(LS^{2\mathfrak{c}})^{2\mathfrak{c}}(S^{2\mathfrak{c}+1}(LS^{2\mathfrak{c}})^{2\mathfrak{c}})^{2\mathfrak{c}} \cdots \text{ and}$$
$$\Gamma_2^\circ = LS^{2\mathfrak{c}}(LS^{2\mathfrak{c}})^{2\mathfrak{c}}(S^{2\mathfrak{c}+1}(LS^{2\mathfrak{c}})^{2\mathfrak{c}})^{2\mathfrak{c}} \cdots$$

with the same language $\mathcal{L}$. See Remark 2.13 why we do not consider analogous substitutions of even length. We set Ω° to be the set of all infinite words with factors in $\mathcal{L}$. This subshift Ω° is minimal and aperiodic.[1]

Fix integers $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{a} \geq 1$ and $\mathfrak{b} \geq 0$, and let α be an irrational with continued fraction expansion $[0; \mathfrak{a}+1, \mathfrak{b}+1, \ldots]$. Recall that Sturmian words of slope α are optimal squareful words with parameters $\mathfrak{a}$ and $\mathfrak{b}$. Let w be a word such that $w \in \{\widetilde{s}_t, L(\widetilde{s}_t)\}$ where $\widetilde{s}_t$ is a reversed standard word of slope α such that $t \geq 3$.[2] This condition $t \geq 3$ is equivalent to insisting that $|\widetilde{s}_t| > |S_6|$ (here S_6 is taken with the parameters $\mathfrak{a}$ and $\mathfrak{b}$ from the continued fraction expansion of α) as is directly observed from the definition of standard words. Recall that w and $L(w)$ are conjugate. Let then σ be the substitution mapping S to w and L to $L(w)$. By substituting the letters S and L in words of Ω°, we obtain a new minimal and aperiodic subshift $\sigma(\Omega^\circ)$, which we denote by Ω_A. We also set $\Gamma_1 = \sigma(\Gamma_1^\circ)$ and $\Gamma_2 = \sigma(\Gamma_2^\circ)$. The subshift Ω_A is generated by both words Γ_1 and Γ_2. The words Γ_1 and Γ_2 differ only by their first two letters. This difference is often irrelevant to us, so we let Γ to stand for either of these words. Further, we let the symbol γ_k to stand for the word $\sigma(\tau^k(S))$ and $\overline{\gamma}_k$ to stand for $\sigma(\tau^k(L))$.

It is easy to see that $\Gamma_1 = \lim_{k \to \infty} \gamma_{2k}$ and $\Gamma_2 = \lim_{k \to \infty} \overline{\gamma}_{2k}$. In what follows, we often consider infinite products of γ_k and $\overline{\gamma}_k$, and we wish to argue independently of the index k. Hence we make a convention that γ and $\overline{\gamma}$ respectively stand for γ_k and $\overline{\gamma}_k$ for some $k \geq 0$. The words γ and $\overline{\gamma}$ are primitive; see [5, Lemma 39]. For simplification, we abuse notation and write S for γ_0 and L for $\overline{\gamma}_0$. It will always be clear from context if letters S and L or words S and L are meant.

It can be shown that the words of Ω_A are optimal squareful words with parameters $\mathfrak{a}$ and $\mathfrak{b}$; see [5, Lemma 40]. Therefore the square root map is defined for words in Ω_A. Let us prove the following crucial properties of the square root map on Ω_A.

---

[1]It is minimal since the generating substitution is primitive; see, e.g., [8]. Aperiodicity is well-known; the proof of [5, Lemma 42] is easily adapted for $\mathfrak{c} > 1$.

[2]Without this condition the subshift Ω, defined below, does not consist of optimal squareful words; see the remark after [5, Lemma 40].

Lemma 2.8. *The following properties hold:*

- $\sqrt{\gamma\gamma} = \gamma,$

- $\sqrt{\gamma\overline{\gamma}} = \gamma,$

- $\sqrt{\overline{\gamma}\gamma} = \overline{\gamma},$ *and*

- $\sqrt{\overline{\gamma}\overline{\gamma}} = \overline{\gamma}.$

Proof. This proof is essentially the proof of [5, Proposition 38]. Say $\gamma = \gamma_k$. Suppose first that $k = 0$. Since S was chosen to be in $\{\widetilde{s}, L(\widetilde{s})\}$ for a reversed standard word $\widetilde{s}$, both of the words S and L are primitive solutions to (3) in the same Sturmian language by Theorem 2.6. Therefore $S^2, L^2 \in \Pi(\mathfrak{a}, \mathfrak{b})$ and $\sqrt{SS} = S$ and $\sqrt{LL} = L$. It is also true that $SL, LS \in \Pi(\mathfrak{a}, \mathfrak{b})$ and that $\sqrt{SL} = S$ and $\sqrt{LS} = L$. This can be proven by an application of Lemma 2.7; see the example preceding the lemma. There are however some nontrivial and lengthy details, so we refer the reader to [5, Lemma 22] for the full story. Thus the claim holds for $k = 0$. Suppose that the claim holds for some $k \geq 0$. Now

$$\gamma_{k+1}^2 = \overline{\gamma}\gamma \cdot (\gamma^2)^{\mathfrak{c}-1} \cdot \gamma\overline{\gamma} \cdot (\gamma^2)^{\mathfrak{c}}$$

so, by the induction hypothesis, we obtain

$$\sqrt{\gamma_{k+1}^2} = \overline{\gamma} \cdot \gamma^{\mathfrak{c}-1} \cdot \gamma \cdot \gamma^{\mathfrak{c}} = \gamma_{k+1}.$$

The other cases are verified similarly by grouping the words into suitable pairs. $\qquad\square$

Lemma 2.8 shows that the words $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_2$ are fixed points of the square root map. Namely, the word γ_{k+2} has γ_k^2 as a prefix and $\overline{\gamma}_{k+2}$ has $\overline{\gamma}_k^2$ as a prefix. Thus by Lemma 2.8, we have, e.g.,

$$\sqrt{\boldsymbol{\Gamma}_1} = \sqrt{\lim_{k\to\infty} \gamma_{2k}^2} = \lim_{k\to\infty} \gamma_{2k} = \boldsymbol{\Gamma}_1.$$

The words in Ω_A can be (uniquely) written as a product of the words S and L up to a shift. We often consider infinite words that are arbitrary products of the words S and L (elements of $\{S, L\}^\omega$) and their shifts (elements of the shift orbit closure $\overline{\{S, L\}^\omega}$). Consider a word $\mathbf{w}$ in $\overline{\{S, L\}^\omega}$ and write $\mathbf{w} = T^\ell(\mathbf{w}')$ for some $\mathbf{w}' \in \{S, L\}^\omega$ and ℓ such that $0 \leq \ell < |S|$. There are four distinct possibilities (types):

(A) $\ell = 0,$

(B) $\ell > 0$ and the prefix of $\mathbf{w}$ of length $|S| - \ell$ is in $\Pi(\mathfrak{a}, \mathfrak{b}),$

(C) $\ell > 0$ and the prefix of $\mathbf{w}$ of length $2|S| - \ell$ is in $\Pi(\mathfrak{a}, \mathfrak{b}),$ or

(D) none of the above apply.

These possibilities are mutually exclusive: cases (B) and (C) cannot simultaneously apply because $S, L \notin \Pi(\mathfrak{a}, \mathfrak{b})$. Indeed, if S (or L) was a product of minimal squares, then it would not be primitive due to the fact that it is a solution to (3).

As the reader might have observed, it is possible that the integer ℓ is not unique, i.e., the type of a word is not well-defined. By the following lemma, this happens only in degenerate cases, so we largely ignore this possibility. It concerns us only briefly in the proofs of Lemma 3.5 and Theorem 3.8; for words in Ω_A, the shift ℓ is always unique.

Lemma 2.9. *Let* $\mathbf{w} \in \overline{\{S, L\}^\omega}$, *and write* $\mathbf{w} = T^\ell(\mathbf{u}) = T^{\ell'}(\mathbf{u}')$ *for some words* $\mathbf{u}, \mathbf{u}' \in \{S, L\}^\omega$ *and integers* ℓ *and* ℓ' *such that* $0 \le \ell, \ell' < |S|$. *If* $\ell \ne \ell'$, *then* $\mathbf{w}$ *is periodic with minimal period conjugate to* S.

Proof. Suppose that $\ell \ne \ell'$. Assume without loss of generality that $\ell < \ell'$, and suppose that $\mathbf{u} = \diamond_1\diamond_2\cdots$ and $\mathbf{u}' = \diamond'_1\diamond'_2\cdots$ with $\diamond_i, \diamond'_i \in \{S, L\}$ for all $i \ge 1$. As $\mathbf{w} = T^\ell(\mathbf{u}) = T^{\ell'}(\mathbf{u}')$, the word $T^{|S|-\ell'}(\mathbf{w})$ has the word $\diamond'_2\diamond'_3$ as a prefix, and this prefix has the word $\diamond_2$ occurring at position $\ell' - \ell$. By repeating the argument, we see that $\diamond_i$ occurs at position $\ell' - \ell$ of $\diamond'_i\diamond'_{i+1}$ for all $i \ge 2$. Symmetrically, the word $\diamond'_i$ occurs at position $|S| - (\ell' - \ell)$ of $\diamond_{i-1}\diamond_i$ for all $i \ge 3$. Let us first argue that $\ell' - \ell \ne 1$. Assume that $\ell' - \ell = 1$. Then the common suffix v of $\diamond_2$ and $\diamond'_2$ of length $|S| - 2$ occurs at consecutive positions 2 and 3 of $\diamond'_2\diamond'_3$. This means that v is unary: it is a power of its first letter. By our assumption $|S| > |S_6|$, both S and L have S_6 as a suffix. This is a contradiction since S_6 is not unary. Hence $\ell' - \ell \ne 1$. The letter of $\diamond'_i\diamond'_{i+1}$ at position $\ell' - \ell$ determines $\diamond_i$ uniquely. The words S and L differ by only their first two letters so, as $\ell' - \ell \ne 1$, we see that the determining letter is the same for all $i \ge 2$, that is, $\diamond_i = \diamond_{i+1}$ for all $i \ge 2$. The prefix of $\mathbf{w}$ of length $|S| - \ell$ is a suffix of S or L. Thus if $\ell > 1$, this prefix is a suffix of both S and L and, in particular, the word $\mathbf{w}$ is periodic with minimal period conjugate to $\diamond_2$. The conclusion follows as S and L are conjugate. Suppose finally that $\ell = 1$. This means that $|S| - (\ell' - \ell) > 1$, so we can deduce that $\diamond'_i = \diamond'_{i+1}$ for all $i \ge 3$. The first letter of the word $\diamond'_2$ is determined by the letter at position $|S| - \ell'$ of $\mathbf{w}$. Since $\ell' > \ell = 1$, this letter must equal the letter at position $|S| - (\ell' - \ell)$ of $\diamond_2\diamond_3$. Therefore $\diamond'_2 = \diamond'_3$. Like previously, the prefix of $\mathbf{w}$ of length $|S| - \ell'$ is a suffix of S or L. Because $\ell' > 1$, we see that this prefix is a suffix of $\diamond'_2$. We conclude that $\mathbf{w}$ is periodic with minimal period conjugate to $\diamond'_2$. $\qquad\square$

Let us then recall two important results from our earlier paper [5]. In [5], we proved these results with slightly different statements and only in the case $\mathfrak{c} = 1$. See Remark 2.12 after the theorems for discussion why they are true also for $\mathfrak{c} > 1$.

Theorem 2.10. *Let* $\mathbf{w} \in \Omega_A$. *If* $\mathbf{w}$ *is of type* (A), (B), *or* (C), *then* $\sqrt{\mathbf{w}} \in \Omega_A$. *If* $\mathbf{w}$ *is of type* (D), *then* $\sqrt{\mathbf{w}}$ *is periodic with minimal period conjugate to* S.

Theorem 2.11. *Let* $\mathbf{w} \in \overline{\{S, L\}^\omega}$. *If* $\mathbf{w}$ *is of type* (D), *then* $\sqrt{\mathbf{w}}$ *is periodic with minimal period conjugate to* S.

Remark 2.12. This remark may be skipped for we only justify here why our earlier proofs almost directly work in the setting of the current paper—nothing in this remark is later referred to.

We claim that the proof of [5, Theorem 44] works as a proof of Theorem 2.10 with almost no modification. We do not provide the full proof, or even its sketch, here because the proof is rather long. The statement of [5, Theorem 44] differs from that of Theorem 2.10: it states that in the case $\mathfrak{c} = 1$ for all $\mathbf{w} \in \Omega_A$ either $\sqrt{\mathbf{w}} \in \Omega_A$ or $\sqrt{\mathbf{w}}$ is periodic with minimal period conjugate to S. The statement of Theorem 2.10 above is more precise, but this more precise version and Theorem 2.11 can be obtained by a careful reading of the proof of [5, Theorem 44], see below.

Let us first point out that the important technical lemmas used in the proof of [5, Theorem 44], namely lemmas 48, 50, 51, and 52, depend only on the combinatorial properties of standard words, not on the parameter $\mathfrak{c}$. Secondly, and more importantly, the proof does not use any structural properties of the words in Ω_A, the arguments go through for arbitrary products of the words S and L. The crucial part is where Lemma 45 is invoked. This lemma states that if a word $\mathbf{w}$ in Ω_A is a product of the words S and L, then $\sqrt{\mathbf{w}} \in \Omega_A$. It can be verified that for the proof to work, it is necessary that a fixed point exists which is built as a product of words γ and $\overline{\gamma}$ whose length is an odd multiple of $|S|$, just as it is the case in this paper. Equally important is the use of Lemma 47. Its statement supposes the word $\mathbf{w}$ to be in Ω_A, but the proof does not use

this assumption, so the claim is true for arbitrary products of S and L. These arguments, while succinct, show that the proof of [5, Theorem 44] goes through in the case $\mathfrak{c} > 1$.

Let us then give a brief guide into reading the proof of [5, Theorem 44] to deduce the statements of Theorems 2.10 and 2.11—the earlier paper [5] uses a different vocabulary. The proof of [5, Theorem 44] is divided into several claims. The first claim, Claim 53, says in the current vocabulary that if $\mathbf{w}$ is of type (B), then $\sqrt{\mathbf{w}} \in \Omega_A$. Claim 55 shows that if $\mathbf{w}$ is of type (C), then $\sqrt{\mathbf{w}} \in \Omega_A$. Lemma 45 (mentioned above) shows that the square root of a word of type (A) is also in Ω_A. Most effort in the proof is put into showing that if the word in question is not of type (A), (B), (C), then there exists a so-called nicely repetitive position (Claims 54 and 57). The existence of such a position shows by Lemma 47 that $\sqrt{\mathbf{w}}$ is periodic with minimal period conjugate to S. This reading of the proof gives the statement of Theorem 2.10. The part showing the existence of a nicely repetitive position applies to arbitrary words in $\overline{\{S, L\}}^\omega$, so we also obtain Theorem 2.11. Notice that by Lemma 2.9 the type of a word in Ω_A is always well-defined. If the type of a word $\mathbf{w}$ is not well-defined, then $\mathbf{w}$ is periodic and the claim of Theorem 2.11 holds anyway.

In order to make Ω_A equipped with the square root map a genuine dynamical system, we need to adjoin a periodic part to it. To this end, we let

$$\Omega_P = \{T^\ell(S^\omega) \colon 0 \leq \ell < |S|\},$$

and define $\Omega = \Omega_A \cup \Omega_P$. Related to Ω_P, we observe the following. Suppose that S or L equals $\widetilde{s}_k$, the kth reversed standard word of slope α, where $\alpha = [0; a_1, a_2, \ldots]$. Let us truncate the continued fraction expansion of α and set $\bar{\alpha} = [0; a_1, a_2, \ldots, a_k]$, so that $\bar{\alpha} = p_k/q_k$ with $|S| = q_k$. Since $\{-q_k\bar{\alpha}\} = 0$, we have that Ω_P equals the codings of rational rotations of slope $\bar{\alpha}$. (See the proof of Theorem 4.3 and Example 4.4 in [3] for the exact details.)

Clearly Ω is compact with respect to the usual product topology and $\sqrt{\Omega_A} \subseteq \Omega$ by Theorem 2.10. On the other hand, for any $w \in \Omega_P$, we have $\sqrt{w} \in \Omega_P$ as noted in Remark 2.4 (recall that S was chosen so that it satisfies $|S| > |S_6|$). Thus $\sqrt{\Omega_P} \subseteq \Omega_P$, and the pair $(\Omega, \sqrt{\cdot})$ is a valid dynamical system. Notice further that $L^\omega \in \Omega_P$ since S and L are conjugate.

Remark 2.13. Let us explain why we do not consider substitutions $\tau \colon S \to LS^{2\mathfrak{c}+1}, L \to S^{2\mathfrak{c}+2}$ for $\mathfrak{c} \geq 0$. The problem is that $\tau^k(S)$ is not a solution to (3) for $k \geq 1$, that is, Lemma 2.8 does not hold. Consider for instance the case $k = 1$. Now $\tau^k(S) = LS^{2\mathfrak{c}+1}$, so $\sqrt{\tau^k(S)^2} = LS^{\mathfrak{c}}LS^{\mathfrak{c}} \neq \tau^k(S)$. This means that Theorem 2.10 is not true either. It is readily checked that between two letters L there is always $S^{2\mathfrak{c}+1}$ or $S^{4\mathfrak{c}+3}$ analogously to Lemma 2.14. However $\sqrt{\tau^k(S)^2}$ contains $LS^{\mathfrak{c}}L$, so the square root map does not preserve the languages of words in $\Omega_A \cap \{S, L\}^\omega$. To obtain a good dynamical system, we need to consider substitutions with odd length.

Let us recall next what is known about the structure of the words in Ω. The word $\mathbf{\Gamma}$ is by definition an infinite product of the words γ_k and $\overline{\gamma}_k$ for all $k \geq 0$. Thus all words in Ω_A are (uniquely) factorizable as products of γ_k and $\overline{\gamma}_k$ up to a shift. Let us for convenience denote by Ω_γ the set $\Omega \cap \{\gamma, \overline{\gamma}\}^\omega$ consisting of words of Ω that are infinite products of γ and $\overline{\gamma}$. Notice that $\Omega_S = \Omega_{\gamma_0}$ by our convention. The following lemma describes two important properties of factorizations of words of Ω_A as products of γ and $\overline{\gamma}$. This result is an immediate property of the substitution τ that generates Ω°.

Lemma 2.14. *Consider a factorization of a word in $\Omega_A \cap \Omega_\gamma$ as a product of γ and $\overline{\gamma}$. Such factorization has the following properties:*

(i) *Between two occurrences of $\overline{\gamma}$ there is always $\gamma^{2\mathfrak{c}}$ or $\gamma^{4\mathfrak{c}+1}$.*

(ii) *Between two occurrences of $\overline{\gamma}\gamma^{4\mathfrak{c}+1}\overline{\gamma}$ there is always $(\gamma^{2\mathfrak{c}}\overline{\gamma})^{2\mathfrak{c}-1} \cdot \overline{\gamma}^{-1}$ or $(\gamma^{2\mathfrak{c}}\overline{\gamma})^{4\mathfrak{c}} \cdot \overline{\gamma}^{-1}$.*

We also need to know how certain factors synchronize or align in a product of γ and $\overline{\gamma}$.

Lemma 2.15 (Synchronizability Properties). *Let $\mathbf{w} \in \Omega_\gamma$. If z is a word in $\{\gamma\gamma, \gamma\overline{\gamma}, \overline{\gamma}\gamma\}$ occurring at position ℓ of $\mathbf{w}$, then the prefix of $\mathbf{w}$ of length ℓ is a product of γ and $\overline{\gamma}$.*[3]

Proof. Suppose that the claim is false, that is, there exists a word $\diamond_1\diamond_2\diamond_3$, with $\diamond_1, \diamond_2, \diamond_3 \in \{\gamma, \overline{\gamma}\}$, occurring at position $r|\gamma|$ of $\mathbf{w}$ for some r such that the word z occurs at position t, $0 < t < |\gamma|$, of $\diamond_1\diamond_2\diamond_3$. This is not possible unless one of the words $\diamond_1$, $\diamond_2$, and $\diamond_3$ equals $\overline{\gamma}$ since γ is primitive and a primitive word cannot occur nontrivially in its square (as is well-known). Assume first that $\diamond_1 = \overline{\gamma}$. Then Lemma 2.14 (i) implies that $\diamond_2 = \diamond_3 = \gamma$. Since γ is primitive, it must be that $z = \gamma\overline{\gamma}$. Recall that the words γ and $\overline{\gamma}$ differ by only their first two letters. Thus if $t > 1$, then the prefix γ of z actually starts at position t of γ^2 which is impossible. Hence $t = 1$. Now $\gamma = abu$ and $\overline{\gamma} = bau$ for some distinct letters a and b and for some nonempty word u. As z occurs at position 1 of $\overline{\gamma}\gamma^2$, we see that $\overline{\gamma}$ begins with bab. The suffix $\overline{\gamma}$ of z occurs at position $1 + |\gamma|$ of $\overline{\gamma}\gamma^2$, so analogously γ must begin by aba. This is a contradiction as $a \neq b$.

Suppose next that $\diamond_3 = \overline{\gamma}$. By using Lemma 2.14 (i), we conclude that $\diamond_1 = \diamond_2 = \gamma$ and, similarly to above, we see that it must be that $z = \overline{\gamma}\gamma$. This means that the suffix of $\overline{\gamma}$ of length t is a prefix of γ and the suffix of γ of length t is a prefix of $\overline{\gamma}$. As γ and $\overline{\gamma}$ have no common prefix and the suffixes of γ and $\overline{\gamma}$ agree up to length $|\gamma| - 2$, it must be that $t = |\gamma| - 1$. However, this means that γ ends with the first letters of γ and $\overline{\gamma}$. This is a contradiction as these letters are distinct.

Consider the final case that $\diamond_1 = \diamond_3 = \gamma$ and $\diamond_2 = \overline{\gamma}$. The same happens as in the second case above: the suffix of length t of the first half of z is a prefix of $\overline{\gamma}$ and the suffix of z of length t is a prefix of γ. This is clearly impossible if $z = \gamma^2$. If $z \neq \gamma^2$, then this leads to the same contradiction as in the second case above. $\qquad\square$

The preceding lemma shows that if $\mathbf{w}$ is a word in Ω_A, then for each k there exists a unique ℓ such that $0 \leq \ell < |\gamma_k|$ and $T^\ell(\mathbf{w}) \in \Omega_{\gamma_k}$. We then say that the γ_k-factorization of $\mathbf{w}$ starts at the position ℓ of $\mathbf{w}$.

Let us conclude this subsection by making a remark regarding the subshift Ω°. It is possible to define a counterpart for the square root map of Ω. Write a word $\mathbf{w}$ of Ω° as a product of pairs of the letters S and L: $\mathbf{w} = X_1 X_1' \cdot X_2 X_2' \cdots$, where $X_i X_i' \in \{SS, SL, LS, LL\}$. We define the square root $\sqrt{\mathbf{w}}$ of $\mathbf{w}$ to be the word $X_1 X_2 \cdots$. By Lemma 2.8, it is not difficult to see that $\sigma(\sqrt{\mathbf{w}}) = \sqrt{\sigma(\mathbf{w})}$ for $\mathbf{w} \in \Omega^\circ$. In other words, the square root map for words in $\Omega_S \cap \Omega_A$ has the same dynamics as the square root map in Ω°. This connection is why we abuse notation and write S for γ_0 and L for $\overline{\gamma_0}$.

3 The Limit Set, Invariant Subsets, and Injectivity

In this section, we consider what happens for words of Ω when the square root map is iterated. We extend Theorem 2.11 and show that also the words of type (B) and type (C) are eventually mapped to a periodic word (Theorem 3.2). In fact, we prove a stronger result: the number of steps required is bounded by a constant depending only on the word S. These results enable us to characterize the limit set of $(\Omega, \sqrt{\cdot})$ as the set Ω_S (Theorem 3.10). In other words, asymptotically the square root map on Ω has the same dynamics as the counterpart mapping on $\Omega^\circ \cup \{S^\omega, L^\omega\}$. We further study invariant subsets and show that there are infinitely many of them (Proposition 3.12). We also show that the square root map is mostly injective on Ω_A, only certain left extensions of $\boldsymbol{\Gamma}$ may have two preimages (Theorem 3.8).

Let us first look at an example.

---

[3]In general, e.g., the word γ^2 can be a factor of $\overline{\gamma}^3$.

$\lvert S \rvert$	8	13	21	34	55	89	144	233	377	610	987	1597	2584	4181	6765
n	3	4	4	5	6	6	7	8	8	9	10	10	11	12	13

Table 1: How $\lvert S \rvert$ and n of Theorem 3.2 relate when S is a reversed Fibonacci word.

Example 3.1. Let $\mathfrak{a} = 1$, $\mathfrak{b} = 0$, and $S = 01010010$. Set $\mathbf{w} = T^4(S^2\mathbf{u})$ for some $S^2\mathbf{u} \in \Omega_S \cap \Omega_A$. The word $\mathbf{w}$ is of type (C) as the word $T^4(S^2)$, which equals $00 \cdot 1001010010$, is in $\Pi(\mathfrak{a}, \mathfrak{b})$. Now $\sqrt{\mathbf{w}} = 010010 \cdot \sqrt{\mathbf{u}}$ and $\sqrt{\mathbf{w}} \in \Omega_A$ by Theorem 2.10. So $\sqrt{\mathbf{w}}$ is of type (B), and $\sqrt[2]{\mathbf{w}} = 010 \cdot \sqrt[2]{\mathbf{u}}$. Still we have $\sqrt[2]{\mathbf{w}} \in \Omega_A$. It is clear now that $\sqrt[2]{\mathbf{w}}$ is not of type (A) or (B). The word $\sqrt[2]{\mathbf{u}}$ begins with S or L, and neither $010 \cdot S$ nor $010 \cdot L$ is in $\Pi(\mathfrak{a}, \mathfrak{b})$, so $\sqrt[2]{\mathbf{w}}$ is not of type (C) either. Thus it is of type (D), so $\sqrt[3]{\mathbf{w}}$ is periodic with period of length $\lvert S \rvert$. The minimal period of $\sqrt[3]{\mathbf{w}}$ is readily checked to be 01010010, that is, $\sqrt[3]{\mathbf{w}} = S^{\omega}$. With some effort it can be verified that in this particular case $\sqrt[3]{\mathbf{v}}$ is periodic for all $\mathbf{v} \in \Omega \setminus \Omega_S$.

Notice that the parameter $\mathfrak{c}$ is irrelevant to all of the arguments in the above example. Notice also that the word $\mathbf{u}$ did not play any special role here, and it could have been any product of the words S and L. Indeed, we formulate the next result for arbitrary products of the words S and L.

Theorem 3.2. *There exists an integer n, depending only on the word S, such that $\sqrt[n]{\mathbf{w}} \in \{S^{\omega}, L^{\omega}\}$ for all $\mathbf{w} \in \overline{\{S, L\}^{\omega}} \setminus \{S, L\}^{\omega}$.*

Before proceeding to prove the theorem, let us remark that the number n in the statement indeed varies when S varies. Let S be a reversed Fibonacci word. Table 1 shows how $\lvert S \rvert$ relates to n. It seems that here $n \to \infty$ as $\lvert S \rvert \to \infty$ even though the growth is slow. We have not attempted seriously to relate $\mathfrak{a}$, $\mathfrak{b}$, S, and n, but we conjecture that n is close to the bound obtained in Remark 3.6; see the final section on open problems.

For the proof, we need three lemmas. The first lemma is the important Embedding Lemma.

Lemma 3.3 (Embedding Lemma). *Let $\mathbf{w} \in \overline{\{S, L\}^{\omega}}$ and u_1 and u_2 respectively be the prefixes of $\mathbf{w}$ and $\sqrt{\mathbf{w}}$ of length $\lvert S \rvert$.*

(i) *If $\mathbf{w}$ begins with 0 and $u_1 \neq u_2$, then $u_1 \lhd u_2$.*

(ii) *If $\mathbf{w}$ begins with 1 and $u_1 \neq u_2$, then $u_1 \rhd u_2$.*

Proof. Suppose that $u_1 \neq u_2$, and let v be the prefix of $\mathbf{w}$ of length $2\lvert S \rvert$. The main idea of this proof is the following idea of embedding: we slightly modify the prefix v so that it belongs to a Sturmian language and the square root of a Sturmian word beginning with this modified prefix has u_2 as a prefix. Then known properties of the lexicographic ordering of factors of Sturmian words together with the dynamics of the function ψ prove the claim. More precisely, we want to find a word v' with the following properties: v' belongs to a Sturmian language $\mathcal{L}(\alpha)$, v' has u_1 as a prefix, $\lvert v \rvert = \lvert v' \rvert$, and the square root of any Sturmian word of slope α beginning with v' has u_2 as a prefix. Once we establish the existence of a word v' with such properties, the claim is then proved as follows. Let $\rho \in [v']$, and consider the word $\mathbf{s}_{\rho,\alpha}$ of intercept ρ and slope α. By Theorem 2.3, the intercept of the word $\sqrt{\mathbf{s}}$ is $\psi(\rho)$. Since $\sqrt{\mathbf{s}}$ has u_2 as a prefix, we see that $\psi(\rho) \in [u_2]$. Because ψ moves points towards the point $1 - \alpha$ and $u_1 \neq u_2$, we see that the interval $[u_2]$ is strictly closer to $1 - \alpha$ than the interval $[u_1]$ is. It thus follows from Proposition 2.1 that $u_1 \lhd u_2$ if $\mathbf{w}$ begins with 0 and $u_1 \rhd u_2$ if $\mathbf{w}$ begins with 1.

We shall fix v' later on; for now, we construct two Sturmian languages in at least one of which the word v' occurs. Suppose that S or L equals $\widetilde{s}_k$, the kth reversed standard word of slope α, where $\alpha = [0; a_1, a_2, \ldots]$. Without loss of generality we may assume $a_{k+1} > 1$. Recall that the

assumption $|S| > |S_6|$ implies that $k \geq 3$. We modify α to obtain two distinct slopes $[0; b_1, b_2, \ldots]$ and $[0; c_1, c_2, \ldots]$, respectively denoted by α_1 and α_2. We set $b_i = a_i$ and $c_i = a_i$ for $1 \leq i < k - 1$. Further we set $b_{k-1} = a_{k-1}$, $b_k = a_k$, and $b_{k+1} \geq 3$; the remaining partial quotients may be chosen arbitrarily. For α_2, we set $c_{k-1} = a_{k-1} + 1$ if $a_k = 1$; otherwise we let $c_{k-1} = a_{k-1}$, $c_k = a_k - 1$, $c_{k+1} = 2$, and $c_{k+2} = 2$. Again, the remaining partial quotients are irrelevant. In the case of slope α_1, by recalling that $\widetilde{s}_i \widetilde{s}_{i+1} = L(\widetilde{s}_{i+1} \widetilde{s}_i)$ for all $i \geq 0$, we see that

$$\widetilde{s}_{k+1} = \widetilde{s}_{k-1} \widetilde{s}_k^{b_{k+1}} = L(\widetilde{s}_k)^{b_{k+1}} \widetilde{s}_{k-1},$$

so $S^3, L^3 \in \mathcal{L}(\alpha_1)$. It is straightforward to show that $\widetilde{s}_{k+1}^2 \in \mathcal{L}(\alpha_1)$. Now

$$\widetilde{s}_{k+1}^2 = (\widetilde{s}_{k-1} \widetilde{s}_k^{b_{k+1}})^2 = \widetilde{s}_{k-1} \cdot \widetilde{s}_k^{b_{k+1}} L(\widetilde{s}_k) \cdot \widetilde{s}_{k-1} \widetilde{s}_k^{b_{k+1}-1},$$

so also $SSL \in \mathcal{L}(\alpha_1)$. Next we want to show that $LLS \in \mathcal{L}(\alpha_2)$. Suppose first that $a_k = 1$. Then $S = \widetilde{s}_{k-2} \widetilde{s}_{k-1}$. The $(k-1)^{\text{th}}$ reversed standard word of slope α_2 now equals L:

$$\widetilde{s}_{k-1} = \widetilde{s}_{k-3} \widetilde{s}_{k-2}^{c_{k-1}} = L(\widetilde{s}_{k-2}) \widetilde{s}_{k-3} \widetilde{s}_{k-2}^{a_{k-1}} = L(\widetilde{s}_{k-2}) \widetilde{s}_{k-1} = L(S).$$

Thus the above argument showing that $SSL \in \mathcal{L}(\alpha_1)$ now shows that $LLS \in \mathcal{L}(\alpha_2)$. Consider then the case $a_k > 1$. Let e be the $(k+2)^{\text{th}}$ reversed standard word of slope α_2. By expanding it as a product of the $(k-1)^{\text{th}}$ and $(k-2)^{\text{th}}$ reversed standard words, we see that

$$e = \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k} (\widetilde{s}_{k-1}(\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2)^2.$$

As $e^2 \in \mathcal{L}(\alpha_2)$, it follows that $\widetilde{s}_{k-1} e \in \mathcal{L}(\alpha_2)$. Now

$$\begin{aligned}
\widetilde{s}_{k-1} e &= \widetilde{s}_{k-1} \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k} \cdot \widetilde{s}_{k-1} (\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2 \cdot \widetilde{s}_{k-1} (\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2 \\
&= \widetilde{s}_{k-1} \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k+1} \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k} \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k+1} \cdot (\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2 \\
&= L(\widetilde{s}_{k-2}) \widetilde{s}_{k-1}^{c_k+1} \cdot L(\widetilde{s}_{k-2}) \widetilde{s}_{k-1}^{c_k+1} \cdot \widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k+1} \cdot (\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2 \\
&= L(S) L(S) S \cdot (\widetilde{s}_{k-2} \widetilde{s}_{k-1}^{c_k})^2,
\end{aligned}$$

so $LLS \in \mathcal{L}(\alpha_2)$.

We now define v' needed for the conclusion of the proof. By Lemma 2.8, the assumption $u_1 \neq u_2$ implies that $\mathbf{w}$ is not in $\{S, L\}^\omega$. Write $\mathbf{w} = T^\ell(\mathbf{w}')$ for some $\mathbf{w}' \in \{S, L\}^\omega$ and integer ℓ such that $0 < \ell < |S|$. Moreover, for $t \geq 1$, we set $\diamond_t$ to be the prefix of $T^{(t-1)|S|}(\mathbf{w}')$ of length $|S|$; notice that $\diamond_t \in \{S, L\}$ for all $t \geq 1$. For now we make the assumption that $\ell > 1$; the case $\ell = 1$ is handled at the end of the proof. As the words S and L differ only by their first two letters, this additional assumption gives us the freedom to substitute $\diamond_1$ by either of the words S and L without affecting u_1 neither the prefix of $\sqrt{\mathbf{w}}$ of length $|S|$. So we substitute $\diamond_1$ by $\diamond_2$, and select v' to be the prefix of $T^\ell(\diamond_2 \diamond_2 \diamond_3)$ of length $2|S|$. Observe that $\diamond_2 \diamond_2 \diamond_3 \in \{SSS, SSL, LLS, LLL\}$, so either $v' \in \mathcal{L}(\alpha_1)$ or $v' \in \mathcal{L}(\alpha_2)$. Clearly v' has u_1 as a prefix and $|v'| = |v|$. Lemma 2.2 implies that the square root of a Sturmian word having v' as a prefix has u_2 as a prefix. Thus the word v' has the desired properties.

Consider finally the case $\ell = 1$. From the definition of standard words, it is straightforward to show that the only possible minimal square prefixes of S and L are S_2^2 and S_5^2. Consequently, the word u_1 has $10^a 10^{a-1}$ or $0^{a+1}(10^a)^b 10^{a+1}(10^a)^b$ as a prefix. In the latter case, the word u_2 clearly has $0^{\lceil (a+1)/2 \rceil} 1$ as a prefix, so $u_1 \lhd u_2$. Consider then the former case, and let t be the largest integer such that $(10^a)^t$ is a prefix of u_1. With some effort, it can be shown that u_2 then has $(10^a)^{\lceil t/2 \rceil} 10^{a+1}$ as a prefix, that is, $u_1 \rhd u_2$. $\qquad\square$

Lemma 3.4. *Let w be any of the words SS, SL, LS, or LL. If ℓ is an odd integer such that $0 < \ell < |S|$, then $T^\ell(w) \notin \Pi(\mathfrak{a}, \mathfrak{b})$.*

13

Proof. Let ℓ be an odd integer such that $0 < \ell < |S|$. Since $|T^\ell(w)| = |S^2| - \ell$, we see that $|T^\ell(w)|$ is odd. Thus it is impossible that $T^\ell(w) \in \Pi(\mathfrak{a}, \mathfrak{b})$. $\qquad\square$

Using the two preceding lemmas, we can prove the next crucial result.

Lemma 3.5. *Let $\mathbf{w} \in \overline{\{S, L\}^\omega} \setminus \{S, L\}^\omega$ and u_1, u_2, and u_3 to respectively be the prefixes of $\mathbf{w}$, $\sqrt{\mathbf{w}}$, and $\sqrt[2]{\mathbf{w}}$ of length $|S|$.*

 (i) If $\mathbf{w}$ begins with 0, then one of the following holds: $u_2 \rhd u_1$, $u_3 \rhd u_1$, or $\sqrt[2]{\mathbf{w}}$ is periodic.

 (ii) If $\mathbf{w}$ begins with 1, then one of the following holds: $u_2 \lhd u_1$, $u_3 \lhd u_1$, or $\sqrt[2]{\mathbf{w}}$ is periodic.

Proof. Let $\mathbf{w} \in \overline{\{S, L\}^\omega} \setminus \{S, L\}^\omega$, and write $\mathbf{w} = T^\ell(\mathbf{w}')$ for some $\mathbf{w}' \in \{S, L\}^\omega$ and integer ℓ such that $0 < \ell < |S|$. Let u_1, u_2, and u_3 respectively be the prefixes of $\mathbf{w}$, $\sqrt{\mathbf{w}}$, and $\sqrt[2]{\mathbf{w}}$ of length $|S|$. Moreover, for $t \geq 1$, we set $\diamond_t$ to be the prefix of $T^{(t-1)|S|}(\mathbf{w}')$ of length $|S|$; notice that $\diamond_t \in \{S, L\}$ for all $t \geq 1$. We suppose for simplicity that $\mathbf{w}$ begins with the letter 0. The proof in the case that the first letter of $\mathbf{w}$ is 1 is the same proof with the lexicographic orderings reversed. Since words of type (D) map to periodic words by Theorem 2.11, we only need to consider words of type (B) and (C). Notice that the type of a word $\mathbf{u}$ in $\overline{\{S, L\}^\omega}$ is not necessarily well-defined. Lemma 2.9 shows that if the type is not well-defined, then $\mathbf{u}$ is periodic and the claim holds. Thus we may assume that the type is well-defined for $\mathbf{w}$, $\sqrt{\mathbf{w}}$, and $\sqrt[2]{\mathbf{w}}$.

Suppose first that $\mathbf{w}$ is of type (B). Our aim is to show that $u_2 \rhd u_1$. By definition, we have $\mathbf{w} = X_1^2 \cdots X_n^2 \cdot T^{|S|}(\mathbf{w}')$ for some minimal squares X_1^2, ..., X_n^2. Observe that $\sqrt{\mathbf{w}}$ has the word $X_1 \cdots X_n \diamond_2$ as a prefix; here we have used Lemma 2.8. If $\ell = 1$, then the argument in the last paragraph of the proof of the Embedding Lemma implies that $u_2 \rhd u_1$. Say $\ell > 1$. Now u_1 is a conjugate of $\diamond_2$ occurring at position ℓ of $\diamond_2 \diamond_2$. If $u_1 = u_2$, then u_1 also occurs at the position $\ell + |X_1 \cdots X_n|$ of $\diamond_2 \diamond_2$. This is not possible as $\diamond_2$ is primitive, so it cannot occur nontrivially in its square. Thus we conclude that $u_1 \neq u_2$. By the Embedding Lemma, we see that $u_2 \rhd u_1$.

Suppose then that $\mathbf{w}$ is of type (C). If $\sqrt{\mathbf{w}}$ is of type (B) then, by applying the arguments of the preceding paragraph to $\sqrt{\mathbf{w}}$, we see that $u_3 \rhd u_2$. If $u_1 = u_2$, then $u_3 \rhd u_1$. Otherwise $u_1 \neq u_2$ and $u_2 \rhd u_1$ by the Embedding Lemma. If $\sqrt{\mathbf{w}}$ is of type (D), then $\sqrt[2]{\mathbf{w}}$ is periodic. Thus we may focus on the case that $\sqrt{\mathbf{w}}$ is also of type (C). Write $\sqrt{\mathbf{w}} = T^{\ell'}(\mathbf{w}'')$ for some word $\mathbf{w}''$ in $\{S, L\}^\omega$ and integer ℓ' such that $0 < \ell' < |S|$. Since the type of $\sqrt{\mathbf{w}}$ is well-defined, we have $\ell' = \ell/2$. Because $\sqrt{\mathbf{w}}$ is of type (C), Lemma 3.4 implies that $\ell/2 > 1$, that is, $\ell \geq 4$. We suppose that $u_1 = u_2$; otherwise $u_2 \rhd u_1$ by the Embedding Lemma. Since $u_1 = u_2$, the word u_1 occurs at position ℓ of $\diamond_2 \diamond_2$ and at position $\ell/2$ of $\diamond_3 \diamond_3$. Since $\diamond_2$ is primitive, we see that necessarily $\diamond_2 \neq \diamond_3$. For now, we make the additional assumption that $\ell \neq 4$. Suppose next on the contrary that $u_3 = u_2 = u_1$. Since $\mathbf{w}$ and $\sqrt{\mathbf{w}}$ are of type (C) and $\ell \neq 4$, it follows that u_1 occurs at position $\ell/4$ of $\diamond_5 \diamond_5$. Now either $\diamond_5 = \diamond_2$ or $\diamond_5 = \diamond_3$, so either u_1 respectively occurs at positions ℓ and $\ell/4$ of $\diamond_2 \diamond_2$ or u_1 respectively occurs at positions $\ell/2$ and $\ell/4$ of $\diamond_3 \diamond_3$. This contradicts the primitivity of $\diamond_2$ and $\diamond_3$. We conclude that $u_3 \neq u_2$, so $u_3 \rhd u_1$ by the Embedding Lemma. What is left is to consider the case $\ell = 4$. Write $\diamond_2 = abcdv$ and $\diamond_3 = bacdv$ for a word v and letters a, b, c, and d such that $a \neq b$. Since u_1 occurs at position 4 of $\diamond_2 \diamond_2$, we see that u_1 has $abcd$ as a suffix. Further, we see that u_1 has cd as a prefix and ba as a suffix because it occurs at position 2 of $\diamond_3 \diamond_3$. Therefore $cd = ba$. Assume now for a contradiction that $u_3 = u_2 = u_1$. The word $\sqrt[2]{\mathbf{w}}$ begins with u_3, so it has cd as a prefix. Since $\ell = 4$, the second letter of $\sqrt[2]{\mathbf{w}}$ is the third letter of S and L, so $d = c$. This is a contradiction as $c \neq d$. Once again, we conclude that $u_3 \neq u_2$, that is, $u_3 \rhd u_2$. $\qquad\square$

Proof of Theorem 3.2. Let $\mathbf{w} \in \overline{\{S, L\}^\omega} \setminus \{S, L\}^\omega$ and u_1, u_2, and u_3 to respectively be the prefixes of $\mathbf{w}$, $\sqrt{\mathbf{w}}$, and $\sqrt[2]{\mathbf{w}}$ of length $|S|$. Suppose $\mathbf{w}$ begins with the letter 0. By Lemma 3.5, one of the following holds: $u_2 \rhd u_1$, $u_3 \rhd u_1$, or $\sqrt[2]{\mathbf{w}}$ periodic. Since the prefixes of length $|S|$ of the words in the orbit of $\mathbf{w}$ can increase lexicographically only finitely many times, it follows that $\mathbf{w}$ eventually

gets mapped to a periodic word. The same conclusion holds if $\mathbf{w}$ begins with the letter 1. We have proved that there exists an integer n, depending only on $|S|$, such that $\sqrt[n]{\mathbf{w}}$ is periodic.

By Theorem 2.11, the word $\sqrt[n]{\mathbf{w}}$ is periodic with minimal periodic conjugate to S. As was observed in defining the periodic part Ω_P, the word $\sqrt[n]{\mathbf{w}}$ is a coding of a rational rotation of some slope $\bar{\alpha}$. By Remark 2.4, the function $\psi\colon \rho \mapsto \frac{1}{2}(\rho + 1 - \bar{\alpha})$ relates the intercepts of $\sqrt[n]{\mathbf{w}}$ and $\sqrt[n+1]{\mathbf{w}}$ (recall that $|S| > |S_6|$). As $\psi^i(\rho)$ tends to $1 - \bar{\alpha}$ as $i \to \infty$, the word $\sqrt[n]{\mathbf{w}}$ eventually gets mapped to S^ω or L^ω as $[S]$ and $[L]$ are the two intervals with endpoint $1 - \bar{\alpha}$; for in-depth details see [5, Section 4]. As it clearly takes a bounded number of steps, depending only on S, for a point to map to $[S]$ or to $[L]$, the proof is complete. $\qquad\square$

Remark 3.6. Observe that in the above system of rotation words of rational slope $\bar{\alpha}$, ψ maps any point to $[S]$ in at most

$$\left\lceil \log_2 \left(\frac{1 - \bar{\alpha}}{\min\{|[S]|, |[L]|\}} \right) \right\rceil$$

many steps. Here $|[S]|$ and $|[L]|$ are respectively the geometric lengths of the intervals $[S]$ and $[L]$ (of slope $\bar{\alpha}$).

Next we turn our attention to injectivity. The results provided next give sufficient information to characterize the limit set. There is a slight imperfection in the following results. Namely, we are unable to characterize the preimage of the periodic part Ω_P, and we believe no nice characterization exists. First of all, the words S^ω and L^ω must have several preimages, periodic and aperiodic, by Theorem 3.2. Secondly, if $\mathbf{w}$ in Ω_A is of type (D), then not only is $\sqrt{\mathbf{w}}$ periodic with minimal period conjugate to S but the square root of any word in Ω_A that shares a prefix of length $3|S|$ with $\mathbf{w}$ is periodic with the same minimal period.[4] Therefore here we only focus on characterizing preimages of words in the aperiodic part Ω_A.

We begin with a lemma. When we say that a word $\mathbf{w}$ is preceded by a finite word u in Ω, we mean that $u\mathbf{w} \in \Omega$. Notice that each word in Ω is preceded by some word.

Lemma 3.7. *Suppose that $\mathbf{u}$ and $\mathbf{v}$ are words in Ω_γ such that $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$. If $\mathbf{u} = \gamma\gamma\cdots$ and $\mathbf{v} = \gamma\bar{\gamma}\cdots$, then $\mathbf{u} = \gamma\gamma\gamma^{2\mathfrak{c}}\bar{\gamma}\cdots$ and $\mathbf{v} = \gamma\bar{\gamma}\gamma^{2\mathfrak{c}}\bar{\gamma}\cdots$ and both $\mathbf{u}$ and $\mathbf{v}$ must be preceded by $\bar{\gamma}\gamma^{2\mathfrak{c}-1}$ in Ω.*

Proof. By Lemma 2.14 (i), the word $\mathbf{v}$ begins with $\gamma\bar{\gamma}\gamma^{2\mathfrak{c}}$. Suppose that $\mathbf{u}$ begins with $\gamma^t\bar{\gamma}$ for $t \geq 2$. Notice that $t \leq 4\mathfrak{c} + 1$ by Lemma 2.14 (i). Assume for a contradiction that t is odd. Since the prefix $\gamma^t\bar{\gamma}$ of $\mathbf{u}$ is followed by $\gamma^{2\mathfrak{c}}$ by Lemma 2.14 (i), we see that $\sqrt{\mathbf{u}} = \gamma^{(t-1)/2}\gamma\gamma^{\mathfrak{c}}\cdots$ so, as $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, we conclude that $\mathbf{v}$ begins with $\gamma\bar{\gamma}\gamma^{2\mathfrak{c}+1}$. Hence Lemma 2.14 (ii) implies that $\mathbf{v}$ has the word $\gamma\bar{\gamma}\gamma^{4\mathfrak{c}+1}\bar{\gamma}\gamma^{2\mathfrak{c}}\bar{\gamma}\gamma^{2\mathfrak{c}}$ as a prefix. Therefore $\sqrt{\mathbf{v}} = \gamma^{3\mathfrak{c}+2}\bar{\gamma}\cdots$. Since $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, we see that the prefix of $\mathbf{u}$ of length $2(3\mathfrak{c} + 2)|\gamma|$ must be followed by $\bar{\gamma}$. Now by Lemma 2.14 (i), the distance between two occurrences of $\bar{\gamma}$ in $\mathbf{u}$ is always a multiple of $(2\mathfrak{c} + 1)|\gamma|$. Hence $t + r(2\mathfrak{c} + 1) = 2(3\mathfrak{c} + 2)$ for some positive integer r. Since $t \geq 2$, we see that $r \leq 2$. If $r = 2$, then $t = 2\mathfrak{c} + 2$, which is impossible as t is odd. Thus the only option is that $r = 1$, that is, $t = 4\mathfrak{c} + 3$. This contradicts the fact that $t \leq 4\mathfrak{c} + 1$. We have thus proved that t must be even.

Now $\sqrt{\mathbf{u}} = \gamma^{t/2}\bar{\gamma}\gamma^{2\mathfrak{c}}\cdots$ and $\sqrt{\mathbf{v}} = \gamma\gamma^{\mathfrak{c}}\cdots$. Since $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, it must be that the prefix of $\mathbf{v}$ of length $t|\gamma|$ must be followed by $\bar{\gamma}$. Like previously, we see that $t = 1 + r(2\mathfrak{c} + 1)$ for some positive integer r. Since $t \leq 4\mathfrak{c} + 1$, we see that the only option is that $r = 1$, that is, $\mathbf{v} = \gamma\bar{\gamma}\gamma^{2\mathfrak{c}}\bar{\gamma}\cdots$ and $\mathbf{u} = \gamma\gamma\gamma^{2\mathfrak{c}}\bar{\gamma}\cdots$. What is left is to show that both $\mathbf{u}$ and $\mathbf{v}$ must be preceded by $\bar{\gamma}\gamma^{2\mathfrak{c}-1}$ in Ω.

Since the word $\mathbf{u}$ begins with $\gamma^{2\mathfrak{c}+2}\bar{\gamma}$, it is clear by Lemma 2.14 (i) that it must be preceded by $\bar{\gamma}\gamma^{2\mathfrak{c}-1}$. Assume for a contradiction that $\mathbf{v}$ is preceded by $\gamma^{2\mathfrak{c}}$. By Lemma 2.14 (ii), either $\mathbf{v}$ has $\gamma\bar{\gamma}(\gamma^{2\mathfrak{c}}\bar{\gamma})^{4\mathfrak{c}}\gamma^{4\mathfrak{c}+1}$ as a prefix or it has $\gamma\bar{\gamma}(\gamma^{2\mathfrak{c}}\bar{\gamma})^{2\mathfrak{c}-1}\gamma^{4\mathfrak{c}+1}$ as a prefix. Consider the former case,

---

[4]See the proof of [5, Theorem 44] for precise details.

where $\sqrt{\mathbf{v}}$ begins with $\gamma^{\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{2\mathfrak{c}}\gamma^{\mathfrak{c}+1}$. Clearly the prefix $\gamma^{2\mathfrak{c}+2}\overline{\gamma}$ of $\mathbf{u}$ must be followed by $(\gamma^{2\mathfrak{c}}\overline{\gamma})^{2\mathfrak{c}-1}\gamma^{2\mathfrak{c}}$ by Lemma 2.14 (ii). The square root of this prefix equals $\gamma^{\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{\mathfrak{c}}$, so as $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, we conclude by Lemma 2.14 (ii) that $\mathbf{u}$ has the word $\gamma^{2\mathfrak{c}+2}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{4\mathfrak{c}}\overline{\gamma}\gamma$ as a prefix. However, now $\sqrt{\mathbf{u}} = \gamma^{\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{2\mathfrak{c}}\overline{\gamma}\cdots \neq \sqrt{\mathbf{v}}$, which is impossible. Therefore we are left with the case where $\mathbf{v}$ begins with $\gamma\overline{\gamma}(\gamma^{2\mathfrak{c}}\overline{\gamma})^{2\mathfrak{c}-1}\gamma^{4\mathfrak{c}+1}$. Set $x = \gamma^{2\mathfrak{c}+2}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{2\mathfrak{c}}$, $y = \gamma\overline{\gamma}(\gamma^{2\mathfrak{c}}\overline{\gamma})^{2\mathfrak{c}-1}\gamma^{4\mathfrak{c}+1}$, $z = (\overline{\gamma}\gamma^{2\mathfrak{c}})^{2\mathfrak{c}}$, and $\overline{z} = \overline{\gamma}\gamma^{4\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{2\mathfrak{c}-2}$. Observe that the prefixes x and y of $\mathbf{u}$ and $\mathbf{v}$ have the same length and that this length divided by $|\gamma|$ is even. Further, notice that $\sqrt{x} = \sqrt{y} = \gamma^{\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{\mathfrak{c}}$ and $\sqrt{z} = \sqrt{\overline{z}} = (\overline{\gamma}\gamma^{2\mathfrak{c}})^{\mathfrak{c}}$. Once again by applying Lemma 2.14 (ii), we see that $\mathbf{v}$ must have yz as a prefix. Since $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, it follows that xz is a prefix of $\mathbf{u}$. Now $xz = \gamma^{2\mathfrak{c}+2}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{4\mathfrak{c}}$, so it must be followed by $\overline{z}$ by Lemma 2.14 (ii). As $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$, the prefix yz of $\mathbf{v}$ must be followed by $\overline{\gamma}$ meaning that $\mathbf{v}$ has yzz as a prefix. The word yzz has $\gamma^{4\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{4\mathfrak{c}}$ as a suffix, so it must be that $yzz\overline{z}$ is a prefix of $\mathbf{v}$. It follows that the square root of the factor of $\mathbf{u}$ of length $|z|$ following the prefix $xz\overline{z}$ equals $\sqrt{\overline{z}}$. The prefix $xz\overline{z}$ of $\mathbf{u}$ must be followed by $(\overline{\gamma}\gamma^{2\mathfrak{c}})^2$ by Lemma 2.14 (ii). As $\sqrt{\overline{z}} = (\overline{\gamma}\gamma^{2\mathfrak{c}})^{\mathfrak{c}}$ and the prefix $yzz\overline{z}$ of $\mathbf{v}$ must be followed by $\overline{\gamma}$, the prefix $xz\overline{z}$ must be followed by $(\overline{\gamma}\gamma^{2\mathfrak{c}})^2\overline{\gamma}$ which means that $xz\overline{z}z$ is a prefix of $\mathbf{u}$. Now the known prefix $yzz\overline{z}$ of $\mathbf{v}$ ends with $\overline{z}$, and an analogous argument shows that $yzz\overline{z}z$ is a prefix of $\mathbf{v}$. By Lemma 2.14 (ii), the prefix $yzz\overline{z}z$ must be followed by $\overline{\gamma}$, and consequently the word $\sqrt{\mathbf{v}}$ must begin with $\gamma^{\mathfrak{c}+1}(\overline{\gamma}\gamma^{2\mathfrak{c}})^{5\mathfrak{c}}\overline{\gamma}$. This contradicts Lemma 2.14 (ii) because $\sqrt{\mathbf{v}} \in \Omega_A$ by Theorem 2.10. Thus $\mathbf{v}$ cannot be preceded by $\gamma^{2\mathfrak{c}}$. By Lemma 2.14 (ii), it must therefore be preceded by $\overline{\gamma}\gamma^{2\mathfrak{c}-1}$ as desired. $\qquad\square$

The next theorem says that the square root map is not injective on Ω_A, i.e. the aperiodic part of Ω, but that it is almost injective: only words of restricted form may have more than one preimage and even then there are at most two preimages. In the Sturmian case, all words have at most one preimage.

Theorem 3.8. *If $\mathbf{w}$ is a word in Ω_A having two preimages $\mathbf{u}$ and $\mathbf{v}$ in Ω under the square root map, then $\mathbf{u} = zS\mathbf{\Gamma}_1$ and $\mathbf{v} = zS\mathbf{\Gamma}_2$ where zS is a suffix of some γ_k such that $z \in \Pi(\mathfrak{a}, \mathfrak{b})$.*

Proof. First of all, notice that $\sqrt{S\Gamma_1} = \sqrt{S\Gamma_2}$ because $\sqrt{SL} = \sqrt{SS}$. Thus we only need to show that words having two preimages must be of the claimed form. Assume that $\mathbf{u}$ and $\mathbf{v}$ are distinct words in Ω having the same square root in Ω_A. Notice that both $\mathbf{u}$ and $\mathbf{v}$ are aperiodic since $\sqrt{\Omega_P} \subseteq \Omega_P$. Suppose first that $\mathbf{u}$ and $\mathbf{v}$ are products of the words S and L. Let next $\gamma = \gamma_k$ for some $k \geq 0$, and assume that the words $\mathbf{u}$ and $\mathbf{v}$ are products of γ and $\overline{\gamma}$ and that they have a minimal common prefix. In other words, we have $\mathbf{u} = \gamma\gamma\cdots$ and $\mathbf{v} = \gamma\overline{\gamma}\cdots$. Lemma 3.7 implies that $\mathbf{u} = \gamma\gamma\gamma^{2\mathfrak{c}}\overline{\gamma}\cdots$ and $\mathbf{v} = \gamma\overline{\gamma}\gamma^{2\mathfrak{c}}\overline{\gamma}\cdots$. In other words, $\mathbf{u} = \gamma\overline{\gamma}_{k+1}\cdots$ and $\mathbf{v} = \gamma\gamma_{k+1}\cdots$. Moreover, both $\mathbf{u}$ and $\mathbf{v}$ must be preceded by $\gamma_{k+1}\gamma^{-1}$. Thus the words $\gamma_{k+1}\gamma^{-1}\mathbf{u}$ and $\gamma_{k+1}\gamma^{-1}\mathbf{v}$ have prefixes $\gamma_{k+1}\overline{\gamma}_{k+1}$ and $\gamma_{k+1}\gamma_{k+1}$ respectively. Since $|\gamma_{k+1}|/|\gamma|$ is odd, we see that $\sqrt{\gamma_{k+1}\gamma^{-1}\mathbf{u}} = \sqrt{\gamma_{k+1}\gamma^{-1}\mathbf{v}}$. Therefore we can repeat our argument so far with γ_{k+1} in place of γ. The conclusion is that $\mathbf{u} = \gamma\mathbf{\Gamma}_1$ and $\mathbf{v} = \gamma\mathbf{\Gamma}_2$ (or $\mathbf{u} = \gamma\mathbf{\Gamma}_2$ and $\mathbf{v} = \gamma\mathbf{\Gamma}_1$). Moreover, we have shown that $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ uniquely extend to the left by γ_i for all $i \geq 1$. Therefore if we allow $\mathbf{u}$ and $\mathbf{v}$ to have arbitrarily long common prefix, it must be that $\mathbf{u} = zS\mathbf{\Gamma}_1$ and $\mathbf{v} = zS\mathbf{\Gamma}_2$, where zS is a suffix of some γ_k such that $z \in \Pi(\mathfrak{a}, \mathfrak{b})$.

Suppose then that one of the words $\mathbf{u}$ and $\mathbf{v}$ is not in Ω_S. Since $\mathbf{u}$ and $\mathbf{v}$ are aperiodic, their types are well-defined. Neither $\mathbf{u}$ nor $\mathbf{v}$ is of type (D) as we assume that $\sqrt{\mathbf{u}}$ is aperiodic. Thus $\mathbf{u} = x\mathbf{w}'$ and $\mathbf{v} = y\mathbf{w}''$ for words $\mathbf{w}'$ and $\mathbf{w}''$ in Ω_S and words x and y such that $x, y \in \Pi(\mathfrak{a}, \mathfrak{b}) \cup \{\varepsilon\}$ and $|x|, |y| < 2|S|$. Now $\sqrt{\mathbf{u}} = \sqrt{\mathbf{v}}$ so, since $|\sqrt{x}|, |\sqrt{y}| < |S|$ and infinite products of S and L synchronize, we conclude that $\sqrt{x} = \sqrt{y}$ and $\sqrt{\mathbf{w}'} = \sqrt{\mathbf{w}''}$. Thus by the arguments of the preceding paragraph, we have, say, $\mathbf{u} = xzS\mathbf{\Gamma}_1$ and $\mathbf{v} = yzS\mathbf{\Gamma}_2$ for some $z \in \{SS, SL, LS\}^*$. Since $\mathbf{\Gamma}$ uniquely extends to the left by γ_i for all $i \geq 1$, we see that $\mathbf{u}$ and $\mathbf{v}$ are of the claimed form. $\qquad\square$

Let us state separately an observation made in the proof of Theorem 3.8 that is helpful when we next characterize the points that are in the limit set.

Corollary 3.9. *In Ω, the word Γ is uniquely extended to the left by γ_k for all $k \geq 0$.*

The *limit set* Λ is the set of words that have arbitrarily long chains of preimages, that is,

$$\Lambda = \bigcap_{n=0}^{\infty} \sqrt[n]{\Omega}.$$

In the Sturmian case, the limit set contains only the two fixed points of the square root map. For the subshift Ω, the limit set is much larger. In fact, the limit set contains all words that are products of the words S and L. Proving this result is our next aim.

Theorem 3.10. *We have $\Lambda = \Omega_S$.*

We begin with a lemma after which we proceed to prove Theorem 3.10.

Lemma 3.11. *If $\mathbf{w} \in \Omega_S$ and $\mathbf{w}$ has Γ as a suffix, then $\mathbf{w} \in \Lambda$.*

Proof. Since both $\mathbf{w}$ and Γ are in Ω_S and the factorization of a word as a product of the words S and L is unique, we see that $\mathbf{w} = z\Gamma$ for some $z \in \{S, L\}^*$. Let z' be a suffix of some γ_k of length $2|z|$. By Corollary 3.9, we have $z'\Gamma \in \Omega$. Now $\sqrt{z'\Gamma} = \sqrt{z'}\Gamma$ and $|\sqrt{z'}| = |z|$, so $\sqrt{z'} = z$ by Corollary 3.9. Thus $\mathbf{w}$ has a preimage of the same form. $\square$

Proof of Theorem 3.10. Suppose first that $\mathbf{w} \in \Omega \setminus \Omega_S$. Since the square root of a word in Ω_S is also in Ω_S, we see that all preimages of $\mathbf{w}$ are in $\Omega \setminus \Omega_S$. It is thus an immediate consequence of Theorem 3.2 that every backward orbit of $\mathbf{w}$ is finite, that is, $\mathbf{w} \notin \Lambda$.

Suppose then that $\mathbf{w} \in \Omega_S$. The only periodic words in Ω_S are the fixed points S^ω and L^ω which clearly have a preimage. It is thus enough to assume that $\mathbf{w}$ is aperiodic and to find two sequences (u_n) and (v_n) with the following properties:

- u_n is a prefix of $\mathbf{w}$ for all $n \geq 1$ and the sequence $(|u_n|)$ is strictly increasing,

- $v_n \in \mathcal{L}(\Omega)$ and $\sqrt{v_n} = u_n$ for all $n \geq 1$.

By compactness, a subsequence of (v_n) converges to an infinite word $\mathbf{v}$ in Ω with the property that $\sqrt{\mathbf{v}} = \mathbf{w}$. Therefore $\mathbf{w}$ has a preimage in Ω so, as $\mathbf{w}$ was arbitrary, we conclude that $\mathbf{w} \in \Lambda$.

We may assume that $\mathbf{w}$ does not have Γ as a suffix by Lemma 3.11. Since $\mathbf{w} \neq \Gamma$, there exists maximal k_1 such that the γ_{k_1}-factorization of $\mathbf{w}$ starts at the beginning of $\mathbf{w}$. Let j_1 be the starting position of the γ_{k_1+1}-factorization of $\mathbf{w}$. By the maximality of k_1, we have $j_1 \in \{1, 2, \ldots, 2\mathfrak{c}\}|\gamma_{k_1}|$. Let u_1 be the prefix of $\mathbf{w}$ of length j_1 and v_1 be the suffix of $\gamma_{k_1+1}^2$ of length $2|u_1|$. We have $v_1 \in \Pi(\mathfrak{a}, \mathfrak{b})$ by Lemma 2.8, so $\sqrt{v_1} = u_1$ as u_1 is a suffix of γ_{k_1+1}. Again since Γ is not a suffix of $\mathbf{w}$, we see that there exists maximal k_2 such that the γ_{k_2}-factorization of $\mathbf{w}$ starts at position j_1. The γ_{k_2+1}-factorization of $\mathbf{w}$ begins at position j_2 where $j_2 = j_1 + t|\gamma_{k_2}|$ with $1 \leq t \leq 2\mathfrak{c}$. Set again u_2 to be the prefix of $\mathbf{w}$ of length j_2 and v_2 to be the suffix of $\gamma_{k_2+1}^2$ of length $2|u_2|$. By the definition of j_2, the word u_2 is a suffix of γ_{k_2+1}. Observe that $|u_2|$ is a multiple of $|\gamma_{k_1}|$, so $|v_2|$ is an even multiple of $|\gamma_{k_1}|$. Since $\gamma_{k_2+1}^2$ is a product of the words γ_{k_1} and $\overline{\gamma}_{k_1}$, it follows that $\sqrt{v_2} = u_2$. By repeating these arguments, we obtain the desired sequences (u_n) and (v_n). $\square$

Finally, we consider invariant subsets and show that the limit set Λ is not simple in the sense that it contains infinitely many invariant subsets. We first show how to decompose Ω_γ into two invariant sets.

Let A_k be the set of words in Ω_{γ_k} that have one of the following words as a prefix: $\gamma_k \gamma_k^{2\mathfrak{c}} \overline{\gamma}_k$, $\overline{\gamma}_k \gamma_k^{2\mathfrak{c}} \overline{\gamma}_k$, or $\overline{\gamma}_k \gamma_k^{2\mathfrak{c}} \gamma_k$. Let us find the prefixes of length $(2\mathfrak{c} + 2)|\gamma_k|$ of the preimages of the words

in A_k. Let $\mathbf{w}$ be a word in Ω_γ (now $\gamma = \gamma_k$) such that $\sqrt{\mathbf{w}} \in A_k$. Say $\sqrt{\mathbf{w}}$ has $\overline{\gamma}\gamma^{2\mathfrak{c}}\overline{\gamma}$ as a prefix so that the prefix of $\mathbf{w}$ of length $2(2\mathfrak{c}+2)|\gamma|$ is of the form $\overline{\gamma}\diamond(\gamma\diamond)^{2\mathfrak{c}}\overline{\gamma}\diamond$. By Lemma 2.14 (i), this prefix must equal $\overline{\gamma}\gamma^{2\mathfrak{c}}\diamond\gamma^{2\mathfrak{c}}\overline{\gamma}\gamma$ with $\diamond \in \{\gamma, \overline{\gamma}\}$, and it follows that $\mathbf{w} \in A_k$. Suppose next that $\mathbf{w}$ has a prefix of the form $\overline{\gamma}\diamond(\gamma\diamond)^{2\mathfrak{c}}\gamma\diamond$. Like previously, this prefix must take the form $\overline{\gamma}\gamma^{2\mathfrak{c}}(\diamond\gamma)^{\mathfrak{c}+1}\diamond$. If $\mathbf{w}$ has $\overline{\gamma}\gamma^{2\mathfrak{c}}\gamma$ as a prefix, then it has $\overline{\gamma}\gamma^{4\mathfrak{c}+1}\overline{\gamma}$ as a prefix by Lemma 2.14 (i). This is clearly a contradiction, so $\overline{\gamma}\gamma^{2\mathfrak{c}}\overline{\gamma}$ is a prefix of $\mathbf{w}$, which in turn implies that $\mathbf{w} \in A_k$. Consider the last case where $\mathbf{w}$ has a prefix of the form $\gamma\diamond(\gamma\diamond)^{2\mathfrak{c}}\overline{\gamma}\diamond$. Again, it must be that the prefix takes the form $(\gamma\diamond)^{\mathfrak{c}+1}\gamma^{2\mathfrak{c}}\overline{\gamma}\gamma$. If the prefix of length $(2\mathfrak{c}+1)|\gamma|$ is followed by γ, then $\mathbf{w}$ must begin with $\overline{\gamma}$ by Lemma 2.14 (i). As this is impossible, we see that again $\mathbf{w} \in A_k$. We have thus proved that $\Omega_\gamma \setminus A_k$ is invariant under the square root map. It is straightforward to see that also A_k is invariant.

Let us show next that $A_k = \Omega_{\gamma_{k+1}}$ for all $k \geq 0$. It is clear that $A_k \subseteq \Omega_{\gamma_{k+1}}$. Let $\mathbf{w} \in \Omega_{\gamma_{k+1}}$, and consider its prefix of length $(2\mathfrak{c}+2)|\gamma_k|$. If $\overline{\gamma}_k$ begins at position $t|\gamma_k|$ of $\mathbf{w}$ with $0 < t \leq 2\mathfrak{c}$, then clearly $\mathbf{w} \notin \Omega_{\gamma_{k+1}}$. If $\overline{\gamma}_k$ is a prefix of $\mathbf{w}$, then obviously $\mathbf{w} \in A_k$. If $\overline{\gamma}_k$ occurs at position $(2\mathfrak{c}+1)|\gamma_k|$ of $\mathbf{w}$, then $\mathbf{w}$ has either $\gamma_k\gamma_k^{2\mathfrak{c}}\overline{\gamma}_k$ or $\overline{\gamma}_k\gamma_k^{2\mathfrak{c}}\overline{\gamma}_k$ as a prefix, and $\mathbf{w} \in A_k$. Thus we are left with the case that $\gamma_k^{2\mathfrak{c}+2}$ is a prefix of $\mathbf{w}$. Now $\mathbf{w}$ has prefix $\gamma_k^{2\mathfrak{c}+2+r}\overline{\gamma}_k$ for some $r \geq 0$. As $\mathbf{w} \in \Omega_{\gamma_{k+1}}$, we see that $2\mathfrak{c}+2+r$ is a multiple of $2\mathfrak{c}+1$. Further by Lemma 2.14 (i), it must be that $2\mathfrak{c}+2+r = 2(2\mathfrak{c}+1)$, which implies that $r = 2\mathfrak{c}$. Thus $\mathbf{w}$ has $\gamma^{4\mathfrak{c}+2}$ as a prefix, which contradicts Lemma 2.14 (i).

Putting together the results of the preceding two paragraphs yields the following result.

Proposition 3.12. *We have the disjoint union*

$$\Lambda = \{\mathbf{\Gamma}_1, \mathbf{\Gamma}_2\} \cup \bigcup_{k=0}^{\infty} \Omega_{\gamma_k} \setminus \Omega_{\gamma_{k+1}}$$

of subsets invariant under the square root map.

Proof. Notice that $\Lambda = \Omega_{\gamma_0}$ by Theorem 3.10. Using the above arguments, we can write $\Omega_{\gamma_k} = A_k \cup (\Omega_{\gamma_k} \setminus A_k) = \Omega_{\gamma_{k+1}} \cup (\Omega_{\gamma_k} \setminus \Omega_{\gamma_{k+1}})$ for all $k \geq 0$. The sets in the union are disjoint and invariant. Clearly the words in $\Lambda \setminus \bigcup_{k=0}^{\infty} \Omega_{\gamma_k} \setminus \Omega_{\gamma_{k+1}}$ are exactly the words whose γ_k-factorization begins at the beginning for all $k \geq 0$. These words are by construction the two fixed points $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$. They clearly form an invariant subset. $\square$

4 Periodic Points

In this section, we characterize the periodic points of the square root map in Ω. The result is that the only periodic points are fixed points. We further characterize asymptotically periodic points and show that all asymptotically periodic points are ultimately periodic points.

Recall that a word $\mathbf{w}$ is a *periodic point* of the square root map with period n if $\sqrt[n]{\mathbf{w}} = \mathbf{w}$.

Theorem 4.1. *If $\mathbf{w}$ is a periodic point in Ω, then $\mathbf{w} \in \{\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, S^\omega, L^\omega\}$.*

Proof. By Theorem 3.2, no word in $\Omega \setminus \Omega_S$ can be a periodic point. Thus we assume that $\mathbf{w}$ is a word in Ω_S such that $\sqrt[n]{\mathbf{w}} = \mathbf{w}$ for some integer $n \geq 1$. Suppose for a contradiction that $\mathbf{w} \notin \{\mathbf{\Gamma}_1, \mathbf{\Gamma}_2, S^\omega, L^\omega\}$. It follows that there exists maximal k such that the γ_k-factorization of $\mathbf{w}$ starts at the beginning of $\mathbf{w}$. Since the square root map acts essentially the same way on products of γ and $\overline{\gamma}$ and on products of S and L due to Lemma 2.8, we may assume that $k = 0$. For $i \geq 1$, let k_i be such that the starting position of the γ_i-factorization of $\mathbf{w}$ equals $k_i|S|$. In particular, we have $k_i \neq 0$ for all $i \geq 1$.

Write $\mathbf{w} = a_0 a_1 \ldots$ for $a_t \in \{S, L\}$, and let $i \geq 1$. Define the infinite word $\mathbf{u}_i$ as the subword $a_{k_i} a_{2k_i} a_{4k_i} \cdots a_{2^t k_i} \cdots$. Since $\sqrt[n]{\mathbf{w}} = \mathbf{w}$, we see that the relation $a_t = a_{2^n t}$ holds for all $t \geq 0$. Hence

the word $\mathbf{u}_i$ has the property that $\mathbf{u}_i = T^{n|S|}(\mathbf{u}_i)$ for all $i \geq 1$, i.e., it is purely periodic. We shall show that this is impossible, and thus that $\mathbf{w}$ does not exist.

Because the γ_i-factorization of $\mathbf{w}$ begins at position $k_i|S|$, we actually know most of the contents of the word $\mathbf{u}_i$ without knowing anything particular about the γ_i-factorization of $\mathbf{w}$. The word $\mathbf{u}_i$ is obtained by concatenating the factors of length $|S|$ of γ_i or $\overline{\gamma}_i$ occurring at positions given by the sequence $(d_t|S|)$ where d_t is given by the sequence $((2^t - 1)k_i)_{t\geq 0}$ modulo $(2\mathfrak{c} + 1)^i$. Thus there is ambiguity only when $(2^t - 1)k_i \equiv 0 \pmod{(2\mathfrak{c} + 1)^i}$. Let p_i be the minimal period of the sequence (d_t). We have

$$\mathbf{u}_i = \prod_{t=1}^{\infty} \diamond_t v_i \tag{4}$$

where $\diamond_t \in \{S, L\}$ and v_i is a word of length $p_i - 1$ over $\{S, L\}$. Notice that $p_i > 1$ because $k_i \neq 0$. Let us next see what the word v_i is like.

Suppose first that $i = 1$. Since $\gamma_1 = LS^{2\mathfrak{c}}$, we see that $v_i = S^{p_1 - 1}$. Let then $i > 1$, so we have $\gamma_i = \overline{\gamma}_{i-1}\gamma_{i-1}^{2\mathfrak{c}}$. Observe that $k_i \in k_{i-1} + \{0, \ldots, 2\mathfrak{c}\}(2\mathfrak{c} + 1)^{i-1}$ so, by basic modular arithmetic, it is straightforward to see that p_{i-1} divides p_i. We see that the word v_i has the word v_{i-1} as a prefix since the prefix of v_i of length $(p_{i-1} - 1)|S|$ is determined by the positions $(2 - 1)k_i|S|$, $(4 - 1)k_i|S|$, $\ldots$, $(2^{p_{i-1}} - 1)k_i|S|$ of γ_i, with the coefficients of $|S|$ taken modulo $(2\mathfrak{c} + 1)^i$, that is, by the same positions of γ_{i-1}. Suppose then that $p_i > p_{i-1}$. By the form of γ_i, the next factor of v_i of length $p_{i-1}|S|$ is determined by the positions $(1 - 1)k_i|S|$, $(2 - 1)k_i|S|$, $\ldots$, $(2^{p_{i-1}} - 1)k_i|S|$ of γ_{i-1}. Repeating this reasoning, it follows that $v_i = v_{i-1}(bv_{i-1})^{p_i/p_{i-1} - 1}$ where $b = S$ if γ_{i-1} begins with S and L otherwise. The words v_i can be generated as follows. Let $v_0' = S$ and $p_0 = 1$. If v_{i-1}' is defined, then $v_i' = L(v_{i-1}')v_{i-1}'^{p_i/p_{i-1} - 1}$. Now v_i is obtained from v_i' by deleting its first $|S|$ letters. The construction of the words v_i' is otherwise the same as the construction of the words γ_i but the exponents are not fixed to $2\mathfrak{c}$. A proof analogous to that showing that the words γ_i are primitive shows that the words v_i' are primitive.

Let us show next that the words $\diamond_t$ of (4) take both values S and L infinitely often. For this, we need to prove the following claim.

Claim. *The sequence (p_i) is unbounded.*

Proof. If the minimal period of the sequence $((2^t - 1)k_i)_{t\geq 0}$ modulo $(p^\ell)^i$ eventually increases as i increases for a prime power p^ℓ dividing $2\mathfrak{c} - 1$, then the minimal period p_i of the sequence (d_t) must also eventually increase by the Chinese Remainder Theorem. Thus it is sufficient to show that the sequence (p_i) is unbounded in the case that $2\mathfrak{c} + 1 = p^\ell$ for a prime p. Suppose that $\gcd(k_1, 2\mathfrak{c} + 1) = p^a$. Since $k_1 < 2\mathfrak{c} + 1$, we have $a < \ell$. Suppose for a contradiction that $\gcd(k_i, (2\mathfrak{c} + 1)^i) > p^a$ for some $i > 1$. Then p^{a+1} divides k_i. Since $k_i = k_{i-1} + rp^{\ell(i-1)}$ for some $r \in \{0, \ldots, 2\mathfrak{c}\}$, it follows that p^{a+1} divides k_{i-1}. By iterating the argument, we see that p^{a+1} divides k_1; a contradiction. Therefore $\gcd(k_i, (2\mathfrak{c} + 1)^i) = p^a$ for all $i \geq 1$. Let $i \geq 1$ and pick j so large that $(2\mathfrak{c} + 1)^j/p^a > 2^{p_i} - 1$. Then it must be that $(2^{p_i} - 1)k_j \not\equiv 0 \pmod{(2\mathfrak{c} + 1)^j}$, so $p_j > p_i$. $\qquad\square$

The claim implies that there exists j such that $p_{j+1} > p_j > p_i$. Since $p_j > p_i$, there exists infinitely many t such that $(2^{tp_i} - 1)k_i \equiv 0 \pmod{(2\mathfrak{c} + 1)^i}$ and $(2^{tp_i} - 1)k_i \not\equiv 0 \pmod{(2\mathfrak{c} + 1)^j}$. Thus for these t, the word $\diamond_t$ equals the first $|S|$ letters of the word γ_{j-1}. Similarly there exists infinitely many t such that $(2^{tp_i} - 1)k_i \equiv 0 \pmod{(2\mathfrak{c} + 1)^i}$ and $(2^{tp_i} - 1)k_i \not\equiv 0 \pmod{(2\mathfrak{c} + 1)^{j+1}}$. For these numbers t, the word $\diamond_t$ equals the first $|S|$ letters of the word γ_j. Since γ_{j-1} begins with S and γ_j begins with L or vice versa, the words $\diamond_t$ indeed take both values S and L infinitely often.

Let us show next that if $\mathbf{u}_i$ is purely periodic with period m, then $p_i|S|$ divides m. Assume on the contrary that $p_i|S|$ does not divide m. This means by (4) that for all large enough r the word $\diamond_r v_i$ is an interior factor of $\diamond_s v_i \diamond_{s+1} v_i$ for some s. Let j be the largest integer such that $|v_j'| < |v_i'|$. Due to the primitivity of the word v_j', we see that the position where $\diamond_r v_i$ occurs at is a multiple of $|v_j'|$. We conclude that $\diamond_r$ is uniquely determined by the first $|S|$ letters of v_j'. This is a contradiction because $\diamond_t$ takes both values S and L infinitely often. Therefore $p_i|S|$ divides m.

Recall that, for all $i \geq 1$, $\mathbf{u}_i$ is purely periodic with period n. By the arguments of the previous paragraph, the number p_i divides n for all $i \geq 1$. This is absurd as the sequence (p_i) is unbounded. This contradiction concludes the proof. □

The case with the Sturmian periodic points is similar: periodic points are fixed points and the fixed points are obtained as limits from solutions of (3).

Next we consider the dynamical notion of an asymptotically periodic point and characterize asymptotically periodic points in Ω.

Definition 4.2. Let (X, f) be a dynamical system. A point x in X is *asymptotically periodic* if there exists a periodic point y in X such that

$$\lim_{n \to \infty} d(f^n(x), f^n(y)) = 0.$$

If this is the case, then we say that the point x is *asymptotically periodic to y*.

The following proposition essentially says that if a word in Ω is asymptotically periodic, then it is an ultimately periodic point. The situation is opposite to the Sturmian case where all words are asymptotically periodic and only periodic points are ultimately periodic points.

Proposition 4.3. *If $\mathbf{w} \in \Omega_S$, then $\mathbf{w}$ is asymptotically periodic if and only if $\mathbf{w} \in \{\Gamma_1, \Gamma_2, S^\omega, L^\omega\}$, that is, if and only if $\mathbf{w}$ is a periodic point. If $\mathbf{w} \in \Omega \setminus \Omega_S$, then $\mathbf{w}$ is asymptotically periodic to S^ω or L^ω.*

Proof. Let $\mathbf{w} \in \Omega \setminus \Omega_S$. By Theorem 3.2, there exists an integer n such that $\sqrt[n]{\mathbf{w}} \in \{S^\omega, L^\omega\}$, so $\mathbf{w}$ is asymptotically periodic to S^ω or L^ω. Suppose then that $\mathbf{w}$ in Ω_S is aperiodic and asymptotically periodic. By Theorem 4.1, this means that the sequence $(\sqrt[n]{\mathbf{w}})_n$ converges to Γ. Observe that if $k > 0$ and $\mathbf{w} \notin \Omega_{\gamma_k}$, then also $\sqrt{\mathbf{w}} \notin \Omega_{\gamma_k}$. From the fact that $\Gamma \in \Omega_{\gamma_k}$ for all $k \geq 0$, we thus conclude that $\mathbf{w} \in \Omega_{\gamma_k}$ for all $k \geq 0$ which means that $\mathbf{w} = \Gamma$. □

5 Solutions to the Word Equation in $\mathcal{L}(\Omega)$

This section contains a characterization of long enough solutions to the word equation (3) in $\mathcal{L}(\Omega)$. The construction of the fixed points Γ_1 and Γ_2 introduces the solutions $S, L, \gamma_1, \gamma_2, \ldots$ into $\mathcal{L}(\Omega)$. The main result of this section, Theorem 5.4, tells that these are essentially all solutions to (3) in $\mathcal{L}(\Omega)$, the construction does not introduce any additional, or accidental, solutions.

Let us first characterize squares in Ω°.

Lemma 5.1. *Let u be primitive. Then $u^2 \in \mathcal{L}(\Omega^\circ)$ if and only if u is conjugate to $\tau^k(S)$ for some $k \geq 0$.*

Proof. Observe that $(\tau^k(S))^3 \in \mathcal{L}(\Omega^\circ)$ for all $k \geq 0$ because $(\tau^k(S))^{4\mathfrak{c}+1}$ occurs between two occurrences of $\tau^k(L)$; see Lemma 2.14 (i). Therefore if u is conjugate to $\tau^k(S)$, then $u^2 \in \mathcal{L}(\Omega^\circ)$.

Suppose that $u^2 \in \mathcal{L}(\Omega^\circ)$ with u primitive. If $u = \tau(v)$ for some word v, then $v^2 \in \mathcal{L}(\Omega^\circ)$ and, by induction, v is conjugate to $\tau^k(S)$ for some $k \geq 0$. This means that u must be conjugate to $\tau^{k+1}(S)$. Assume that u is not of the form $\tau(v)$. If $u = S$, then the claim holds. Otherwise u must contain at least one occurrence of the letter L, and it is possible to factorize $u = x\tau(u')ay$ for some words u', x, and y and letter a such that ayx equals $\tau(S)$ or $\tau(L)$. Let b be a letter such that

$b \neq a$. By the simple form of the substitution τ, we see that $(u'b)^2 \in \mathcal{L}(\Omega^\circ)$. By induction, $u'b$ is conjugate to $\tau^k(S)$ for some $k \geq 0$. Consequently, bu' is conjugate to $\tau^k(S)$, which in turn implies that $\tau(bu')$, which equals $ayx\tau(u')$, is conjugate to $\tau^{k+1}(S)$. Now u and $ayx\tau(u')$ are conjugate, proving the claim. $\qquad\square$

The next theorem, which is quite general, could be of independent interest in characterizing solutions to (3) more generally.

Theorem 5.2. *If u is a word that is a product of the words S and L and a primitive solution to* (3), *then none of its proper conjugates are solutions to* (3), *except in the case that $u \in \{S, L\}$ when S and L are the only conjugates of u that are solutions to* (3).

Proof. Suppose that u is a word that is a product of the words S and L and a primitive solution to (3). If $u \in \{S, L\}$, then the only conjugates of u that are solutions to (3) are S and L by [5, Theorem 18]. We may thus suppose that $u \notin \{S, L\}$. Let v be a proper conjugate of u, and assume for a contradiction that v is a solution to (3).

Suppose first that v is a product of the words S and L. Consider u and v as words over the alphabet $\{S, L\}$. The length of u (as a word over $\{S, L\}$) must be odd, as otherwise we would have $u = (\sqrt{u})^2$ contradicting the primitivity of u. Since both u and v are solutions to (3), we have $u[i] = u[2i]$ and $v[i] = v[2i]$ for $i \in \{0, 1, \ldots, |u| - 1\}$ (here the indices are naturally interpreted modulo $|u|$). Since v is a proper conjugate of u, we have $v[i] = u[i + \ell]$ for some $\ell \in \{1, 2, \ldots, |u| - 1\}$. Therefore $u[2i] = u[i] = v[i - \ell] = v[2i - 2\ell] = u[2i - \ell]$ so, as the length of u is odd, we conclude that $u[j] = u[j - \ell]$ for all $j \in \{0, 1, \ldots, |u| - 1\}$. This implies that u is a power of a word of length $\gcd(\ell, |u|)$, which contradicts the primitivity of u.

Assume then that v is not a product of the words S and L. Consider an arbitrary infinite word $\mathbf{w}$ that is a product of the words S and L and has u^5 as a prefix. Suppose that v occurs at position ℓ of u^2, so that the word $T^\ell(\mathbf{w})$ has v^4 as a prefix. Since v is not a product of the words S and L, the word $T^\ell(\mathbf{w})$ is not in $\{S, L\}^\omega$. Moreover, as v is a solution to (3), the words $\sqrt{T^\ell(\mathbf{w})}$ and $\sqrt[2]{T^\ell(\mathbf{w})}$ have v as a prefix. Now Lemma 3.5 implies that $\sqrt[2]{T^\ell(\mathbf{w})}$ is periodic. By Theorem 2.11, the minimal period of $\sqrt[2]{T^\ell(\mathbf{w})}$ is conjugate to S. As $|S|$ divides $|v|$ and v is primitive, we conclude that $|v| = |S|$. This contradicts the assumption that $u \notin \{S, L\}$. $\qquad\square$

Notice that the final paragraph of the proof of Theorem 5.2 does not use the fact that u is a solution to (3). Thus we obtain the following corollary.

Corollary 5.3. *If u is not a product of the words S and L but is conjugate to such a product, then u is not a solution to* (3).

Theorem 5.4. *If u is a primitive solution to* (3) *in $\mathcal{L}(\Omega)$ such that $|u| \geq 2|S|$, then $u = \gamma_k$ for some $k \geq 1$.*

Proof. Suppose that u is a primitive solution to (3) in $\mathcal{L}(\Omega)$ such that $|u| \geq 2|S|$. In particular, $u^2 \in \mathcal{L}(\Omega)$. First we aim to show that $|u|$ is a multiple of $|S|$. The word u^2 is a factor of a product of the words S and L. If two consecutive words S or L from the product are completely contained in u, then $|u|$ is a multiple of $|S|$ by Lemma 2.15. Suppose this is not the case. Since $|u| \geq 2|S|$, the word u^2 is a factor of a word $\diamond_1\diamond_2\diamond_3\diamond_4\diamond_5$, $\diamond_i \in \{S, L\}$, such that $\diamond_2$ and $\diamond_4$ are completely contained in u. It follows that $\diamond_4$ is an interior factor of $\diamond_1\diamond_2$ or $\diamond_2\diamond_3$. Consider the former case. Now also $\diamond_2$ is an interior factor of $\diamond_4\diamond_5$. Observe that $\diamond_4$ cannot occur at position 1 of $\diamond_1\diamond_2$ because $\diamond_1$ and $\diamond_4$ may differ only by their first two letters. Their common suffix of length $|S| - 2$ would otherwise be unary, and this is not possible as they both have the nonunary suffix S_6. Thus $\diamond_2 \neq \diamond_4$ as otherwise $\diamond_2$ would be an interior factor of its square contradicting the primitivity of $\diamond_2$. Exactly symmetric argument shows that $\diamond_2 \neq \diamond_5$. By Lemma 2.14 (i), LL does not occur, so we see that

21

$\diamond_4 = \diamond_5 = S$ and $\diamond_2 = L$. Further, we have $\diamond_1 = \diamond_3 = S$. The occurrence of $\diamond_2$ as an interior factor of $\diamond_4\diamond_5$ must be followed by the first letter of $\diamond_2$. Hence the letter following the prefix $\diamond_1\diamond_2$ of $\diamond_1\diamond_2\diamond_3\diamond_4\diamond_5$ must be the first letter of $\diamond_2$. This is a contradiction as $\diamond_2 \neq \diamond_3$. In the latter case, we analogously conclude that $\diamond_3 \neq \diamond_4$ and $\diamond_2 \neq \diamond_4$. This implies that $\diamond_2 = \diamond_3 = \diamond_5 = S$ and $\diamond_4 = L$. Again the occurrence of $\diamond_4$ as an interior factor of $\diamond_2\diamond_3$ must be followed by the first letter of $\diamond_4$, but this is impossible as the prefix $\diamond_1\diamond_2\diamond_3\diamond_4$ of $\diamond_1\diamond_2\diamond_3\diamond_4\diamond_5$ is followed by the first letter of $\diamond_5$. Thus we have shown that $|u|$ is a multiple of $|S|$.

Now u is conjugate to a product of the words S and L. If u is not itself a product of the words S and L, then it is not a solution to (3) by Corollary 5.3. Therefore u is a product of the words S and L, and there exists a word v in $\mathcal{L}(\Omega^\circ)$ such that $\sigma(v) = u$ and $v^2 \in \mathcal{L}(\Omega^\circ)$. Lemma 5.1 implies that u is conjugate to γ_k for some $k \geq 1$, and Theorem 5.2 shows that $u = \gamma_k$. $\qquad\square$

It is certainly possible that $\mathcal{L}(\Omega)$ contains short solutions to (3) that are not conjugates of S or L. First of all, the solution S, as a reversed standard word, can have squares of shorter reversed standard words as factors; these are also solutions to (3) in $\mathcal{L}(\Omega)$. Secondly, it is possible that there is a solution u such that $|S| < |u| < 2|S|$. For instance, if $S = 01010010$, then $\mathcal{L}(\Omega)$ contains the solution 01010010010 of length 11 as a factor of $SLSS$.

6   Further Remarks

In this paper, we constructed the fixed points $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$ using the simple substitution $S \mapsto LS^{2\mathfrak{c}}$, $L \mapsto S^{2\mathfrak{c}+1}$. We observe that any word $\mathbf{w} = a_0 a_1 a_2 \cdots$ over $\{S, L\}$ satisfying $a_i = a_{2i}$ for all i has the property $\sqrt{\mathbf{w}} = \mathbf{w}$ (see the remark of the final paragraph of Subsection 2.3). There are many other substitutions with suitable properties. Indeed, let n be an odd positive integer, and consider a word u of length n over $\{S, L\}$ having the property that $\sqrt{u^2} = u$. For each multiplicative set of $\mathbb{Z}_n$ given by the element 2, we can make an independent choice of a letter in $\{S, L\}$. For instance, if $n = 7$, then we have sets $\{0\}$, $\{1, 2, 4\}$, and $\{3, 5, 6\}$ meaning that the word $\diamond_0\diamond_1\diamond_1\diamond_2\diamond_1\diamond_2\diamond_2$ is suitable for any choice of $\diamond_0$, $\diamond_1$, and $\diamond_2$ in $\{S, L\}$. Letting $\diamond_1 = S$ and $\diamond_2 = L$ thus gives the substitution

$$S \mapsto LSSLSLL, \quad L \mapsto SSSLSLL,$$

which again generates two fixed points of the square root map (after S and L are substituted by suitable solutions to (3)). Mixing the applications of substitutions generated like this gives rise to even larger class of fixed points.

It is quite unclear which of the results given generalize to this larger class of fixed points and the associated subshifts. Theorem 3.2 still applies, but it is unknown if variants of Theorems 3.8 and 4.1 hold in general. It seems that a proof analogous to that of Theorem 4.1 works in the case of a single generating substitution, but it is unclear if it works with two substitutions.

Unlike in Theorem 5.4, there can in general be other solutions to (3) than those given directly by the generating substitution. This is because the substitution might itself have other solution generating patterns embedded. Consider for instance the substitution

$$\tau : \begin{aligned} S &\mapsto LSSLSSLSS \\ L &\mapsto SSSLSSLSS \end{aligned}$$

and the associated subshift Ω with the periodic part adjoined. The images of the letters have the word $(LSS)^2$ as a suffix. Now LSS is a solution to (3), so $\tau^k(LSS)$ is a solution to (3) for all $k \geq 0$. Since $(\tau^k(LSS))^2 \in \mathcal{L}(\Omega)$ for all $k \geq 0$, we see that $\mathcal{L}(\Omega)$ contains arbitrarily long solutions to (3) that are not equal to $\tau^k(S)$ or $\tau^k(L)$. Notice that these additional solutions cannot be used to produce additional fixed points as limits.

$\lvert S\rvert$	8	13	21	34	55	89	144	233	377	610	987	1597	2584	4181	6765
n	3.47	4.16	4.85	5.55	6.24	6.94	7.63	8.33	9.02	9.71	10.41	11.11	11.80	12.50	13.19

Table 2: Estimation of the relation of $\lvert S\rvert$ and n of Theorem 3.2 when S is a reversed Fibonacci word. The first two digits are correct.

There are many additional dynamical systems concepts that could be studied in our symbolic square root map setting. This time we finish our inquiry by making a remark on topological transitivity and topological mixing. A dynamical system (X, f) is *topologically transitive* if for every nonempty open sets A and B there exists an integer n such that $f^n(A) \cap B \neq \emptyset$. The dynamical system $(\Omega, \sqrt{\cdot})$ is not topologically transitive because the words beginning with 0 never map to words beginning with 1 (the cylinder sets $[0]$ and $[1]$ are open sets); Proposition 3.12 provides additional open sets with the same property. The same examples show that $(\Omega, \sqrt{\cdot})$ is not topologically mixing. A dynamical system (X, f) is *topologically mixing* if for any two nonempty open sets A and B there exists an integer n such that $f^N(A) \cap B \neq \emptyset$ for all $N \geq n$.

7 Open Problems

In Section 3, we gave some example values for the time n it takes any word in $\overline{\{S, L\}^\omega} \setminus \{S, L\}^\omega$ to map to S^ω or L^ω in the case that S is a reversed Fibonacci word. We conjecture that the quantity

$$\log_2\left(\frac{1 - \overline{\alpha}}{\min\{\lvert[S]\rvert, \lvert[L]\rvert\}}\right)$$

given in Remark 3.6 is close to the real value in general. Let us compute this quantity for the reversed Fibonacci words for comparison. In the case of the Fibonacci words, the slope α has continued fraction expansion $[0; 2, 1, 1, 1, \ldots]$, that is, $\alpha = 2 - \varphi$ where φ is the golden ratio $(1 + \sqrt{5})/2$. We replace the truncation $\overline{\alpha}$ by α. It is not difficult to see that $\min\{\lvert[S]\rvert, \lvert[L]\rvert\} = \lVert F_k \alpha \rVert$ for a Fibonacci number F_k; here the norm $\lVert \cdot \rVert$ measures the distance to closest integer. From elementary properties of continued fractions, it is straightforward to derive that $\lVert F_k \alpha \rVert^{-1} = \varphi F_k + F_{k-1}$ (save for some small values of k). Thus the above quantity now (approximately) equals

$$\log_2((\varphi - 1)(\varphi F_k + F_{k-1}))$$

for k such that $\lvert S\rvert = F_k$. Table 2 contains the values of this quantity for several F_k. Comparing these to the real values in Table 1, we see that at least in this special case our conjecture seems valid. The values of Table 1 seem to match the sequence A020909 in Sloane's *On-Line Encyclopedia of Integer Sequences* [11]. This sequence gives the number of bits in the base 2 representations of the Fibonacci numbers suggesting that the real value is $1 + \log_2 F_{k-1}$. The difficulty here lies in the first part of the proof of Theorem 3.2 where we showed that eventually $\sqrt[n]{\mathbf{w}}$ must be periodic. Unless we know that $\sqrt[n]{\mathbf{w}}$ is periodic, we cannot transfer the situation to the system of rational rotations where it is easy to estimate how long it takes for a word to map to S^ω or L^ω. Based on our computer experiments, it seems that typically $\mathbf{w}$ maps to a periodic quite early suggesting that the above quantity should be close to the truth. We have observed examples where it takes a longer time for $\mathbf{w}$ to become periodic, but in these cases the interval corresponding to the minimal period is already quite close to $[S]$ or $[L]$ on the circle balancing the situation. Overall, we do not have a clear picture of the situation. We propose the following open problem.

Open Problem. *Prove a good estimate on the number of steps required for any word in $\overline{\{S, L\}^\omega} \setminus \{S, L\}^\omega$ to map to S^ω or L^ω.*

Regarding the preimages of the word in the periodic part Ω_P we left uncharacterized, it would be interesting to know which words of Ω_P are images of words in Ω_A. Thus we propose the following open problem.

Open Problem. *Characterize the set $\sqrt{\Omega} \setminus \Omega_A$. How large is this set?*

In the case that S is a reversed Fibonacci word, it seems that

$$|\sqrt{\Omega} \setminus \Omega_A| \in \{|S|/2, (|S|+1)/2, (|S|-1)/2\},$$

so the size of $\sqrt{\Omega} \setminus \Omega_A$ seems to be approximately half of the size of Ω_P.

Acknowledgments

References

[1] G. Fici et al. Abelian powers and repetitions in Sturmian words. *Theoretical Computer Science* 635 (2016), 16–34.
DOI: 10.1016/j.tcs.2016.04.039.

[2] M. Lothaire. *Algebraic Combinatorics on Words*. Encyclopedia of Mathematics and Its Applications 90. Cambridge University Press, 2002.

[3] J. Peltomäki. Characterization of repetitions in Sturmian words: A new proof. *Information Processing Letters* 115.11 (2015), 886–891.
DOI: 10.1016/j.ipl.2015.05.011.

[4] J. Peltomäki. *Privileged Words and Sturmian Words*. Ph.D. dissertation. Turku, Finland: Turku Centre for Computer Science, University of Turku, 2016.
URL: http://urn.fi/URN:ISBN:978-952-12-3422-4.

[5] J. Peltomäki and M. Whiteland. A square root map on Sturmian words. *The Electronic Journal of Combinatorics* 24.1 (2017).
URL: http://www.combinatorics.org/ojs/index.php/eljc/article/view/v24i1p54.

[6] J. Peltomäki and M. Whiteland. More on the dynamics of the symbolic square root map. (Extended abstract). *Combinatorics on Words. Proceedings of the 11th International Conference, WORDS 2017*. Lecture Notes in Computer Science 10432. Springer, 2017, pp. 97–108.
DOI: 10.1007/978-3-319-66396-8.

[7] N. Pytheas Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*. Lecture Notes in Mathematics 1794. Springer, 2002.
DOI: 10.1007/b13861.

[8] M. Queffélec. *Substitution Dynamical Systems — Spectral Analysis*. 2nd edition. Lecture Notes in Mathematics 1294. Springer, 2010.
DOI: 10.1007/978-3-642-11212-6.

[9] K. Saari. *On the Frequency and Periodicity of Infinite Words*. Ph.D. dissertation. Turku, Finland: Turku Centre for Computer Science, University of Turku, 2008.
URL: http://users.utu.fi/kasaar/pubs/phdth.pdf.

[10]   K. Saari. Everywhere α-repetitive sequences and Sturmian words. *European Journal of Combinatorics* 31 (2010), 177–192.
DOI: 10.1016/j.ejc.2009.01.004.

[11]   N. J. A. Sloane. *The On-Line Encyclopedia of Integer Sequences.*
URL: http://oeis.org.