

STRUCTURE OF ATTRACTORS FOR (a, b) -CONTINUED FRACTION TRANSFORMATIONS

SVETLANA KATOK AND ILIE UGARCOVICI
(Communicated by Jens Marklof)

ABSTRACT. We study a two-parameter family of one-dimensional maps and related (a, b) -continued fractions suggested for consideration by Don Zagier. We prove that the associated natural extension maps have attractors with finite rectangular structure for the entire parameter set except for a Cantor-like set of one-dimensional Lebesgue zero measure that we completely describe. We show that the structure of these attractors can be “computed” from the data (a, b) , and that for a dense open set of parameters the Reduction theory conjecture holds, *i.e.*, every point is mapped to the attractor after finitely many iterations. We also show how this theory can be applied to the study of invariant measures and ergodic properties of the associated Gauss-like maps.

CONTENTS

1. Introduction	637
2. Theory of (a, b) -continued fractions	640
3. Attractor set for $F_{a,b}$	644
4. Cycle property	647
5. Finiteness condition implies finite rectangular structure	655
6. Finite rectangular structure of the attracting set	666
7. Reduction theory conjecture	670
8. Set of exceptions to the finiteness condition	671
9. Invariant measures and ergodic properties	687
References	689

1. INTRODUCTION

The standard generators $T(x) = x + 1$, $S(x) = -1/x$ of the modular group $SL(2, \mathbb{Z})$ were used classically to define piecewise continuous maps acting on

Received March 29, 2010; revised September 26, 2010.

2000 *Mathematics Subject Classification*: 37E05, 11A55, 11K50.

Key words and phrases: Continued fractions, attractor, natural extension, invariant measure.

We are grateful to Don Zagier for helpful discussions and the Max Planck Institute for Mathematics in Bonn for its hospitality and support.

IU: Partially supported by the NSF grant DMS-0703421.

the extended real line $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ that led to well-known continued fraction algorithms. In this paper we present a general method of constructing such maps suggested by Don Zagier, and study their dynamical properties and associated generalized continued fraction transformations.

Let \mathcal{P} be the two-dimensional parameter set

$$\mathcal{P} = \{(a, b) \in \mathbb{R}^2 \mid a \leq 0 \leq b, b - a \geq 1, -ab \leq 1\},$$

and consider the map $f_{a,b}: \bar{\mathbb{R}} \rightarrow \bar{\mathbb{R}}$ defined as

$$(1.1) \quad f_{a,b}(x) = \begin{cases} x+1 & \text{if } x < a, \\ -1/x & \text{if } a \leq x < b, \\ x-1 & \text{if } x \geq b. \end{cases}$$

Using the first-return map of $f_{a,b}$ to the interval $[a, b]$, denoted by $\hat{f}_{a,b}$, we introduce a two-dimensional family of continued fraction algorithms and study their properties. We mention here three classical examples: the case $a = -1/2$, $b = 1/2$ gives the “nearest-integer” continued fractions considered first by Hurwitz in [7], the case $a = -1$, $b = 0$ described in [23, 9] gives the “minus” (backward) continued fractions, while the situation $a = -1$, $b = 1$ was presented in [21, 10] in connection with a method of symbolically coding the geodesic flow on the modular surface following Artin’s pioneering work [3]. We remark that in the case $b - a = 1$, the class of one-parameter maps $f_{b-1,b}$ with $b \in [0, 1]$ is conceptually similar to the “ α -transformations” introduced by Nakada in [18] and studied subsequently in [4, 5, 15, 16, 17, 19, 20, 22]. Also, in the case $a = -b$, the one parameter family $f_{-b,b}$ with $1/2 \leq b \leq 1$ is related to a class of maps (factors of α -transformations) studied in [4, 17].

The main object of our study is a two-dimensional realization of the natural extension map of $f_{a,b}$, $F_{a,b}: \bar{\mathbb{R}}^2 \setminus \Delta \rightarrow \bar{\mathbb{R}}^2 \setminus \Delta$, $\Delta = \{(x, y) \in \bar{\mathbb{R}}^2 \mid x = y\}$, defined by

$$(1.2) \quad F_{a,b}(x, y) = \begin{cases} (x+1, y+1) & \text{if } y < a, \\ (-1/x, -1/y) & \text{if } a \leq y < b, \\ (x-1, y-1) & \text{if } y \geq b. \end{cases}$$

Numerical experiments led Don Zagier to conjecture that such a map $F_{a,b}$ has several interesting properties for all parameter pairs $(a, b) \in \mathcal{P}$:

REDUCTION THEORY CONJECTURE.

- (1) *The map $F_{a,b}$ possesses a global attractor set $D_{a,b} = \bigcap_{n=0}^{\infty} F^n(\bar{\mathbb{R}}^2 \setminus \Delta)$ on which $F_{a,b}$ is essentially bijective.*
- (2) *The set $D_{a,b}$ consists of two (or one, in degenerate cases) connected components each having finite rectangular structure, i.e., bounded by nondecreasing step-functions with a finite number of steps.*
- (3) *Every point (x, y) of the plane ($x \neq y$) is mapped to $D_{a,b}$ after finitely many iterations of $F_{a,b}$.*

REMARK 1.1. If one identifies a pair $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$ with a geodesic on the upper half-plane from x to y , then $F_{a,b}$ maps geodesics to geodesics. The attractor of

$F_{a,b}$ corresponds to reduced geodesics, hence the name of the conjecture and the reason why $F_{a,b}$ is also called the *reduction map*. This approach also has applications to coding symbolically geodesics on the modular surface [14].

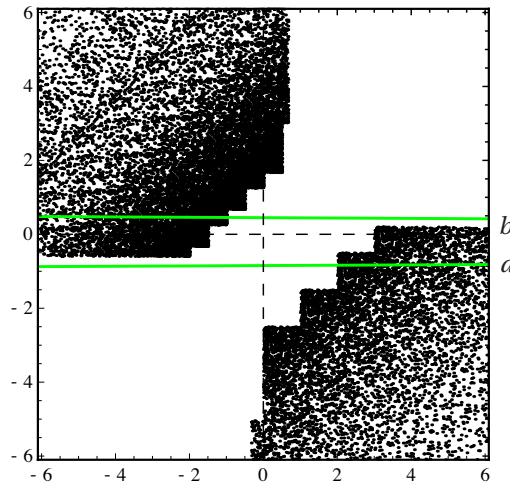


FIGURE 1. Attracting domain for Zagier's example: $a = -\frac{4}{5}$, $b = \frac{2}{5}$

Figure 1 shows the computer picture of such a the set $D_{a,b}$ with $a = -4/5$ and $b = 2/5$. It is worth mentioning that the complexity of the domain $D_{a,b}$ increases as (a, b) approach the line segment $b - a = 1$ in \mathcal{P} , a situation fully analyzed in what follows. The main result of this paper is the following:

MAIN RESULT. *There exists an explicit one-dimensional Lebesgue measure zero, uncountable set \mathcal{E} that lies on the diagonal boundary $b = a + 1$ of \mathcal{P} such that:*

- (i) *for all $(a, b) \in \mathcal{P} \setminus \mathcal{E}$, the map $F_{a,b}$ has an attractor $D_{a,b}$ satisfying properties (1) and (2) above;*
- (ii) *for an open and dense set in $\mathcal{P} \setminus \mathcal{E}$, property (3), and hence the Reduction theory conjecture, holds. For the rest of $\mathcal{P} \setminus \mathcal{E}$, property (3) holds for almost every point of $\mathbb{R}^2 \setminus \Delta$.*

We point out that this approach gives explicit conditions for the set $D_{a,b}$ to have finite rectangular structure that are satisfied, in particular, for all pairs (a, b) in the interior of the maximal parameter set \mathcal{P} . At the same time, it provides an effective algorithm for finding $D_{a,b}$, independent of the complexity of its boundary (*i.e.*, number of horizontal segments). The simultaneous properties satisfied by $D_{a,b}$, attracting set and bijectivity domain for $F_{a,b}$, is an essential feature that has not been exploited in earlier works. This approach makes the notions of reduced geodesic and dual expansion natural and transparent, with a potential for generalization to other Fuchsian groups. We remark that for “ α -transformations” [18, 16], explicit descriptions of the domain of the natural

extension maps have been obtained only for a subset of the parameter interval $[0, 1]$ (where the boundary has low complexity).

The paper is organized as follows. In Section 2, we develop the theory of (a, b) -continued fractions associated to the map $f_{a,b}$. In Section 3 we prove that the natural extension map $F_{a,b}$ possesses a *trapping region*; it will be used in Section 6 to study the attractor set for $F_{a,b}$. In Section 4 we further study the map $f_{a,b}$. Although it is discontinuous at $x = a, b$, one can look at two orbits of each of the discontinuity points. For generic (a, b) , these orbits meet after finitely many steps, forming a *cycle* that can be *strong* or *weak*, depending on whether the composite transformations at the end of each side of the cycle are equal or not. The values appearing in these cycles play a crucial role in the theory. Theorems 4.2 and 4.5 give necessary and sufficient conditions for b and a to have the *cycle property*. In Section 5 we introduce the *finiteness condition* using the notion of *truncated orbits* and prove that under this condition the map $F_{a,b}$ has a bijectivity domain $A_{a,b}$ with a finite rectangular structure that can be “computed” from the data (a, b) (Theorem 5.5). In Section 6 we define the attractor for the map $F_{a,b}$ by iterating the trapping region, and identify it with the earlier constructed set $A_{a,b}$ assuming the finiteness condition (Theorem 6.4). In Section 7 we prove that the Reduction theory conjecture holds under the assumption that both a and b have the strong cycle property, and that under the finiteness condition, property (3) holds for almost every point of $\mathbb{R}^2 \setminus \Delta$. In Section 8 we prove that the finiteness condition holds for all $(a, b) \in \mathcal{P}$ except for an uncountable set of one-dimensional Lebesgue measure zero that lies on the boundary $b = a + 1$ of \mathcal{P} (Theorem 8.11), and we present a complete description of this exceptional set. We conclude the proof of the Main Result by showing that the set of $(a, b) \in \mathcal{P}$ where a and b have the strong cycle property is open and dense in \mathcal{P} (Proposition 8.12). And, finally, in Section 9 we show how these results can be applied to the study of invariant measures and ergodic properties of the associated Gauss-like maps.

2. THEORY OF (a, b) -CONTINUED FRACTIONS

Consider $(a, b) \in \mathcal{P}$. The map $f_{a,b}$ defines what we call (a, b) -continued fractions using a generalized integral part function $\lfloor x \rfloor_{a,b}$: for any real x , let

$$\lfloor x \rfloor_{a,b} = \begin{cases} \lfloor x - a \rfloor & \text{if } x < a \\ 0 & \text{if } a \leq x < b \\ \lfloor x - b \rfloor & \text{if } x \geq b, \end{cases}$$

where $\lfloor x \rfloor$ denotes the integer part of x and $\lceil x \rceil = \lfloor x \rfloor + 1$.

Let us remark that the first-return map of $f_{a,b}$ to the interval $[a, b)$, $\hat{f}_{a,b}$, is given by the function

$$\hat{f}_{a,b}(x) = -\frac{1}{x} - \left\lfloor -\frac{1}{x} \right\rfloor_{a,b} = T^{-\lfloor -1/x \rfloor_{a,b}} S(x) \quad \text{if } x \neq 0, f(0) = 0.$$

We prove that any irrational number x can be expressed in a unique way as an infinite (a, b) -continued fraction

$$x = n_0 - \frac{1}{n_1 - \frac{1}{n_2 - \frac{1}{\ddots}}}$$

which we will denote by $[n_0, n_1, \dots]_{a,b}$ for short. The “digits” n_i , $i \geq 1$, are nonzero integers determined recursively by

$$(2.1) \quad n_0 = [x]_{a,b}, \quad x_1 = -\frac{1}{x - n_0}, \quad \text{and} \quad n_i = [x_i]_{a,b}, \quad x_{i+1} = -\frac{1}{x_i - n_i}.$$

In what follows, the notation $(\alpha_0, \alpha_1, \dots, \alpha_k)$ is used to write formally a “minus” continued fraction expression, where α_i are real numbers.

THEOREM 2.1. *Let x be an irrational number, $\{n_i\}$ the associated sequence of integers defined by (2.1) and $r_k = (n_0, n_1, \dots, n_k)$. Then the sequence r_k converges to x .*

Proof. We¹ start by proving that none of the pairs of type $(p, 1)$, $(-p, -1)$, with $p \geq 1$ are allowed to appear as consecutive entries of the sequence $\{n_i\}$. Indeed, if $n_{i+1} = 1$, then

$$b \leq x_{i+1} = -\frac{1}{x_i - n_i} < b + 1.$$

Therefore, $-\frac{1}{b} \leq x_i - n_i < -\frac{1}{b+1} \leq b - 1$ and $n_i < 0$. If $n_{i+1} = -1$, then

$$a - 1 \leq x_{i+1} = -\frac{1}{x_i - n_i} < a,$$

so $-\frac{1}{a-1} \leq x_i - n_i < -\frac{1}{a}$. But $a + 1 \leq -\frac{1}{a-1}$, thus $n_i > 0$.

With these two restrictions, the argument follows the lines of the proof for the classical case of minus (backward) continued fractions [9], where $n_i \geq 2$, for all $i \geq 1$. We define inductively two sequences of integers $\{p_k\}$ and $\{q_k\}$ for $k \geq -2$:

$$(2.2) \quad \begin{aligned} p_{-2} &= 0, & p_{-1} &= 1, & p_k &= n_k p_{k-1} - p_{k-2} & \text{for } k \geq 0; \\ q_{-2} &= -1, & q_{-1} &= 0, & q_k &= n_k q_{k-1} - q_{k-2} & \text{for } k \geq 0. \end{aligned}$$

We have the following properties:

- (i) there exists $l \geq 1$ such that $|q_l| < |q_{l+1}| < \dots < |q_k| < \dots$;
- (ii) $(n_0, n_1, \dots, n_k, \alpha) = (\alpha p_k - p_{k-1}) / (\alpha q_k - q_{k-1})$ for any real number α ;
- (iii) $p_k q_{k+1} - p_{k+1} q_k = 1$.

¹The authors proved initially the convergence statement assuming $-1 \leq a \leq 0 \leq b \leq 1$, and two Penn State REU students, Tra Ho and Jesse Barbour, worked on the proof for a and b outside of this compact triangular region. The unified proof presented here uses some of their ideas.

Let us prove property (i) since the proof is more intricate than in the classical case. Obviously $1 = q_0 \leq |q_1| = |n_1|$, $q_2 = n_2 q_1 - q_0 = n_2 n_1 - 1$. Notice that $|q_2| > |q_1|$ unless $n_1 = 1$, $n_2 = 2$ or $n_1 = -1$, $n_2 = -2$. We analyze the situation $n_1 = 1$, $n_2 = 2$. This implies that $q_3 = n_3(n_2 n_1 - 1) - n_1 = n_3 - n_1$, so $|q_3| > |q_2|$, unless $n_3 = 2$. Notice that it is impossible to have $n_i = 2$ for all $i \geq 2$, because x is irrational and the minus continued fraction expression consisting only of two's, $(2, 2, \dots)$, has numerical value 1. Therefore, there exists $l \geq 1$ such that $n_{l+1} \neq 1, 2$. This implies that $|q_{l+1}| > |q_l|$. We continue to proceed by induction. Assume that property (i) is satisfied up to k -th term, $k > l$. If $|n_{k+1}| \geq 2$, then

$$|q_{k+1}| \geq |n_{k+1}| \cdot |q_k| - |q_{k-1}| \geq 2|q_k| - |q_{k-1}| > |q_k|.$$

If $n_{k+1} = 1$, then $q_{k+1} = q_k - q_{k-1}$. Since $q_k = n_k q_{k-1} - q_{k-2}$ with $n_k < 0$, one gets

$$q_{k-1} = \frac{q_k + q_{k-2}}{n_k}.$$

We analyze the two possible situations

- If $q_k > 0$ then $|q_{k-2}| < q_k$, so $q_k + q_{k-2} > 0$ and $q_{k-1} < 0$. This implies that $q_{k+1} = q_k - q_{k-1} \geq q_k > 0$.
- If $q_k < 0$, then $|q_{k-2}| < -q_k$, so $q_k + q_{k-2} < 0$ and $q_{k-1} > 0$. This implies that $q_{k+1} = q_k - q_{k-1} < q_k < 0$.

Thus $|q_k| < |q_{k+1}|$. A similar argument shows that the inequality remains true if $n_{k+1} = -1$.

To finish the proof we use properties (i)–(iii). First, we see that $r_k = p_k/q_k$ for $k \geq 0$. Moreover, the sequence r_k is a Cauchy sequence because

$$|r_{k+1} - r_k| = \frac{1}{|q_k q_{k+1}|} \leq \frac{1}{(k-l)^2} \quad \text{for } k > l.$$

Hence r_k is convergent.

In order to prove that r_k converges to x , we write $x = (n_0, n_1, \dots, n_k, x_{k+1})$, and look only at those terms $(n_0, n_1, \dots, n_k, x_{k+1})$ with $|x_{k+1}| \geq 1$. There are infinitely many such terms: indeed, if $-1 \leq a < b \leq 1$, then $|x_{k+1}| \geq 1$ for all $k \geq 1$; if $a < -1$ and $|x_{k+1}| < 1$, then $b \leq x_{k+1} < 1$, so $x_{k+2} = -1/(x_{k+1} - 1) \geq 1$; if $b > 1$ and $|x_{k+1}| < 1$, then $-1 < x_{k+1} < a$, so $x_{k+2} = -1/(x_{k+1} + 1) \geq 1$. Therefore, the corresponding subsequence $r_k = p_k/q_k$ satisfies

$$\begin{aligned} \left| \frac{p_k}{q_k} - x \right| &= \left| \frac{p_k}{q_k} - \frac{p_k x_{k+1} - p_{k-1}}{q_k x_{k+1} - q_{k-1}} \right| = \frac{1}{|q_k(q_k x_{k+1} - q_{k-1})|} \\ &\leq \frac{1}{|q_k|(|q_k||x_{k+1}| - |q_{k-1}|)} \leq \frac{1}{|q_k|} \rightarrow 0. \end{aligned}$$

We showed that the convergent sequence $r_k = p_k/q_k$ has a subsequence convergent to x , therefore the whole sequence converges to x . \square

REMARK 2.2. One can construct (a, b) -continued fraction expansions for rational numbers, too. However, such expansions will terminate after finitely many steps if $b \neq 0$. If $b = 0$, the expansions of rational numbers will end with a tail of

2's, since $0 = (1, 2, 2, \dots)$. In this latter case, properties (i)–(iii) listed above still hold.

REMARK 2.3. It is easy to see that if the (a, b) -continued fraction expansion of a real number is eventually periodic, then the number is a quadratic irrationality.

It is not our intention to present in this paper some of the typical number-theoretical results that can be derived for the class of (a, b) -continued fractions. However, we state and prove a simple version about (a, b) -continued fractions with “bounded digits”. For the regular continued fractions, this is a classical result due to Borel and Bernstein (see [6, Theorem 196] for an elementary treatment). We are only concerned with (a, b) -expansions that are written with two consecutive digits, a result explicitly needed in Sections 7 and 8.

PROPOSITION 2.4. *The set $\Gamma_{a,b}^{(m)} = \{x = [0, n_1, n_2, \dots]_{a,b} \mid n_k \in \{m, m+1\}\}$ has zero Lebesgue measure for every $m \geq 1$.*

Proof. First, notice that if $m = 1$, then the set $\Gamma_{a,b}^{(1)}$ has obviously zero measure, since the pairs $(2, 1)$ and $(-2, -1)$ are not allowed in the (a, b) -expansions.

Assume $m \geq 2$. Notice that $\Gamma_{a,b}^{(m)} \subset \Gamma_{0,-1}^{(m)}$ since a formal continued fraction $x = (0, n_1, n_2, \dots)$ with $n_k \in \{m, m+1\}$ coincides with its “minus” (backward) continued fraction expansion ($a = -1, b = 0$), $x = [0, n_1, n_2, \dots]_{-1,0}$. The reason is that any sequence of digits $n_i \geq 2$ gives a valid “minus” continued fraction expansion.

In what follows, we study the set $\Gamma_{0,-1}^{(m)}$. For practical reasons we will drop the subscript $(0, -1)$. It is worth noticing that the result for $\Gamma_{0,-1}^{(m)}$ does not follow automatically from the result about regular continued fractions, since there are numbers for which the $(0, -1)$ -expansion has only digits 2 and 3, while the regular continued fractions expansion has unbounded digits. We follow the approach of [6, Theorem 196] and estimate the size of the set $\Gamma_{n_1, n_2, \dots, n_k}^{(m)} \subset \Gamma^{(m)}$ with the digits $n_1, n_2, \dots, n_k \in \{m, m+1\}$ being fixed. In this particular case, the recursive relation (2.2) implies that $1 = q_1 < q_2 < \dots < q_k$. If $x \in \Gamma_{n_1, n_2, \dots, n_k}^{(m)}$, then

$$(0, n_1, n_2, \dots, n_k - 1) \leq x < (0, n_1, n_2, \dots, n_k).$$

Using property (ii), the end-points of such an interval $I_{n_1, \dots, n_k}^{(m)}$ are given by

$$\frac{(n_k - 1)p_{k-1} - p_{k-2}}{(n_k - 1)q_{k-1} - q_{k-2}} \quad \text{and} \quad \frac{n_k p_{k-1} - p_{k-2}}{n_k q_{k-1} - q_{k-2}}$$

and the length of this interval is

$$l(I_{n_1, \dots, n_k}^{(m)}) = \frac{1}{(n_k q_{k-1} - q_{k-2})((n_k - 1)q_{k-1} - q_{k-2})} = \frac{1}{q_k(q_k - q_{k-1})}$$

by using that $p_{k-2}q_{k-1} - p_{k-1}q_{k-2} = 1$ and $q_k = n_k q_{k-1} - q_{k-2}$.

Denote by $\Gamma_k^{(m)}$ the set of numbers in $[-1, 0)$ with $(-1, 0)$ -continued fraction digits $n_1, n_2, \dots, n_k \in \{m, m+1\}$. The set $\Gamma_k^{(m)}$ is part of the set

$$I_k^{(m)} = \bigcup_{n_1, \dots, n_k \in \{m, m+1\}} I_{n_1, \dots, n_k}^{(m)}.$$

We have the following relation:

$$I_{k+1}^{(m)} = \bigcup_{n_1, \dots, n_k \in \{m, m+1\}} I_{n_1, \dots, n_k, m}^{(m)} \cup I_{n_1, \dots, n_k, m+1}^{(m)}.$$

If x lies in $I_{n_1, \dots, n_k, m}^{(m)} \cup I_{n_1, \dots, n_k, m+1}^{(m)}$, then

$$(0, n_1, n_2, \dots, n_k, m-1) \leq x < (0, n_1, n_2, \dots, n_k, m+1).$$

The length of this interval is

$$l(I_{n_1, \dots, n_k, m}^{(m)} \cup I_{n_1, \dots, n_k, m+1}^{(m)}) = \frac{2}{((m+1)q_k - q_{k-1})((m-1)q_k - q_{k-1})}.$$

Now we estimate the ratio

$$\begin{aligned} \frac{l(I_{n_1, \dots, n_k, m}^{(m)} \cup I_{n_1, \dots, n_k, m+1}^{(m)})}{l(I_{n_1, n_2, \dots, n_k}^{(m)})} &= \frac{2q_k(q_k - q_{k-1})}{((m+1)q_k - q_{k-1})((m-1)q_k - q_{k-1})} \\ &\leq \frac{2q_k}{(m+1)q_k - q_{k-1}} \\ &\leq \frac{2q_k}{3q_k - q_{k-1}} = \frac{2}{3 - q_{k-1}/q_k} \\ &\leq \frac{2k}{2k+1} \end{aligned}$$

since $q_{k-1}/q_k \leq (k-1)/k$. Indeed, if $n_1 = \dots = n_k = 2$, then $q_{k-1}/q_k = (k-1)/k$; if some $n_j > 2$, then $q_{k-1}/q_k \leq 1/2$ from (2.2). This proves that for every $k \geq 1$

$$I_{k+1}^{(m)} \leq \frac{2k}{2k+1} I_k^{(m)}$$

so

$$l(I_k^{(m)}) \leq \frac{2 \cdot 4 \cdots (2k-2)}{3 \cdot 5 \cdots (2k-1)} \cdot l(I_1^{(m)}) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Therefore, in all cases, $l(I_k^{(m)}) \rightarrow 0$ as $k \rightarrow \infty$. Since $\Gamma^{(m)} \subset I_k^{(m)}$ for every $k \geq 1$, the proposition follows. \square

REMARK 2.5. By a similar argument, the set

$$\Gamma_{a,b}^{(-m)} = \{x = [0, n_1, n_2, \dots]_{a,b} \mid n_k \in \{-m, -m-1\}\}$$

has zero Lebesgue measure for every $m \geq 1$.

3. ATTRACTOR SET FOR $F_{a,b}$

The reduction map $F_{a,b}$ defined by (1.2) has a trapping domain, *i.e.*, a closed set $\Theta_{a,b} \subset \bar{\mathbb{R}}^2 \setminus \Delta$ with the following properties:

- (i) for every pair $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$, there exists a positive integer N such that $F_{a,b}^N(x, y) \in \Theta_{a,b}$;
- (ii) $F_{a,b}(\Theta_{a,b}) \subset \Theta_{a,b}$.

THEOREM 3.1. *The region $\Theta_{a,b}$ consisting of two connected components (or one if $a = 0$ or $b = 0$) defined as*

$$\Theta_{a,b}^u = \begin{cases} [-\infty, -1] \times [b-1, \infty] \cup [-1, 0] \times [-1/a, \infty] & \text{if } b \geq 1, a \neq 0, \\ \emptyset & \text{if } a = 0, \\ [-\infty, -1] \times [b-1, \infty] \cup [-1, 0] \times [\min(-\frac{b}{b-1}, -\frac{1}{a}), \infty] & \text{if } 0 < b < 1, \\ \cup [0, 1] \times [-\frac{1}{b-1}, \infty] & \end{cases}$$

and

$$\Theta_{a,b}^l = \begin{cases} [0, 1] \times [-\infty, -1/b] \cup [1, \infty] \times [-\infty, a+1] & \text{if } a \leq -1, b \neq 0, \\ \emptyset & \text{if } b = 0, \\ [-1, 0] \times [-\infty, -\frac{1}{a+1}] \cup [0, 1] \times [-\infty, \max(\frac{a}{a+1}, -\frac{1}{b})] & \text{if } a > -1, \\ \cup [1, \infty] \times [-\infty, a+1] & \end{cases}$$

is the trapping region for the reduction map $F_{a,b}$.

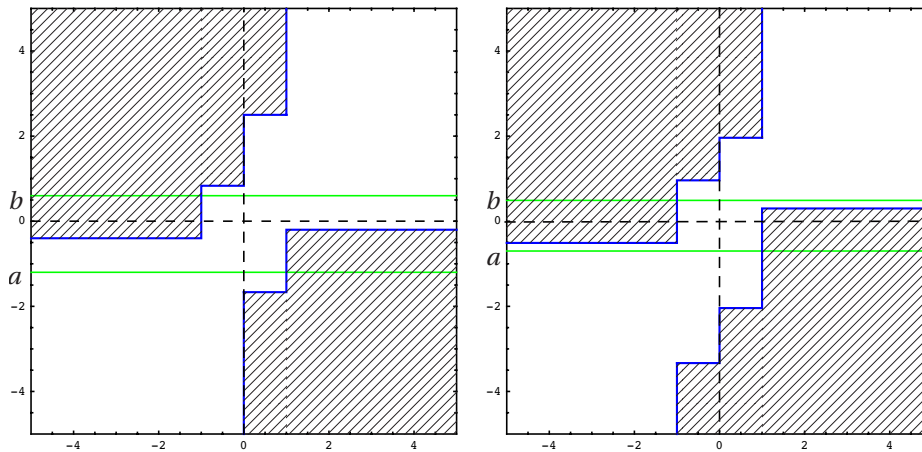


FIGURE 2. Typical trapping regions: case $a < -1, 0 < b < 1$ (left); case $-1 < a < 0 < b < 1$ (right)

Proof. The fact that the region $\Theta_{a,b}$ is $F_{a,b}$ -invariant is verified by a direct calculation. We focus our attention on the attracting property of $\Theta_{a,b}$. Let $(x, y) \in \mathbb{R}^2 \setminus \Delta$, write $y = [n_0, n_1, \dots]_{a,b}$, and construct the following sequence of real pairs $\{(x_k, y_k)\}$ ($k \geq 0$) defined by $x_0 = x, y_0 = y$, and

$$y_{k+1} = ST^{-n_k} \dots ST^{-n_1} ST^{-n_0} y, \quad x_{k+1} = ST^{-n_k} \dots ST^{-n_1} ST^{-n_0} x.$$

If y is rational and its (a, b) -expansion terminates $y = [n_0, n_1, \dots, n_l]_{a,b}$, then $y_{l+1} = \pm\infty$, so (x, y) lands in $\Theta_{a,b}$ after finitely many iterations. If y has an infinite (a, b) -expansion, then $y_{k+1} = [n_{k+1}, n_{k+2}, \dots]_{a,b}$, and $y_{k+1} \geq -1/a$

$y_{k+1} \leq -1/b$ for $k \geq 0$. Also,

$$y = T^{n_0} S T^{n_1} S \dots T^{n_k} S(y_{k+1}) = \frac{p_k y_{k+1} - p_{k-1}}{q_k y_{k+1} - q_{k-1}}$$

$$x = T^{n_0} S T^{n_1} S \dots T^{n_k} S(x_{k+1}) = \frac{p_k x_{k+1} - p_{k-1}}{q_k x_{k+1} - q_{k-1}}.$$

Hence,

$$x_{k+1} = \frac{q_{k-1}x - p_{k-1}}{q_k x - p_k} = \frac{q_{k-1}}{q_k} + \frac{1}{q_k^2(p_k/q_k - x)} = \frac{q_{k-1}}{q_k} + \varepsilon_k.$$

Since $p_k/q_k \rightarrow y \neq x$, and $|q_k| \rightarrow \infty$, we obtain that $|\varepsilon_k| < 1/q_k$ for large enough k . But one of the properties of the sequence $\{q_k\}$ is that $|q_{k-1}| < |q_k|$ for large enough k (see property (i) in the proof of Theorem 2.1 and Remark 2.2.) Therefore,

$$|x_{k+1}| \leq \frac{|q_{k-1}|}{|q_k|} + \frac{1}{|q_k|} \leq 1 \quad \text{for large enough } k.$$

We proved that there exists $N > 0$, such that

$$F_{a,b}^N(x, y) = ST^{-n_k} \dots ST^{-n_1} ST^{-n_0}(x, y) \in [-1, 1] \times ([-1/a, \infty] \cup [-\infty, -1/b]).$$

The point $F_{a,b}^N(x, y) =: (\tilde{x}, \tilde{y})$ belongs to $\Theta_{a,b}$, unless $b < 1$ and $(\tilde{x}, \tilde{y}) \in [0, 1] \times [-1/a, -1/(b-1)]$ or $a > -1$ and $(\tilde{x}, \tilde{y}) \in [-1, 0] \times [-1/b, -1/(a+1)]$.

Let us study the next iterates of $(\tilde{x}, \tilde{y}) \in [0, 1] \times [-1/a, -1/(b-1)]$. If $\tilde{y} \geq b+1$ then

$$F_{a,b}^2(\tilde{x}, \tilde{y}) = (\tilde{x} - 2, \tilde{y} - 2) \in [-1, 1] \times [b-1, \infty],$$

so $F_{a,b}^2(\tilde{x}, \tilde{y}) \in \Theta_{a,b}$. If it so happens that $-1/a \leq \tilde{y} < b+1$, then

$$F_{a,b}(\tilde{x}, \tilde{y}) = (\tilde{x} - 1, \tilde{y} - 1) \in [-1, 0] \times [0, b]$$

and

$$F_{a,b}^2(\tilde{x}, \tilde{y}) = ST^{-1}(\tilde{x}, \tilde{y}) \in [0, \infty] \times [-1/b, \infty] \subset \Theta_{a,b}.$$

Similarly, if $(x, y) \in [-1, 0] \times [-1/b, -1/(a+1)]$, then $F_{a,b}^2(x, y) \in \Theta_{a,b}$.

Notice that if $a = 0$, then $y_{k+1} \leq -1/b$ for all $k \geq 0$ (so $\Theta_{a,b}^u = \emptyset$) and if $b = 0$, then $y_{k+1} \geq -1/a$ for all $k \geq 0$ (so $\Theta_{a,b}^l = \emptyset$). □

Using the trapping region described in Theorem 3.1 we define the associated *attractor set*

$$(3.1) \quad D_{a,b} = \bigcap_{n=0}^{\infty} D_n,$$

where $D_n = \bigcap_{i=0}^n F_{a,b}^i(\Theta_{a,b})$.

REMARK 3.2. In the particular cases when $a = 0$ and $b \geq 1$, or $b = 0$ and $a \leq -1$, or $(a, b) = (-1, 1)$, the trapping regions

$$\Theta_{0,b} = [-1, 0] \times [-\infty, -1] \cup [0, 1] \times [-\infty, 0] \cup [1, \infty] \times [-\infty, 1]$$

$$\Theta_{a,0} = [-\infty, -1] \times [-1, \infty] \cup [-1, 0] \times [0, \infty] \cup [0, 1] \times [1, \infty]$$

$$\Theta_{-1,1} = [-\infty, -1] \times [-1, \infty] \cup [-1, 0] \times [1, \infty] \cup [0, 1] \times [-\infty, -1] \cup [1, \infty] \times [-\infty, 0]$$

are also bijectivity domains for the corresponding maps $F_{a,b}$. Therefore, in these cases the attractor $D_{a,b}$ coincides with the trapping region $\Theta_{a,b}$, so the properties mentioned in the introduction are obviously satisfied. In what follows, all our considerations will exclude these degenerate cases.

4. CYCLE PROPERTY

In what follows, we simplify the notations for $f_{a,b}$, $[\cdot]_{a,b}$, $\hat{f}_{a,b}$ and $F_{a,b}$ to f , $[\cdot]$, \hat{f} and F , respectively, assuming implicitly their dependence on parameters a, b . We will use the notation f^n (or \hat{f}^n) for the n -times composition operation of f (or \hat{f}). Also, for a given point $x \in (a, b)$ the notation $\hat{f}^{(k)}$ means the transformation of type $T^i S$ (i is an integer) such that

$$\hat{f}^k(x) = \hat{f}^{(k)} \hat{f}^{(k-1)} \dots \hat{f}^{(2)} \hat{f}^{(1)}(x),$$

where $\hat{f}^{(1)}(x) = \hat{f}(x)$.

The map f is discontinuous at $x = a, b$. However, we can associate to each a and b two forward orbits: to a we associate the *upper orbit* $\mathcal{O}_u(a) = \{f^n(Sa)\}$, and the *lower orbit* $\mathcal{O}_\ell(a) = \{f^n(Ta)\}$; to b the *lower orbit* $\mathcal{O}_\ell(b) = \{f^n(Sb)\}$ and the *upper orbit* $\mathcal{O}_u(b) = \{f^n(T^{-1}b)\}$. We use the convention that if an orbit hits one of the discontinuity points a or b , then the next iterate is computed according to the lower or upper location. For example, if the lower orbit of b hits a , then the next iterate is Ta , if the upper orbit of b hits a then the next iterate is Sa .

Now we explore the patterns in the above orbits. The following property plays an essential role in studying the map f .

DEFINITION 4.1. We say that the point a has the *cycle property*² if for some nonnegative integers m_1 and k_1 ,

$$f^{m_1}(Sa) = f^{k_1}(Ta) = c_a.$$

We will refer to the set $\{Ta, fTa, \dots, f^{k_1-1}Ta\}$ as the *lower side of the a -cycle*, to the set $\{Sa, fSa, \dots, f^{m_1-1}Sa\}$ as the *upper side of the a -cycle*, and to c_a as the *end of the a -cycle*. If the transformations $f^{k_1}T$ and $f^{m_1}S$ at the end of each side of the a -cycle are equal, we say that a has the *strong cycle property*. Otherwise, we say that a has the *weak cycle property*.

Similarly, we say that b has the *cycle property* if for some nonnegative integers m_2 and k_2 ,

$$f^{k_2}(Sb) = f^{m_2}(T^{-1}b) = c_b.$$

We will refer to the set $\{Sb, fSb, \dots, f^{k_2-1}Sb\}$ as the *lower side of the b -cycle*, to the set $\{T^{-1}b, fT^{-1}b, \dots, f^{m_2-1}T^{-1}b\}$ as the *upper side of the b -cycle*, and to c_b as the *end of the b -cycle*. If the transformations $f^{k_2}T$ and $f^{m_2}S$ at the end of each side of the b -cycle are equal, we say that b has the *strong cycle property*. Otherwise, we say that b has the *weak cycle property*.

²A similar “matching” condition is used by the authors of [5, 20] to study the measure-theoretic entropy of α -transformations.

It turns out that the cycle property is the prevalent pattern. It can be analyzed and described explicitly by partitioning the parameter set \mathcal{P} based on the first digits of Sb , STa , and Sa , $ST^{-1}b$, respectively. Figure 3 shows a part of the countable partitions, with B_{-1}, B_{-2}, \dots denoting the regions where Sb has the first digit $-1, -2, \dots$, and A_1, A_2, \dots , denoting the regions where Sa has the first digit $1, 2, \dots$. For most of the parameter region, the cycles are short: everywhere except for the narrow triangular regions shown in Figure 3 the cycles for both a and b end after the first return to $[a, b)$. However, there are Cantor-like recursive sets where the lengths of the cycles can be arbitrarily long. Part of this more complex structure, studied in details in Section 8, can be seen as narrow triangular regions close to the boundary segment $b - a = 1$.

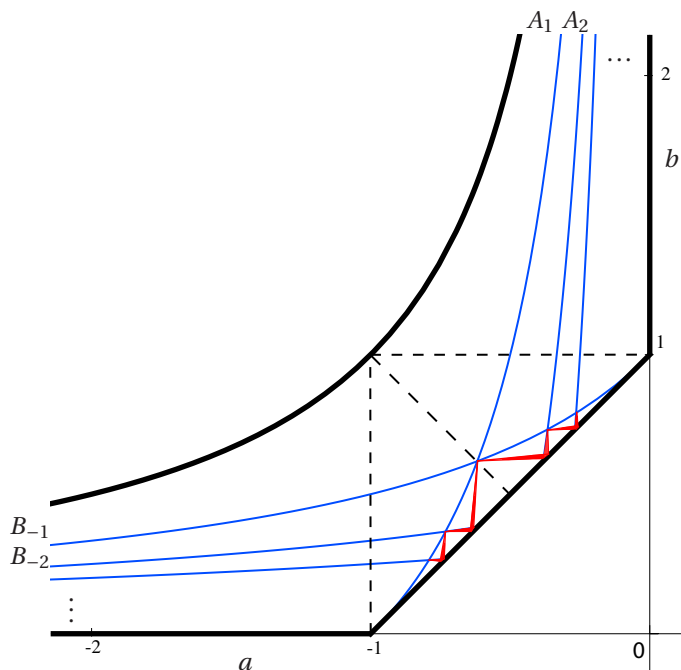


FIGURE 3. The parameter set \mathcal{P} and its partition

By symmetry of the parameter set \mathcal{P} with respect to the line $b = -a$, $(a, b) \mapsto (-b, -a)$, we may assume that $b \leq -a$ and concentrate our attention to this subset of \mathcal{P} .

The structure of the set where the cycle property holds for b is described next for the part of the parameter region with $0 < b \leq -a < 1$. We make use extensively of the first-return map \hat{f} .

THEOREM 4.2. *Let $(a, b) \in \mathcal{P}$, $0 < b \leq -a < 1$, and $m \geq 1$ with $a \leq T^m Sb < a + 1$.*

(I) *Suppose that there exists $n \geq 0$ such that*

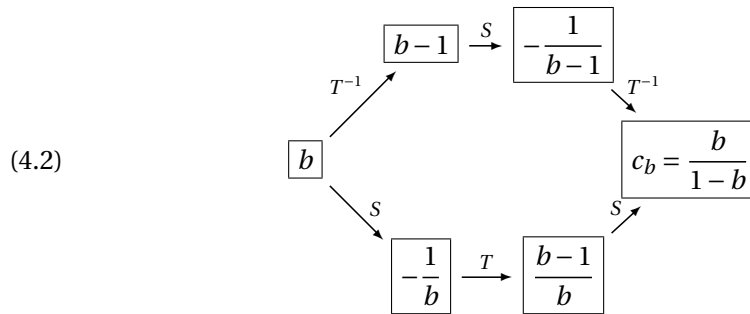
$$\hat{f}^k T^m Sb \in \left(\frac{b}{b+1}, a+1 \right) \text{ for } k < n, \text{ and } \hat{f}^n T^m Sb \in \left[a, \frac{b}{b+1} \right].$$

- (i) If $\hat{f}^n T^m S b \in (a, b/(b+1))$, then b has the cycle property; the cycle property is strong if and only if $\hat{f}^n T^m S b \neq 0$.
 - (ii) If $\hat{f}^n T^m S b = a$, then b has the cycle property if and only if a has the cycle property.
 - (iii) If $\hat{f}^n T^m S b = b/(b+1)$, then b does not have the cycle property, but the orbits of Sb and $T^{-1}b$ are periodic.
- (II) If $\hat{f}^k T^m S b \in (b/(b+1), a+1)$ for all $k \geq 0$, then b does not have the cycle property.

Proof. (I). In the case $m = 1$, and assuming $a < TSb < a + 1$ we have

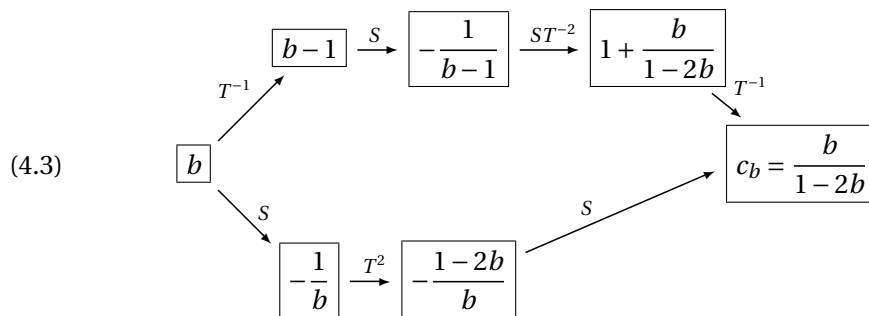
$$(4.1) \quad a < 1 - \frac{1}{b} < \frac{b}{b+1},$$

and the cycle relation for b can be explicitly described as

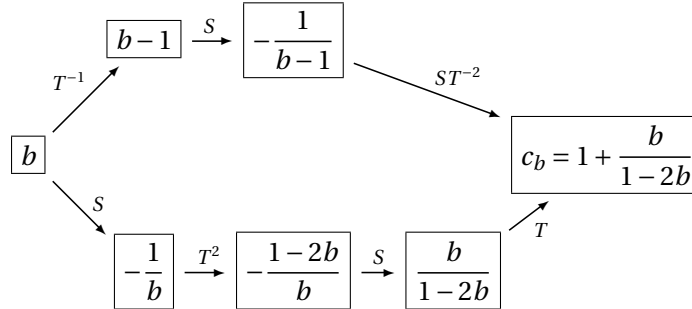


In the particular situation that $TSb = a$, the lower orbit of b hits a and continues to $a + 1$, while the upper orbit hits $b/(1 - b) = -1/a$. This means that the iterates will follow the lower and upper orbits of a , respectively, thus statement (ii) holds. Since the second inequality (4.1) is strict, the case (iii) cannot occur.

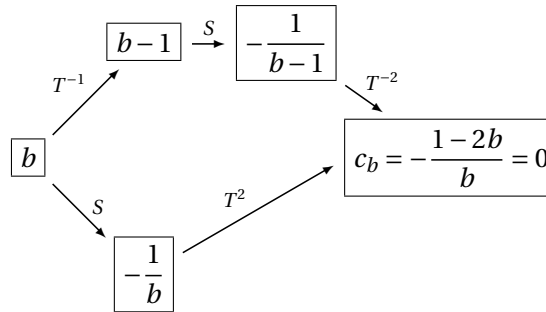
For the case $m = 2$ (and assuming $T^2 S b \neq a$), we analyze the following situations: if $b < 1/2$, then $2 - 1/b < 0$, and the cycle relation is



If $b > 1/2$ we have $0 < 2 - 1/b \leq b/(b + 1)$, since we must also have $2 - 1/b < a + 1$, i.e., $b \leq 1/(1 - a)$, and the cycle relation is



The above cycles are strong. If $b = 1/2$ the cycle relation is



It is easy to check that this cycle is weak. In the particular situation when $T^2Sb = a$, the lower orbit of b hits a , and continues with $a + 1$, while the upper orbit still hits $b/(1 - 2b) = -1/a$. This means that the iterates will follow the lower and upper orbits of a , respectively, and statement (ii) holds. The relation $2 - 1/b = b/(b + 1)$ implies $b = \frac{-1 + \sqrt{5}}{2}$ that does not have the cycle property and the orbits of Sb and $T^{-1}b$ are periodic; this is the only possibility for (iii) to hold.

The situation for $m \geq 3$ is more intricate. First we will need the following lemmas.

LEMMA 4.3. *Suppose $STSx = y$. Then each of the following are true.*

- (a) *If $TSb \leq x < a$, then $b - 1 \leq y < a/(1 - a)$.*
- (b) *If $a \leq x < b/(b + 1)$, then $a/(1 - a) \leq y < b$.*
- (c) *If $b/(b + 1) \leq x < a + 1$, then $b \leq y < a/(1 - a) + 1$.*
- (d) *If $x = 0$, then $y = 0$.*

Proof. Applying STS to the corresponding inequalities, we obtain

(a)
$$b - 1 = STSTSb \leq y < STSa = \frac{a}{1 - a}$$

(b)
$$\frac{a}{1 - a} = STSa \leq y < STSTSTb = b$$

(c)
$$b = STSTSTb \leq y < STSTa = T^{-1}Sa \leq \frac{1}{1 - a} = \frac{a}{1 - a} + 1,$$

where the last inequality is valid for $a \leq \frac{1-\sqrt{5}}{2}$, which is true in the considered region $b \leq 1/(2-a)$. Relation (d) is obvious. \square

LEMMA 4.4. *Suppose that for all $k < n$*

$$(4.4) \quad \frac{b}{b+1} < \hat{f}^k T^m S b < a+1.$$

Then

- (i) *for $0 \leq k \leq n$, in the lower orbit of b , $\hat{f}^{(k)} = T^m S$ or $T^{m+1} S$; in the upper orbit of b , $\hat{f}^{(k)} = T^{-i} S$ with $i = 2$ or 3 ;*
- (ii) *there exists $p > 1$ such that*

$$(4.5) \quad (STS) \hat{f}^n T^m S = (T^{-2} S) \hat{f}^p T^{-1}.$$

Proof. (i). Applying $T^m S$ to the inequality (4.4), we obtain

$$a-1 \leq T^{m-1} S b = T^m S T S T b < T^m S \hat{f}^k T^m S b \leq T^m S T a \leq T^m S b < a+1.$$

Therefore $\hat{f}^{(k+1)} = T^m S$ or $T^{m+1} S$. Since $\hat{f}^{(0)} = T^m S$, we conclude that $\hat{f}^{(k)} = T^m S$ or $T^{m+1} S$ for $0 \leq k \leq n$.

(ii). In order to determine the upper side of the b -cycle, we will use the following relation in the group $SL(2, \mathbb{Z})$ obtained by concatenation of the “standard” relations (from right to left)

$$(4.6) \quad (STS) T^i S = (T^{-2} S)^{i-1} T^{-1} \quad (i \geq 1),$$

and Lemma 4.3 repeatedly. The proof is by induction on n . For the base case $n = 1$, we have

$$\frac{b}{b+1} < T^m S b < a+1.$$

Then for $1 \leq i \leq m-1$, $T^i S b$ satisfies Lemma 4.3(a). Hence

$$b-1 < (T^{-2} S)^{i-2} T^{-1} b < \frac{a}{1-a},$$

which means that on the upper side of the b -cycle $\hat{f}^{(1)} = T^{-1}$ and $\hat{f}^{(i)} = T^{-2} S$ for $1 < i \leq m-1$. Using (4.6) for $i = m$, we obtain

$$(STS) T^m S = (T^{-2} S)^{m-1} T^{-1} = (T^{-2} S) \hat{f}^{m-2} T^{-1}.$$

In other words, (4.5) holds with $p = m-2$. Now suppose the statement holds for $n = n_0$, and for all $k < n_0 + 1$ we have

$$\frac{b}{b+1} < \hat{f}^k T^m S b < a+1.$$

By the induction hypothesis, there exists $p_0 > 1$ such that

$$(4.7) \quad (STS) \hat{f}^{n_0} T^m S = (T^{-2} S) \hat{f}^{p_0} T^{-1}.$$

But since

$$\frac{b}{b+1} < \hat{f}^{n_0} T^m S b < a+1,$$

the condition of Lemma 4.3(c) is satisfied. Hence,

$$b < (T^{-2} S) \hat{f}^{p_0} T^{-1} b < \frac{a}{1-a} + 1,$$

which is equivalent to

$$b - 1 < (T^{-3}S)\hat{f}^{p_0}T^{-1}b < \frac{a}{1-a},$$

i.e., $\hat{f}^{(p_0+1)} = T^{-3}S$. Using the relation $(STS)T^2S = T^{-1}(STS)$, we can rewrite equation (4.7) as

$$(4.8) \quad (STS)T^2S\hat{f}^{n_0}T^mS = (T^{-3}S)\hat{f}^{p_0}T^{-1} = \hat{f}^{p_0+1}T^{-1}.$$

Let $\hat{f}^{(n_0+1)} = T^qS$. We have proved in (i) that $q = m$ or $m + 1$, hence $q \geq 3$. Let

$$b_0 = T^2S\hat{f}^{n_0}T^mSb \quad \text{and} \quad c_0 = (T^{-3}S)\hat{f}^{p_0}T^{-1}b.$$

Then $(STS)b_0 = c_0$ by (4.8). Using the relation $(STS)T = T^{-2}S(STS)$, we obtain $(STS)T^i = (T^{-2}S)^i(STS)$, and therefore,

$$(4.9) \quad (STS)T^i b_0 = (T^{-2}S)^i(STS)b_0 = (T^{-2}S)^i c_0.$$

Since, for $0 \leq i < q - 2$, $T^i b_0$ satisfies condition (a) of Lemma 4.3, we conclude that

$$b - 1 < (T^{-2}S)^i c_0 < \frac{a}{1-a}.$$

Therefore $\hat{f}^{(i)} = T^{-2}S$ and (4.9) for $i = q - 2$ give us the desired relation

$$(STS)\hat{f}^{n_0+1}T^mS = (T^{-2}S)\hat{f}^{p_0+q}T^{-1}$$

with $p = p_0 + q$. □

Now we complete the proof of part (I) of Theorem 4.2. In what follows we introduce the notations

$$I_\ell = \left(a, \frac{b}{b+1}\right) \quad \text{and} \quad I_u = \left(\frac{a}{1-a}, b\right)$$

and write \bar{I}_ℓ, \bar{I}_u for the corresponding closed intervals. If $\hat{f}^n T^m S b \in I_\ell$, then condition (b) of Lemma 4.3 is satisfied, and $(T^{-2}S)\hat{f}^p T^{-1}b \in I_u$. It follows that $\hat{f}^{(p+1)} = T^{-2}S$. Therefore (4.5) can be rewritten as $(STS)\hat{f}^n T^m S = \hat{f}^{p+1}T^{-1}$, which means that we reached the end of the cycle. More precisely,

(i) If $\hat{f}^n T^m S b \in (0, b/(b+1))$, then $\hat{f}^p T^{-1}b \in (0, b)$,

$$TS\hat{f}^n T^m S b = S\hat{f}^p T^{-1}b = c_b, \quad \text{and} \quad b - 1 < \hat{f}^j T^{-1}b < \frac{a}{1-a},$$

for $j < p$. In this case, $c_b < Sb$.

If $\hat{f}^n T^m S b \in (a, 0)$, then $\hat{f}^p T^{-1}b \in (a/(1-a), 0)$,

$$S\hat{f}^n T^m S b = T^{-1}S\hat{f}^p T^{-1}b = c_b, \quad \text{and} \quad b - 1 < \hat{f}^j T^{-1}b < \frac{a}{1-a},$$

for $j < p$. In this case, $c_b > Sa$. Since the cycle relation in both cases is equivalent to the identity (4.5), the cycle property is strong, and (i) is proved.

If $\hat{f}^n T^m S b = 0$, then $\hat{f}^n T^m S b = \hat{f}^p T^{-1}b = 0$ is the end of the cycle and $b - 1 < \hat{f}^j T^{-1}b < \frac{a}{1-a}$ for $j < p$. In this case the cycle ends “before” the identity (4.5) is complete, therefore the product over the cycle is not equal to identity, and the cycle is weak.

- (ii) If $\hat{f}^n T^m S b = a$, then following the argument in (i) and using relation (4.5) we obtain that the upper orbit of b hits $T^{-1} S \hat{f}^p T^{-1} b = S \hat{f}^n T^m S b = S a = -1/a$, while the lower orbit hits the value $a + 1$, hence b satisfies the cycle property if and only if a does.
- (iii) If $\hat{f}^n T^m S b = b/(b + 1)$, then following the argument in (i), we obtain that $(T^{-2} S) \hat{f}^p T^{-1} b = b$. However, one needs to apply one more T^{-1} to follow the definition of the map f , hence $\hat{f}^{(p+1)} = T^{-3} S$, not $T^{-2} S$, and the cycle will not close. One also observes that in this case the (a, b) -expansions of $S b$ and $T^{-1} b$ will be periodic, and therefore the cycle will never close.

We prove now part (II). If $\hat{f}^k T^m S b \notin \bar{I}_\ell$ for all $k \geq 0$, by the argument in the part (I) of the proof, on the lower orbit of b each $\hat{f}^{(k)} = T^q S$, where $q = m$ or $m + 1$, and on the upper orbit of b each $\hat{f}^{(p)} = T^{-r} S$, where $r = 2$ or 3 , and for all $p \geq 1$, $\hat{f}^p T^{-1} b \notin \bar{I}_u$. This means that for all images under the original map f on the lower orbit of b we have

$$f^k S b \in \left(-1 - \frac{1}{b}, a\right) \cup \left(\frac{b}{b+1}, a+1\right)$$

while for the images on the upper orbit of b

$$f^k T^{-1} b \in \left(b-1, \frac{a}{1-a}\right) \cup \left(b, 1 - \frac{1}{a}\right).$$

Since these ranges do not overlap, the cycle cannot close, and b has no cycle property. □

A similar result holds for the a -cycles. First, if $S a$ has the first digit 1, i.e., $b \leq S a < b + 1$, then one can easily write the a -cycle, similarly to (4.1). For the rest of the parameter region we have the following.

THEOREM 4.5. *Let $(a, b) \in \mathcal{P}$, $0 < b \leq -a < 1$ with $S a \geq b + 1$ and $m \geq 1$ such that $a \leq T^m S T a < a + 1$.*

(I) *Suppose that there exists $n \geq 0$ such that*

$$\hat{f}^k T^m S T a \in \left(\frac{b}{b+1}, a+1\right) \text{ for } k < n, \text{ and } \hat{f}^n T^m S T a \in \left[a, \frac{b}{b+1}\right].$$

- (i) *If $\hat{f}^n T^m S T a \in (a, b/(b + 1))$, then a has the cycle property; the cycle property is strong if and only if $\hat{f}^n T^m S T a \neq 0$.*
 - (ii) *If $\hat{f}^n T^m S T a = b/(b + 1)$, then a has the cycle property if and only if b has the cycle property.*
 - (iii) *If $\hat{f}^n T^m S T a = a$, then a does not have the cycle property, but the (a, b) -expansions of $S a$ and $T a$ are eventually periodic.*
- (II) *If $\hat{f}^k T^m S T a \in (b/(b + 1), a + 1)$ for all $k \geq 0$, then a does not have the cycle property.*

Proof. The proof follows the proof of Theorem 4.2 with minimal modifications. In particular, the relation (4.5) should be replaced by relation

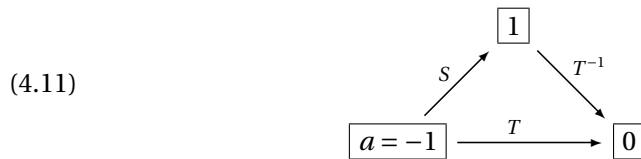
$$(4.10) \quad (STS) \hat{f}^n T^m S T = (T^{-2} S) \hat{f}^p.$$

For (ii), since $\hat{f}^n T^m S T a = \frac{b}{b+1}$, on the lower side we have $T S \hat{f}^n T^m S T a = S b$, and on the upper side, using (4.10), $(T^{-2} S) \hat{f}^p b = b$. As in the proof of Theorem 4.2, $\hat{f}^{p+1} = T^{-3} S$, so $(T^{-3} S) \hat{f}^p b = T^{-1} b$. Therefore a has (strong or weak) cycle property if and only if b does. \square

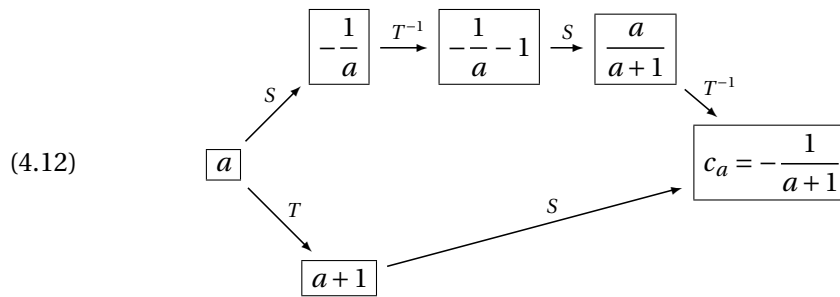
Let us now describe the situation when $a \leq -1$.

THEOREM 4.6. *Let $(a, b) \in \mathcal{P}$ with $0 < b \leq -a$ and $a \leq -1$. Then a and b satisfy the cycle property.*

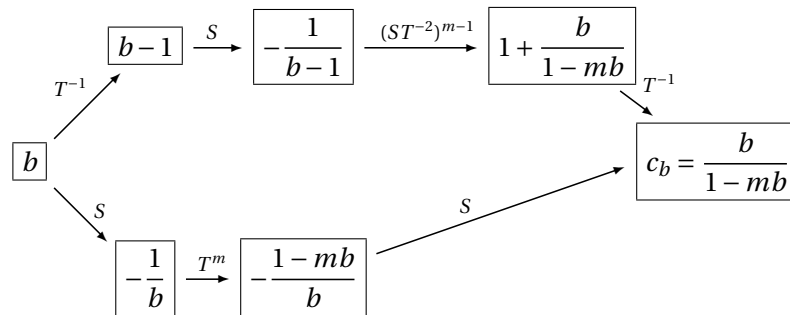
Proof. It is easy to see that $a = -1$ has the degenerate weak cycle:



while $a < -1$ satisfies the following strong cycle relation:

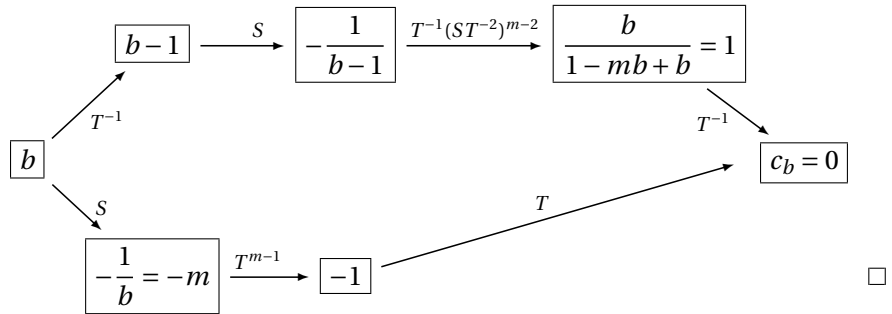


In order to study the orbits of b , let $m \geq 0$ such that $a \leq T^m S b < a + 1$. If $m = 0$, then $S b = a$ (since $S b \leq a$), and the cycle of b is identical to the one described by (4.11). If $m \geq 1$, then one can use relation (4.6) to construct the b -cycle. More precisely, if $a < T^m S b < a + 1$, then we have:



If $T^m S b = a$, then it happens again that the lower orbit of b hits a , and then $T a$, while the upper orbit hits $S a$. Following now the cycle of a described by (4.12), we conclude that b satisfies the strong cycle property.

If $T^m Sb = 0$, i.e., $b = 1/m$, then a minor modification of the above b -cycle gives us the following weak cycle relation:



The following corollaries are immediate from the proof of Theorems 4.2, 4.5, and 4.6.

COROLLARY 4.7. *If b has the cycle property, then the upper side of the b -cycle*

$$\{T^{-1}b, fT^{-1}b, \dots, f^{m_2-1}T^{-1}b\}$$

and the lower side of the b -cycle

$$\{Sb, fSb, \dots, f^{k_2-1}Sb\}$$

do not have repeating values.

COROLLARY 4.8. *If a has the cycle property, then the upper side of the a -cycle*

$$\{Sa, fSa, \dots, f^{m_1-1}Sa\}$$

and the lower side of the a -cycle

$$\{Ta, fTa, \dots, f^{k_1-1}Ta\}$$

do not have repeating values.

5. FINITENESS CONDITION IMPLIES FINITE RECTANGULAR STRUCTURE

In order to state the condition under which the natural extension map $F_{a,b}$ has an attractor with finite rectangular structure mentioned in the Introduction, we follow the split orbits of a and b , and define the *truncated orbits* \mathcal{L}_a and \mathcal{U}_a by

$$\mathcal{L}_a = \begin{cases} \mathcal{O}_\ell(Ta) & \text{if } a \text{ has no cycle property,} \\ \text{lower part of } a\text{-cycle} & \text{if } a \text{ has strong cycle property,} \\ \text{lower part of } a\text{-cycle} \cup \{0\} & \text{if } a \text{ has weak cycle property,} \end{cases}$$

$$\mathcal{U}_a = \begin{cases} \mathcal{O}_u(Sa) & \text{if } a \text{ has no cycle property,} \\ \text{upper part of } a\text{-cycle} & \text{if } a \text{ has strong cycle property,} \\ \text{lower part of } a\text{-cycle} \cup \{0\} & \text{if } a \text{ has weak cycle property.} \end{cases}$$

Similarly, \mathcal{L}_b and \mathcal{U}_b are defined by

$$\mathcal{L}_b = \begin{cases} \mathcal{O}_\ell(Sb) & \text{if } b \text{ has no cycle property,} \\ \text{lower part of } b\text{-cycle} & \text{if } b \text{ has strong cycle property,} \\ \text{lower part of } b\text{-cycle} \cup \{0\} & \text{if } b \text{ has weak cycle property,} \end{cases}$$

$$\mathcal{U}_b = \begin{cases} \mathcal{O}_u(T^{-1}b) & \text{if } b \text{ has no cycle property,} \\ \text{upper part of } b\text{-cycle} & \text{if } b \text{ has strong cycle property,} \\ \text{lower part of } b\text{-cycle} \cup \{0\} & \text{if } b \text{ has weak cycle property.} \end{cases}$$

We find it useful to introduce the map $\rho_{a,b}: \bar{\mathbb{R}} \rightarrow \{T, S, T^{-1}\}$

$$\rho_{a,b}(x) = \begin{cases} T & \text{if } x < a \\ S & \text{if } a \leq x < b \\ T^{-1} & \text{if } x \geq b \end{cases}$$

in order to write $f_{a,b}(x) = \rho_{a,b}(x)x$ and $F_{a,b}(x, y) = (\rho(y)x, \rho(y)y)$.

REMARK 5.1. It follows from the above definitions that $\rho(y) = S$ or T if $y \in \mathcal{L}_a \cup \mathcal{L}_b$, and $\rho(y) = S$ or T^{-1} if $y \in \mathcal{U}_a \cup \mathcal{U}_b$.

DEFINITION 5.2. We say that the map $f_{a,b}$ satisfies the *finiteness condition* if the sets of values in all four truncated orbits \mathcal{L}_a , \mathcal{L}_b , \mathcal{U}_a , and \mathcal{U}_b are finite.

PROPOSITION 5.3. *The following statements are equivalent.*

- (1) *The set \mathcal{L}_b is finite.*
- (2) *The set \mathcal{U}_b is finite.*
- (3) *Either b has the cycle property, or the upper and lower orbits of b are eventually periodic.*

Similar statements hold for the sets \mathcal{L}_a and \mathcal{U}_a .

Proof. The two properties follow from Theorem 4.2 and its proof. If b does not have the cycle property, but its lower orbit is eventually periodic, then one uses the argument in Lemma 4.4 to conclude that the upper orbit of b has to be eventually periodic. \square

REMARK 5.4. If b has the strong cycle property, then the set \mathcal{L}_b coincides with the lower side of the b -cycle and \mathcal{U}_b coincides with the upper side of the b -cycle. If b does not have the cycle property, but the lower and upper orbits of b are eventually periodic then \mathcal{L}_b and \mathcal{U}_b are identified with these orbits accordingly, until the first repeat.

THEOREM 5.5. *Let $(a, b) \in \mathcal{P}$, $a \neq 0$, $b \neq 0$, and assume that the map $f_{a,b}$ satisfies the finiteness condition. Then there exists a set $A_{a,b} \subsetneq \bar{\mathbb{R}}^2$ with the following properties.*

- (i) *The set $A_{a,b}$ consists of two connected components each having finite rectangular structure, i.e., bounded by nondecreasing step-functions with a finite number of steps.*

- (ii) $F_{a,b}: A_{a,b} \rightarrow A_{a,b}$ is a bijection except for some images of the boundary of $A_{a,b}$.

Proof. We construct a set $A_{a,b}$ whose upper connected component is bounded by a step-function with values in the set $\mathcal{U}_{a,b} = \mathcal{U}_a \cup \mathcal{U}_b$ that we refer to as *upper levels*, and whose lower connected component is bounded by a step-function with values in the set $\mathcal{L}_{a,b} = \mathcal{L}_a \cup \mathcal{L}_b$ that we refer to as *lower levels*. Notice that each level in \mathcal{U}_a and \mathcal{U}_b appears exactly once, but if the same level appears in both sets, we have to count it twice in $\mathcal{U}_{a,b}$. The same remark applies to the lower levels.

Now let $y_\ell \in \mathcal{L}_{a,b}$ be the closest y -level to Sb with $y_\ell \geq Sb$, and $y_u \in \mathcal{U}_{a,b}$ be the closest y -level to Sa with $y_u \leq Sa$. Since each level in \mathcal{U}_a and in \mathcal{L}_b appears only once, if $y_u = Sa$, y_u can only belong to \mathcal{U}_b , and if $y_\ell = Sb$, y_ℓ can only belong to \mathcal{L}_a .

We consider the rays $[-\infty, x_b] \times \{b\}$ and $[x_a, \infty] \times \{a\}$, where x_a and x_b are unknown, and “transport” them (using the special form of the natural extension map $F_{a,b}$) along the sets \mathcal{L}_b , \mathcal{U}_b , \mathcal{L}_a , and \mathcal{U}_a , respectively, until we reach the levels y_u and y_ℓ (see Figure 4). Now we set-up a system of two fractional linear equations by equating the right end of the segment at the level Sb with the left end of the segment at the level y_ℓ , and, similarly, the left end of the segment at the level Sa and the right end of the level y_u .

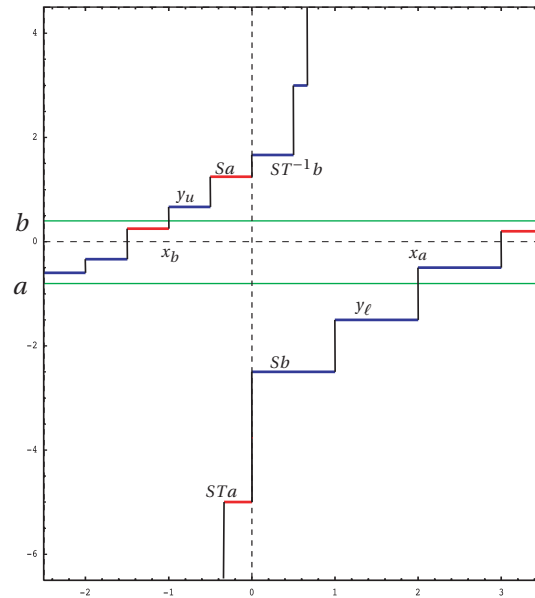


FIGURE 4. Construction of the domain $A_{a,b}$

In what follows, we present the proof assuming that $0 < b \leq -a < 1$. The situation $a \leq -1$ is less complex due to the explicit cycle expressions described in Theorem 4.6 and will be discussed at the end of this section.

LEMMA 5.6. *The system of two equations at the consecutive levels y_u and Sa , and y_ℓ and Sb , has a unique solution with $x_a \geq 1$ and $x_b \leq -1$.*

Proof. Let m_a, m_b be positive integers such that $a \leq T^{m_a} S T a < a + 1$ and $a \leq T^{m_b} S b < a + 1$. For the general argument we assume that $m_a, m_b \geq 3$, the cases m_a or $m_b \in \{1, 2\}$ being considered separately. The level y_u may belong to \mathcal{U}_a or \mathcal{U}_b , and the level y_ℓ may belong to \mathcal{L}_a or \mathcal{L}_b , therefore we need to consider the following four possibilities.

Case 1: $y_u \in \mathcal{U}_a, y_\ell \in \mathcal{L}_a$. Then we have

$$Sx_a = T^{-1} S \hat{f}_-^{n_1}(\infty) \quad \text{and} \quad Sx_b = T S \hat{f}_+^{n_2} T x_a,$$

where $\hat{f}_-^{n_1}$ is a product of factors $T^{-i} S$ (that appear on the upper orbit of a) with $i = 2$ or 3 , and $\hat{f}_+^{n_2}$ is a product of factors $T^i S$ (that appear on the lower orbit of a) with $i = m$ or $m + 1$. Using (4.10), we rewrite the first equation as

$$x_a = S T^{-1} S \hat{f}_-^{n_1}(\infty) = S T^{-1} S S T^2 S T S \hat{f}_+^{k_1} T^m S T(\infty) = T^{-1} \hat{f}_+^{k_1} T^m S T(\infty).$$

Since $\hat{f}_+^{k_1}$ is a product of factors $T^i S$ with $i = m$ or $m + 1, m \geq 3$, we conclude that $T x_a$ has a finite formal continued fraction expansion starting with $m' \geq 3$, i.e., $T x_a > 2$, and $x_a > 1$. Furthermore, it follows from the second equation that $x_b = T S \hat{f}_+^{n_2} T x_a$. Hence $\hat{f}_+^{n_2} T x_a$ has a finite formal continued fraction expansion starting with $m' \geq 3$, i.e., $\hat{f}_+^{n_2} T x_a > 2$, and $x_b < -2$.

Case 2: $y_u \in \mathcal{U}_a, y_\ell \in \mathcal{L}_b$. Then

$$Sx_a = T^{-1} S \hat{f}_-^{n_1}(\infty) \quad \text{and} \quad Sx_b = T S \hat{f}_+^{n_2}(-\infty).$$

Like in Case 1 we see that $x_a > 1$, and $x_b = T S \hat{f}_+^{n_2}(-\infty) < -2$, since $\hat{f}_+^{n_2}(-\infty)$ has a formal continued fraction expansion starting with $m' \geq 3$, and therefore is > 2 .

Case 3: $y_u \in \mathcal{U}_b, y_\ell \in \mathcal{L}_a$. Then

$$Sx_a = T^{-1} S \hat{f}_-^{n_1} T^{-1} x_b \quad \text{and} \quad Sx_b = T S \hat{f}_+^{n_2} T x_a.$$

Using (4.5),

$$x_a = S T^{-1} S \hat{f}_-^{n_1} T^{-1} x_b = S T^{-1} S S T^2 S T S \hat{f}_+^{k_2} T^m S x_b,$$

and using the second equation and simplifying, we obtain

$$T x_a = \hat{f}_+^{k_2} T^m S S T S \hat{f}_+^{n_2}(T x_a) = \hat{f}_+^{k_2} T^{m+1} S \hat{f}_+^{n_2}(T x_a).$$

Since all its factors are of the form $T^i S$ with $i \geq 3$, the matrix $\hat{f}_+^{k_2} T^{m+1} S \hat{f}_+^{n_2}$ is hyperbolic and its attracting fixed point $T x_a$ has periodic formal continued fraction expansion starting with $m' \geq 3$ (see Proposition 1.3 of [11]), hence $x_a > 1$. Finally, as in Case 1, $x_b = T S \hat{f}_+^{n_2} T x_a < -2$, since $\hat{f}_+^{n_2} T x_a$ has formal continued fraction expansion starting with $m' \geq 3$, hence > 2 .

Case 4: $y_u \in \mathcal{U}_b, y_\ell \in \mathcal{L}_b$. Then

$$Sx_a = T^{-1} S \hat{f}_-^{n_1} T^{-1} x_b \quad \text{and} \quad Sx_b = T S \hat{f}_+^{n_2} S(-\infty).$$

From the second equation we obtain $x_b = STS\hat{f}_+^{n_2}S(-\infty) < -2$ since $\hat{f}_+^{n_2}S(-\infty)$ has formal continued fraction expansion starting with $m' \geq 3$, hence > 2 . Finally,

$$x_a = ST^{-1}S\hat{f}_-^{n_1}T^{-1}x_b = T^{-1}\hat{f}^{k_2}T^{m+1}S\hat{f}_+^{n_2}S(-\infty).$$

Hence $Tx_a = \hat{f}^{k_2}T^{m+1}S\hat{f}_+^{n_2}S(-\infty) > 2$ since it has formal continued fraction expansion starting with $m' \geq 3$, therefore $x_a > 1$.

Now we analyze the particular situations when m_a or $m_b \in \{1, 2\}$, using the explicit cycle descriptions that exist for these situations as described by Theorems 4.2 and 4.5.

- (i) If $m_a = m_b = 1$, then relation (4.2) for the b -cycle and a similar one for the a -cycle shows that $y_\ell = -\frac{1}{b} + 1$ and $y_u = -\frac{1}{a} - 1$, therefore $x_a = 1$ and $x_b = -1$.
- (ii) If $m_a = 1$, $m_b = 2$, following the explicit cycles given by (4.3) we obtain $y_\ell = -1/b + 1$, and $y_u = -1/(b-1) - 1$, therefore $x_a = 2$, $x_b = -1$.
- (iii) If $m_a = 1$, $m_b \geq 3$, using the cycle structure in Theorem 4.2 we obtain $y_\ell = 1/b + 1$ and $y_u = T^{-1}(ST^{-2})^{m_b-2}ST^{-1}b$, therefore, $x_a = m_b$, and $x_b = -1$.
- (iv) If $m_a = 2$, $m_b = 2$, using the cycle structure in Theorems 4.2 and 4.5 we obtain $y_\ell = -\frac{1}{a+1} + 1$ and $y_u = -\frac{1}{b-1} - 1$, and a calculation in this particular case, like in Lemma 5.6, Case 3 implies that $x_a > 1$ and $x_b < -1$.
- (v) If $m_a = 2$, $m_b > 2$, an analysis of the four cases above for this particular situation (with an explicit cycle relation for a) yields $x_a \geq 1$ and $x_b \leq -1$. Indeed, in Case 1, we have $y_u = -1/a - 1$, hence $x_a = 1$ and $x_b = -2$. In Case 2, we get $x_a = 1$ and $x_b < -2$. Cases 3 and 4 are treated similarly. \square

Now, since x_a and x_b are uniquely determined, by “transporting” the rays $[-\infty, x_b] \times \{b\}$ and $[x_a, \infty] \times \{a\}$ along the sets \mathcal{L}_b , \mathcal{U}_b , \mathcal{L}_a , and \mathcal{U}_a , we obtain the x -coordinates of the right and left end of the segments on each level.

DEFINITION 5.7. We say that two consecutive levels $y_1 \leq y_2$ of $\mathcal{L}_{a,b}$, respectively, $\mathcal{U}_{a,b}$, are *connected by a vertical segment* (we will refer to this as *connected*) if the x -coordinate of the right end-point of the horizontal segment on the level y_1 is equal to the the x -coordinate of the left end-point of the horizontal segment on the level y_2 .

In order to prove part (i) of Theorem 5.5, we show that all levels of $\mathcal{L}_{a,b}$ and all levels of $\mathcal{U}_{a,b}$ are connected. We first look at the levels in $\mathcal{L}_{a,b}$. By Lemma 5.6 the levels y_u and Sa , and the levels Sb and y_ℓ are connected.

LEMMA 5.8. *The levels $Sb \in \mathcal{L}_b$ and $STa \in \mathcal{L}_a$ are two consecutive levels of $\mathcal{L}_{a,b}$ connected by a vertical segment at $x = 0$. The levels $Sa \in \mathcal{U}_a$ and $ST^{-1}b \in \mathcal{U}_b$ are two consecutive levels of $\mathcal{U}_{a,b}$ connected by a vertical segment at $x = 0$.*

Proof. Suppose there is $y \in \mathcal{L}_{a,b}$ such that $STa \leq y \leq Sb$. Then $y \in \mathcal{L}_a$ or \mathcal{L}_b . In either case, since by Lemmas 4.8 and 4.7 the truncated orbits $\mathcal{L}_a, \mathcal{L}_b$ do not have repeated values, neither $STa = y$ nor $y = Sb$ is possible. Thus the only case we need to consider is $STa < y < Sb$. Then, either $y = Sy'$ for some $y' \in \mathcal{L}_{a,b}$ ($0 < y' \leq a + 1$) or $y = Ty''$ for some $y'' \in \mathcal{L}_{a,b}$. These would imply

that either $y' > Ta$, which is impossible, or $Ty'' < Sb$, i.e., $y'' < T^{-1}Sb$, which is also impossible (if $y'' < T^{-1}Sb$ then $y = Ty''$ must be the end of the a -cycle, by Theorem 4.5). The x -coordinate of the right end-point of the segment at the level STa and of the left end-point of the segment at the level Sb is equal to 0. The second part of the proof is similar. \square

The following proposition will be used later in the proof.

PROPOSITION 5.9. *Suppose that the set $\mathcal{L}_{a,b}$ is finite and $y \in \mathcal{L}_{a,b}$ with $y > STa$.*

- (1) *If $y \in \mathcal{L}_a$, then there exists $n_0 > 0$ such that $\rho(f^n y) = \rho(f^n STa)$ for all $0 < n < n_0$ and $\rho(f^{n_0} y) \neq \rho(f^{n_0} STa)$, or $f^{n_0} y = 0$.*
- (2) *If $y \in \mathcal{L}_b$, then $y > Sb$ and there exists $n_0 > 0$ such that $\rho(f^n y) = \rho(f^n Sb)$ for all $n < n_0$ and $\rho(f^{n_0} y) \neq \rho(f^{n_0} Sb)$, or $f^{n_0} y = 0$.*

Proof. Suppose that $y \in \mathcal{L}_a$ and a satisfies the cycle property. It follows that such an n_0 exists or $f^{n_0} y$ is the end of the a -cycle. We will show that the latter is possible only if $f^{n_0} y = 0$, i.e., it is the end of a weak cycle. Suppose $f^{n_0} y$ is the end of the a -cycle. Then if

$$\rho(f^{n_0-1} y) = \rho(f^{n_0-1} STa) = S,$$

we must have $f^{n_0-1} y < 0$ since otherwise the cycle would not stop at S , but $f^{n_0-1}(STa) > 0$ since for STa we have not reached the end of the cycle. This contradicts the monotonicity of f^{n_0-1} and the original assumption $y > STa$, thus is impossible. The other possibility is

$$\rho(f^{n_0-1} y) = \rho(f^{n_0-1} STa) = T.$$

But this either implies that $f^{n_0-1} y < T^{-1}Sb$, and by monotonicity of f^{n_0-1} , $f^{n_0-1}(STa) < f^{n_0-1} y < T^{-1}Sb$, which implies that we have reached the end of the cycle of STa as well, a contradiction, or, $f^{n_0} y = 0$, i.e., it is the end of a weak cycle.

Now suppose $y \in \mathcal{L}_b$. Then by Lemma 5.8 $y \geq Sb$, but since each level in \mathcal{L}_b appears only once, we must have but $y > Sb$. Now the argument that $f^{n_0} y$ cannot be the end of the b -cycle is exactly the same as for the a -cycle.

In the periodic case, assume that no such n_0 exists. Then, in case (1), the (a, b) -expansions of STa and y , which is the lower part of the former, are the same, i.e., (a, b) -expansions of STa is invariant under a left shift. In case (2), we have seen already that we must have $y > Sb$. Then the (a, b) -expansions of Sb and y , which is the lower part of the former, are the same, i.e., (a, b) -expansions of Sb is invariant under a left shift. The proof that this is impossible is based on the following simple observation: if $\sigma = (a_1, a_2, \dots, a_k, \overline{a_{k+1}}, \overline{a_{k+2}}, \dots, \overline{a_{k+n}})$ is an eventually periodic symbolic sequence with the minimal period n and invariant under a left shift by m , then σ is purely periodic and m is a multiple of n . By the uniqueness property of (a, b) -expansions, this would imply that $y = STa$ or $y = Sb$, a contradiction. \square

Let $y_b^-, y_b^+ \in \mathcal{U}_{a,b}$ be two consecutive levels with $y_b^- \leq b < y_b^+$, and $y_a^-, y_a^+ \in \mathcal{L}_{a,b}$ be two consecutive levels with $y_a^- < a \leq y_a^+$.

LEMMA 5.10. *There is always one level connected with level $a + 1$, and the levels y_a^- and y_a^+ are connected by the vertical segment at x_a .*

Proof. By Lemmas 5.6 and 5.8, we know that three consecutive levels $STa \leq Sb \leq y_\ell$ are connected. Moreover, their images remain connected under the same transformations in $SL(2, \mathbb{Z})$. Since each level in \mathcal{U}_a and in \mathcal{L}_a appears only once, at least one of the two inequalities must be strict, i.e., if $STa = Sb$, then $STa = Sb < y_\ell$, and if $Sb = y_\ell$, then $STa < Sb = y_\ell$.

First we prove that $y_\ell < TSb$. Suppose $y_\ell \geq TSb$. Its preimage must be $y'_\ell = T^{-1}y_\ell$ since for any y , $0 < y < Ta$, $Sy < STa \leq Sb < TSb$, and we would have $Sb \leq y'_\ell < y_\ell$ that contradicts the assumption that y_ℓ is the next level above Sb . Therefore, if the first digit in the (a, b) -expansion of Sb is $-m$, then the first digit of y_ℓ is $-(m-1)$ or $-m$. In the first case, the three levels

$$T^{m-1}Sb < a \leq T^{m-1}y_\ell$$

are connected and satisfy $T^{m-1}Sb = y_a^-$, $T^{m-1}y_\ell = y_a^+$. Therefore, the levels T^mSb and $a + 1$ are connected.

For the second case, we know that $Sb \leq y_\ell$ and

$$a \leq T^mSb \leq T^m y_\ell < a + 1.$$

If $Sb = y_\ell$, then $y_\ell \in \mathcal{L}_a$, and $STa < y_\ell$. If $Sb < y_\ell$, then $y_\ell \in \mathcal{L}_b$, or $y_\ell \in \mathcal{L}_a$ and $STa < y_\ell$.

Let us assume that y_ℓ belongs to \mathcal{L}_a . Since $STa < y_\ell$, by Proposition 5.9, there are two possibilities:

- (1) $f^{n_0}y_\ell$ is the end of a weak cycle.
- (2) There exists n_0 such that $\rho(f^n y_\ell) = \rho(f^n STa)$ for all $n < n_0$, and $\rho(f^{n_0} y_\ell) \neq \rho(f^{n_0} STa)$.

In the first case, we have $f^{n_0}STa = y_a^-$ and $f^{n_0}Sb = y_a^+$, or $f^{n_0}Sb = y_a^-$ and $f^{n_0}y_\ell = y_a^+$. Therefore, either $f^{n_0+1}STa$ or $f^{n_0+1}Sb$ is connected with level $a + 1$. In the second case, we notice that

$$\rho(f^{n_0-1}y_\ell) = \rho(f^{n_0-1}STa) = T$$

otherwise, $\rho(f^{n_0-1}y_\ell) = \rho(f^{n_0-1}STa) = S$ would imply

$$\rho(f^{n_0}y_\ell) = \rho(f^{n_0}STa) = T$$

in contradiction with the choice of n_0 . Further, there are two possibilities:

- (i) $\rho(f^{n_0}STa) = S$, $\rho(f^{n_0}y_\ell) = T$;
- (ii) $\rho(f^{n_0}STa) = T$, $\rho(f^{n_0}y_\ell) = S$.

In case (i) we obtain

$$f^{n_0}y_\ell < a \leq f^{n_0}STa$$

which contradicts the monotonicity of f and the original assumption $y_\ell > STa$. Thus the only possibility is

$$f^{n_0}y_\ell \geq a > f^{n_0}STa.$$

By using the monotonicity of f^{n_0} we have

$$f^{n_0} y_\ell > f^{n_0} S b > f^{n_0} S T a$$

and conclude that $f^{n_0} S T a = y_a^-$ and $f^{n_0} S b = y_a^+$, or $f^{n_0} S b = y_a^-$ and $f^{n_0} y_\ell = y_a^+$. Therefore, either $f^{n_0+1} S T a$ or $f^{n_0+1} S b$ is connected with level $a + 1$. The case when y_ℓ belongs to \mathcal{L}_b is similar, and in this case $f^{n_0} S b = y_a^-$, $f^{n_0} y_\ell = y_a^+$, and $f^{n_0+1} S b$ is connected with $a + 1$. By construction, in both cases the common x -coordinate of the end-points is equal to x_a . \square

After an application of S the level connected with $a + 1$ will be connected with $S T a$, and now, instead of 3 connected levels $S T a \leq S b \leq y_\ell$ (with at least one strict inequality) we have at least 4 connected levels $y' \leq S T a \leq S b \leq y_\ell$ (with no more than two equalities in a row).

The process continues with a growing number of connected levels, the highest being $a + 1$. Since on each step we cannot have more than two equalities in a row, the number of distinct levels in this sequence will also increase. Therefore, we obtain a sequence of connected levels

$$(5.1) \quad a + 1 \geq y_1 \geq \dots \geq y_s > \frac{b}{b+1} \geq y_{s+1}.$$

It is evident from the construction that there are no unaccounted levels $y \in \mathcal{L}_{a,b}$, $a + 1 \geq y \geq y_{s+1}$.

Now we prove a similar result for $\mathcal{U}_{a,b}$.

LEMMA 5.11. *There is always one level connected with level $b - 1$, and the levels y_b^- and y_b^+ are connected by a vertical segment at x_b .*

Proof. By Lemmas 5.6 and 5.8, we know that the three consecutive levels $y_u \leq S a \leq S T^{-1} b$ are connected. It is easy to see that the first digit in (a, b) -expansion of $S T^{-1} b$ is 2, and the first digit in (a, b) -expansion of $S a$ is either 1 or 2. Therefore, the first digit in (a, b) -expansion of y_u is either 1 or 2. In the first case, either

$$T^{-1} S a < b \leq T^{-1} S T^{-1} b \quad \text{or} \quad T^{-1} y_u < b \leq T^{-1} S a$$

are the connected levels. Therefore either $T^{-1} S a = y_b^-$ and $T^{-1} S T^{-1} b = y_b^+$, or $T^{-1} y_u = y_b^-$ and $T^{-1} S a = y_b^+$ are connected. So either $T^{-2} S T^{-1} b$ or $T^{-2} S a$ is connected with level $b - 1$.

In the second case, we know that $y_u \leq S a$ and

$$b - 1 \leq T^{-2} y_u \leq T^{-2} S a < b.$$

If $y_u = S a$, y_u must belong to \mathcal{U}_b , in which case $y_u < S T^{-1} b$. If $y_u < S a$, then $y_u \in \mathcal{U}_a$, or $y_u \in \mathcal{U}_b$ and $y_u < S T^{-1} b$.

Let us assume that y_u belongs to \mathcal{U}_b . Since $y_u < S T^{-1} b$, by Proposition 5.9 there are two possibilities:

- (1) $f^{n_0} y_u$ is the end of a weak cycle;
- (2) there exists n_0 such that we have $\rho(f^n y_u) = \rho(f^n S T^{-1} b)$ for all $n < n_0$, and $\rho(f^{n_0} y_u) \neq \rho(f^{n_0} S T^{-1} b)$.

In the first case, either $f^{n_0}ST^{-1}b = y_b^+$ and $f^{n_0}Sa = y_b^-$, or $f^{n_0}Sa = y_b^+$ and $f^{n_0}y_u = y_b^-$, so either $f^{n_0+1}ST^{-1}b$ or $f^{n_0+1}Sa$ is connected with level $b - 1$. In the second case, we first notice that

$$\rho(f^{n_0-1}y_u) = \rho(f^{n_0-1}ST^{-1}b) = T^{-1}$$

since if we had $\rho(f^{n_0-1}y_u) = \rho(f^{n_0-1}ST^{-1}b) = S$, then we would have

$$\rho(f^{n_0}y_u) = \rho(f^{n_0}ST^{-1}b) = T^{-1}$$

in contradiction with the choice of n_0 . Further, there are two possibilities:

- (i) $\rho(f^{n_0}ST^{-1}b) = S, \rho(f^{n_0}y_u) = T^{-1}$;
- (ii) $\rho(f^{n_0}ST^{-1}b) = T^{-1}, \rho(f^{n_0}y_u) = S$.

In the first case we obtain

$$f^{n_0}y_u > b > f^{n_0}ST^{-1}b$$

which contradicts the monotonicity of f^{n_0} and the original assumption $y_u < ST^{-1}b$. Thus the only possibility is

$$f^{n_0}y_u < b < f^{n_0}ST^{-1}b.$$

By monotonicity of f^{n_0} we have

$$f^{n_0}y_u < f^{n_0}Sa < f^{n_0}ST^{-1}b.$$

Therefore either $f^{n_0}y_u = y_b^-$ and $f^{n_0}Sa = y_b^+$, or $f^{n_0}Sa = y_b^-$ and $f^{n_0}ST^{-1}b = y_b^+$ are connected. So either $T^{-1}f^{n_0}ST^{-1}b$ or $T^{-1}f^{n_0}Sa$ is connected with level $b - 1$. The case when y_u belongs to the a -cycle is similar, and in this case $f^{n_0}y_u = y_b^-$ and $f^{n_0}Sa = y_b^+$ and $T^{-1}f^{n_0}Sa$ is connected with level $b - 1$. By construction, in both cases the common x -coordinate of the end-points of the segments at the levels y_b^- and y_b^+ is x_b . □

After an application of S the level connected with $b - 1$ will be connected with $ST^{-1}b$, and now, instead of 3 connected levels $y_u \leq Sa \leq ST^{-1}b$ we have at least 4 connected levels $y_u \leq Sa \leq ST^{-1}b \leq y''$.

The process continues with a growing number of connected levels, the lowest being $b - 1$. Also the number of distinct levels will increase, and we obtain a sequence of connected levels

$$b - 1 \leq \bar{y}_1 \leq \dots \leq \bar{y}_t < \frac{a}{1 - a} \leq \bar{y}_{t+1}.$$

It is evident from the construction that there are no unaccounted levels $y \in \mathcal{U}_{a,b}, b - 1 \leq y \leq \bar{y}_{t+1}$.

Now we complete the proof that all levels of $\mathcal{L}_{a,b}$ are connected. For that it is sufficient to find a sequence of connected levels with the distance between the highest and the lowest level ≥ 1 and the lowest level $\geq T^{-1}Sb$. This is because the set of levels in $y \in \mathcal{L}_{a,b}$ satisfying $T^{-1}Sb \leq y \leq a + 1$ is periodic with period 1, and each $y \in \mathcal{L}_{a,b}$ uniquely determines a horizontal segment on level y , as was explained just before Lemma 5.8.

If $y_{s+1} \leq a$, then all levels in $\mathcal{L}_{a,b}$ are connected. Suppose now that $y_{s+1} > a$. If $y_{s+1} = y_a^+$, then, since y_a^+ is already connected with y_a^- , all levels of $\mathcal{L}_{a,b}$ are connected. Now assume that $y_{s+1} > y_a^+$. Then either

$$y_{s+1} = \frac{b}{b+1} \quad \text{or} \quad y_{s+1} < \frac{b}{b+1}.$$

In the first case either $TSy_{s+1} = y_\ell = Sb$ (this can only happen if $y_{s+1} \in \mathcal{L}_a$), or $TSy_s > Sb$ is the next level above Sb , and hence $TSy_s = y_\ell$. In either case $Sy_{s+1} \leq Sy_s \leq \dots \leq STa \leq Sb = TSy_{s+1}$ are the connected levels with the distance between the lowest and the highest equal to 1, thus we conclude that all levels of $\mathcal{L}_{a,b}$ are connected.

In the second case, the two levels $y_a^+ < y_{s+1}$ will produce the ends of the cycles (one of them can be weak if one of y_a^+ or y_{s+1} is equal to 0). By the cycle property (Lemma 4.4(ii)), there exists a level $z \in \mathcal{U}_{a,b}$, $\frac{a}{1-a} < z < b$ such that $z = (STS)y_{s+1}$. We claim that $z = y_b^-$. Suppose not, and $z < y_b^-$. Then y_b^- gives rise to the second cycle, and again by the cycle property, there exists $y \in \mathcal{L}_{a,b}$, $y < \frac{b}{b+1}$, such that $y_b^- = STSy$. Since $STS(z) = -\frac{z}{z-1}$ is monotone increasing for $z < 1$, we conclude that $y > y_{s+1}$ in contradiction with (5.1). Thus $y_b^- = (STS)y_{s+1}$. Then $TSy_{s+1} = Sy_b^-$ which implies that the right end of the segment at the level Sy_b^- , which is equal to the right end of the segment at the level Sb , is equal to the right end of the segment at the level TSy_{s+1} (notice that this level may belong to $\mathcal{L}_{a,b}$, $\mathcal{U}_{a,b}$ or be at infinity if $y_{s+1} = 0$). Since y_s and y_{s+1} were connected, the left end of the segment at the level TSy_s is equal to the right end of the segment at the level TSy_{s+1} even though they may belong to the boundaries of different connected components. Since $TSy_s \in \mathcal{L}_{a,b}$, we conclude that the segment at the level TSy_s is adjacent to the segment at the level Sb , i.e., $TSy_s = y_\ell$. Thus $Sy_s \leq Sy_{s-1} \leq \dots \leq STa \leq Sb \leq TSy_s$ are the connected levels with the distance between the lowest and the highest equal to 1, and therefore all levels in $\mathcal{L}_{a,b}$ are also connected. The proof for $\mathcal{U}_{a,b}$ follows exactly the same lines.

Now we continue with the proof of part (ii) of Theorem 5.5. In order to prove the bijectivity of the map F on $A_{a,b}$ we write it as a union of the upper and lower connected components, $A_{a,b} = A_{a,b}^u \cup A_{a,b}^\ell$, and subdivide each component into 3 pieces: $A_{a,b}^u = \bigcup_{i=1}^3 U_i$, and $A_{a,b}^\ell = \bigcup_{i=1}^3 L_i$, where

$$\begin{aligned} U_1 &= \{(x, y) \in A_{a,b}^u \mid y \geq b\} & L_1 &= \{(x, y) \in A_{a,b}^\ell \mid y \leq a\} \\ U_2 &= \{(x, y) \in A_{a,b}^u \mid b-1 \leq y \leq 0\} & L_2 &= \{(x, y) \in A_{a,b}^\ell \mid 0 \leq y \leq a+1\} \\ U_3 &= \{(x, y) \in A_{a,b}^u \mid 0 \leq y \leq b\} & L_3 &= \{(x, y) \in A_{a,b}^\ell \mid a \leq y \leq 0\}. \end{aligned}$$

Let

$$\begin{aligned} U'_1 &= T^{-1}(U_1), & U'_2 &= S(U_2), & U'_3 &= S(U_3), \\ L'_1 &= T(L_1), & L'_2 &= S(L_2), & L'_3 &= S(L_3) \end{aligned}$$

be their images under the transformation F (see Figure 5).

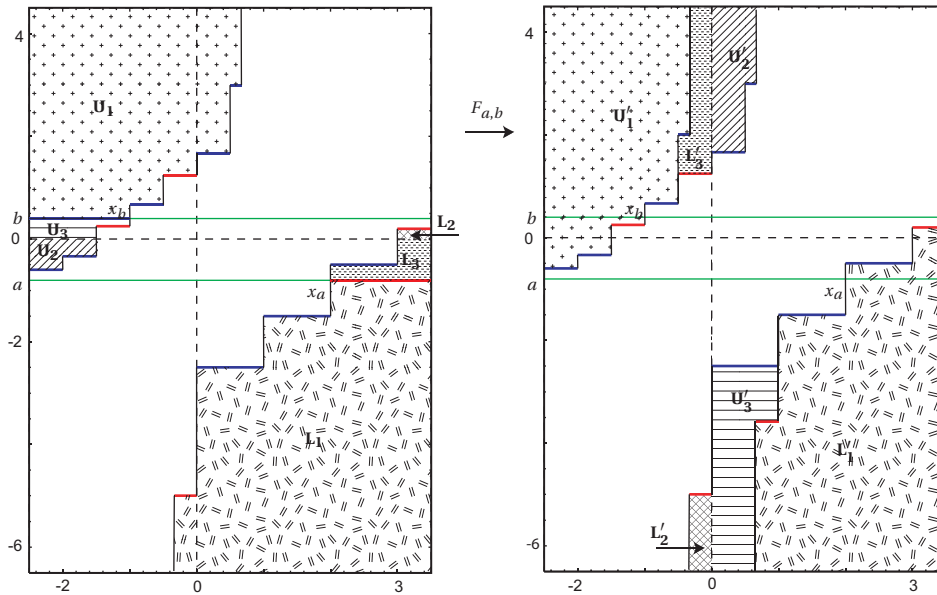


FIGURE 5. Bijectivity of the map $F_{a,b}$

Since the set $A_{a,b}$ is bounded by step-functions with finitely many steps, each of the pieces U_i, L_i have the same property, and so do their images under F . By the construction of the set $A_{a,b}$ we know that the levels corresponding to the ends of the cycles c_a and c_b , if the cycles are strong, do not appear as horizontal boundary levels; the corresponding horizontal segments, let us call them the *locking segments* lie in the interior of the set $A_{a,b}$. Furthermore, the images of all levels except for the levels next to the ends of the cycles, $f^{k_1-1}Ta$, $f^{m_1-1}Sa$, $f^{m_2-1}Sb$, and $f^{k_2-1}T^{-1}b$, also belong to $\mathcal{U}_{a,b} \cup \mathcal{L}_{a,b}$. The exceptional levels are exactly those between 0 and b and above $T Sa$ in $\mathcal{U}_{a,b}$, and between a and 0 and below $T^{-1} Sb$ in $\mathcal{L}_{a,b}$. The images of the horizontal segments belonging to these levels are the locking segments. Notice that the exceptional levels between 0 and b and between a and 0 constitute the horizontal boundary of the regions U_3 and L_3 .

Transporting the rays $[-\infty, x_b]$ and $[x_a, \infty]$ (with x_a and x_b uniquely determined by Lemma 5.6), along the corresponding cycles, and using the strong cycle property, we see that the “locking segment” in the horizontal boundary of U'_1 coincides with the locking segment of the horizontal boundary of L'_3 , and the locking segment in the horizontal boundary of L'_1 coincides with the locking segment of the horizontal boundary of U'_3 . It can happen that both “locking segments” belong to $A^u_{a,b}$ or $A^\ell_{a,b}$. If only one of the numbers a or b has the strong cycle property, then there will be only one locking segment.

If the cycle property is weak or the (a, b) -continued fraction expansion of one or both a and b is periodic, then all levels of $\mathcal{L}_a, \mathcal{L}_b, \mathcal{U}_a$, and \mathcal{U}_b will belong to the boundary of $A_{a,b}$, and there will be no locking segments. In these

cases $L_3 = [x_1, \infty) \times [a, 0]$, and $L'_3 = [-1/x_1, 0] \times [-1/a, \infty)$, where $x_1 = x_a$. Let x_2 be the x -coordinate of the right vertical boundary segment of U_2 . Then the x -coordinate of the right vertical boundary segment of U_1 is $-1/x_2$. Let us denote the highest level in $\mathcal{U}_{a,b}$ by y_2 . Since $y_2 \leq -1/a + 1$, $y_2 - 1 \leq -1/a$ is the next level after $-1/a$ in $\mathcal{U}_{a,b}$. This is since if we had $y \in \mathcal{U}_{a,b}$ such that $y_2 - 1 < y < -1/a$, its preimage $y' = Ty$ would satisfy $y_2 < y' < -1/a + 1$, a contradiction. By construction of the region $A_{a,b}$ the segments at the levels $y_2 - 1$ and $-1/a$ are connected, therefore $Sx_1 = T^{-1}Sx_2$. This calculation shows that L'_3 and U'_1 do not overlap and fit together by this vertical ray.

Thus in all cases the images U'_i, L'_i do not overlap, and $A_{a,b} = (\cup_{i=1}^3 U'_i) \cup (\cup_{i=1}^3 L'_i)$. This proves the bijectivity of the map F on $A_{a,b}$ except for some images of its boundary. This completes the proof in the case $0 < b \leq -a < 1$.

Now we return to the case $a \leq -1$ that was dropped from consideration before Lemma 5.6. The explicit cycle relations for this case have been described in Theorem 4.6. Notice that all lower levels are connected, and $T^m Sb$ is connected with $a + 1$. Therefore $y_\ell = TSb$, and this implies that $x_a = m$. The upper levels in the positive part are

$$ST^{-1}b < ST^{-2}ST^{-1}b < \dots < (ST^{-2})^{m-1}ST^{-1}b < a/(a+1)$$

and $y_u = T^{-1}(ST^{-2})^{m-2}ST^{-1}b$. Lemma 5.6 in this case holds with $x_a = m$ and $x_b = -1$ since the equation for adjacency of the levels y_u and Sa is

$$T^{-1}(ST^{-2})^{m-2}ST^{-1}x_b = ST^{m-1}Sx_b = -1/m,$$

which implies $x_b = -1$. Lemma 5.10 also holds with $y_a^- = ST^{m-1}b$ and $y_a^+ = ST^m b$. Lemma 5.11 holds with $y_b^- = T^{-1}Sa$, $y_b^+ = T^{-1}ST^{-1}b$ and all upper level will be connected by an argument similar to one described above. To prove the bijectivity of F on $A_{a,b}$ one proceeds the same way as above, the only modification being that level L_2 does not exist, and $L_3 = \{(x, y) \in A_{a,b}^\ell, a \leq y \leq a + 1\}$. \square

The following corollary is evident from the proof of part (ii) of Theorem 5.5.

COROLLARY 5.12. *If both a and b have the strong cycle property, then for any boundary component h of $A_{a,b}$ (vertical or horizontal) there exists $N > 0$ such that $F^N(h)$ is in the interior of $A_{a,b}$.*

6. FINITE RECTANGULAR STRUCTURE OF THE ATTRACTING SET

Recall that the attracting set $D_{a,b}$ was defined by (3.1): starting with the trapping region $\Theta_{a,b}$ described in Theorem 3.1, one has

$$D_{a,b} = \bigcap_{n=0}^{\infty} D_n, \quad \text{with} \quad D_n = \bigcap_{i=0}^n F^i(\Theta_{a,b}).$$

LEMMA 6.1. *Suppose that the map f satisfies the finiteness condition. Then, for each $n \geq 0$, D_n is a region consisting of two connected components, the upper one, D_n^u , and the lower one, D_n^ℓ , bounded by nondecreasing step-functions.*

Proof. The proof is by induction on n . The base of induction holds by the definition of the trapping region $\Theta_{a,b}$. For the induction step, let us assume that the region D_n consists of two connected components, the upper one D_n^u and the lower one D_n^ℓ , bounded by nondecreasing step-functions. We will show that the region D_{n+1} consists of two connected components, D_{n+1}^u and D_{n+1}^ℓ , bounded by nondecreasing step-functions.

In what follows, we present the proof assuming that $0 < b \leq -a < 1$. The situation $a \leq -1$ is less complex due to the explicit cycle expressions described in Theorem 4.6 and can be treated similarly with some minor modifications. We decompose the regions D_n^u and D_n^ℓ as follows

$$\begin{aligned} U_n^{11} &= \{(x, y) \in D_n^u \mid y \geq TSa\} & L_n^{11} &= \{(x, y) \in D_n^\ell \mid y \leq T^{-1}Sb\} \\ U_n^{12} &= \{(x, y) \in D_n^u \mid b \leq y \leq TSa\} & L_n^{12} &= \{(x, y) \in D_n^\ell \mid T^{-1}Sb \leq y \leq a\} \\ U_n^3 &= \{(x, y) \in D_n^u \mid 0 \leq y \leq b\} & L_n^3 &= \{(x, y) \in D_n^\ell \mid a \leq y \leq 0\} \\ U_n^{21} &= \left\{ (x, y) \in D_n^u \mid \frac{a}{1-a} \leq y \leq 0 \right\} & L_n^{21} &= \left\{ (x, y) \in D_n^\ell \mid 0 \leq y \leq \frac{b}{b+1} \right\} \\ U_n^{22} &= \left\{ (x, y) \in D_n^u \mid b-1 \leq y \leq \frac{a}{1-a} \right\} & L_n^{22} &= \left\{ (x, y) \in D_n^\ell \mid \frac{b}{b+1} \leq y \leq a+1 \right\}. \end{aligned}$$

By induction hypothesis, the regions U_n^{12} , U_n^3 , U_n^{21} , and U_n^{22} are bounded below and above, and U_n^{11} only below, by a ray and on the right by a nondecreasing step-function. Similarly, the regions L_n^{12} , L_n^3 , L_n^{21} , and L_n^{22} bounded above and below, and L_n^{11} only above, by a ray and on the left by a nondecreasing step-function.

If $B \subset D_n^u$ is one of the upper subregions, let ∂B be the union of the boundary components of B that belong to the boundary of D_n^u , and, similarly, if $B \subset D_n^\ell$ is one of the lower subregions, let ∂B be the union of the boundary components of B that belong to the boundary of D_n^ℓ .

Since $\Theta_{a,b}$ is a trapping region, $F(\Theta_{a,b}) \subset \Theta_{a,b}$, $D_{n+1} = F(D_n) \subset D_n$, and hence $D_{n+1}^u \subset D_n^u$ and $D_{n+1}^\ell \subset D_n^\ell$.

The natural extension map F is piecewise fractional-linear, hence it maps regions bounded by nondecreasing step-functions to regions bounded by nondecreasing step-functions. More precisely, we have

$$\begin{aligned} U_{n+1}^u &= S(U_n^{22} \cup U_n^{21}) \cup T^{-1}(U_n^{11} \cup U_n^{12}) \cup S(L_n^3) \\ U_{n+1}^\ell &= S(L_n^{22} \cup L_n^{21}) \cup T(L_n^{11} \cup L_n^{12}) \cup S(U_n^3). \end{aligned}$$

In order to show that the region D_{n+1}^u is connected, we notice that the region $T^{-1}(U_n^{11} \cup U_n^{12})$ is inside the ‘‘quadrant’’ $[-\infty, 0] \times [b-1, \infty]$ while $S(U_n^{22} \cup U_n^{21})$ is inside the strip $[0, 1] \times [ST^{-1}b, \infty]$. Therefore, they either intersect by a ray of the y -axis, or are disjoint. In the first case, either $T^{-1}ST^{-1}b < Sa$, which implies that $S(L_n^3)$ is inside the connected region $S(U_n^{22} \cup U_n^{21}) \cup T^{-1}(U_n^{11} \cup U_n^{12})$, or $Sa \leq T^{-1}ST^{-1}b$ which implies that the level Sa belongs to the boundary of the trapping region, and again $S(L_n^3)$ is inside the connected region $S(U_n^{22} \cup U_n^{21}) \cup T^{-1}(U_n^{11} \cup U_n^{12})$. Now suppose that the regions $T^{-1}(U_n^{11} \cup U_n^{12})$ and $S(U_n^{22} \cup U_n^{21})$

are disconnected. Notice that the right vertical boundary of the region $S(L_n^3)$ is a ray of the y -axis, thus $S(L_n^3) \cup S(U_n^{22} \cup U_n^{21})$ is a connected region bounded by a nondecreasing step-function. Since $T^{-1}(U_n^{12}) \cap S(L_n^3) = \emptyset$, the nonconnectedness situation may only appear from the intersection of $T^{-1}(U_n^{11})$ and $S(L_n^3)$, i.e., inside the strip $[-1, 0] \times [-1/a, \infty]$. Since f satisfies the finiteness condition, Theorem 5.5 is applicable, and the set $A_{a,b}$ constructed there belongs to each D_n . This is because $A_{a,b} \subset \Theta_{a,b}$, and if $A_{a,b} \subset D_n$, we have $A_{a,b} = F(A_{a,b}) \subset F(D_n) = D_{n+1}$. The set $A_{a,b}$ has finite rectangular structure and contains the strip $[-1, 0] \times [-1/a, \infty]$. Thus the connectedness of the region D_{n+1}^u is proved. Moreover, this argument shows that $\partial T^{-1}(U_n^{11})$ is inside D_{n+1}^u and therefore does not contribute to its boundary, and

$$\partial U_{n+1}^u = \partial(T^{-1}(U_n^{12})) \cup \partial(S(U_n^{22} \cup U_n^{21}) \cup S(L_n^3)).$$

Since $\partial(T^{-1}(U_n^{12}))$ and $\partial(S(U_n^{22} \cup U_n^{21}) \cup S(L_n^3))$ are given by nondecreasing step-functions, one $< Sa$, and the other $\geq Sa$, it follows that ∂U_{n+1}^u is also given by a nondecreasing step-function. A similar argument proves that D_{n+1}^ℓ is connected and bounded by a nondecreasing step-function. \square

LEMMA 6.2. *Suppose that for each n , D_n consists of two connected components as in Lemma 6.1.*

- (1) *All horizontal levels of the boundary of D_n^u belong to $\mathcal{U}_{a,b}$ (resp., D_n^ℓ belong to $\mathcal{L}_{a,b}$) and remain as horizontal levels of D_{n+1}^u (resp., D_{n+1}^ℓ).*
- (2) *All levels of $\mathcal{U}_{a,b}$ appear in the boundary of some D_n^u , and all levels of $\mathcal{L}_{a,b}$ appear in the boundary of some D_n^ℓ .*
- (3) *The attractor $D_{a,b}$ consists of two connected components bounded by nondecreasing step-functions. The upper boundary function takes all values from the set $\mathcal{U}_{a,b}$, and the lower boundary function takes all values from the set $\mathcal{L}_{a,b}$.*
- (4) *The map $F: D_{a,b} \rightarrow D_{a,b}$ is surjective.*

Proof. (1). We prove this by induction. For the base case, D_0^u contains the horizontal levels $T^{-1}b$, $ST^{-1}b$, and $\min(T^{-1}ST^{-1}b, Sa)$. The levels $T^{-1}b$ and $ST^{-1}b$ belong to the boundary of D_1^u . If $Sa < T^{-1}ST^{-1}b$, then $ST^{-1}b > T^{-1}Sa$ and therefore is the end of the cycle and does not belong to $\mathcal{U}_{a,b}$. If $Sa > T^{-1}ST^{-1}b$, then $T^{-1}ST^{-1}b$ appears as a boundary segment of D_1^u . A similar argument applies to D_0^ℓ that contains the horizontal levels Ta , STa , and either $TSTa$ or Sb .

For the induction step, we assume that (1) holds for $k = n - 1$, and prove that it holds for $k = n$. Let $y \in \partial D_n$ be a horizontal segment of the boundary, $y \geq ST^{-1}b$, and $y \in \mathcal{U}_{a,b}$. Then $y = Sy'$, where $y' \in \partial D_{n-1}$, $b - 1 \leq y' < 0$. By inductive hypothesis, $y' \in \partial D_n$, hence $y = Sy' \in \partial D_{n+1}$. Now let $y \in \partial D_n$ be a horizontal segment of the boundary, $b - 1 \leq y < Sa$. Then $y = T^{-1}y'$, where $y' \in \partial D_{n-1}$, $0 < y' < T^{-1}Sa$. By inductive hypothesis, $y' \in \partial D_n$, hence $y = Sy' \in \partial D_{n+1}$.

The level $y = Sa$ appears as a boundary segment of D_n^u since $T^{-1}(\partial(U_{n-1}^{11}) \cup \partial(U_{n-1}^{12}))$ and $S(\partial(L_{n-1}^3))$ do not overlap. Then $y = Sy'$, where $y' = a$ is the y -coordinate of the horizontal lower boundary of L_{n-1}^3 . Since $L_n^3 \subset L_{n-1}^3$ and $U_n^{11} \cup$

$U_n^{12} \subset U_{n-1}^{11} \cup U_{n-1}^{12}$, we get that $T^{-1}(\partial(U_n^{11}) \cup \partial(U_n^{12}))$ and $S(\partial(L_n^3))$ do not overlap, and $y = Sa$ will appear as a boundary segment of D_{n+1}^u .

On the other hand, assume $y \in \partial D_{n+1}$ was not a horizontal level of ∂D_n . Then $y = Sy'$ for some $y' \in \partial(U_n^{22} \cup U_n^{21})$, $y = T^{-1}y'$ for some $y' \in \partial(U_n^{12})$, or $y = Sa$. In all cases $y \in \mathcal{U}_{a,b}$ by the structure of the sets \mathcal{U}_a and \mathcal{U}_b established in Theorems 4.5 and 4.2.

(2). We start with level $-1/(b-1)$ which belongs to the boundary of the trapping region $\Theta_{a,b}$ by definition. We have seen that if $T^{-1}ST^{-1}b \in \mathcal{U}_b$, then the level appears in the boundary of D_1^u . Now, if $b-1 < T^{-k}ST^{-1}b < a/(1-a)$ (for the smallest $k = 2$ or 3), then the expansion continues, each $T^{-i}ST^{-1}b$, $i \leq k$, appears for the first time in the boundary of D_i^u for $i \leq k$, and the next element in the cycle, $ST^{-k}ST^{-1}b$, appears in the boundary of D_{k+1}^u . Using the structure of the set \mathcal{U}_b established in Theorem 4.2 we see that all levels of the set \mathcal{U}_b appear as boundary levels of some D_n^u . We use the same argument for level $-1/a$ which appears for the first time in the boundary of some $D_{n_0}^u$, to see that all elements of the set \mathcal{U}_a appear as boundary levels of all successive sets D_n^u . The same argument works for the lower boundary.

(3). Thus starting with some n , all sets D_n have two connected components bounded by nondecreasing step-functions whose y levels coincide with the sets $\mathcal{U}_{a,b}$ and $\mathcal{L}_{a,b}$. Therefore, the attractor $D_{a,b} = \bigcap_{n=0}^\infty D_n$ has the same property.

(4). The surjectivity of the map F on $D_{a,b}$ follows from the nesting property of the sets D_n . □

A priori, the map F on $D_{a,b}$ does not have to be injective, but in our case it will be since we will identify $D_{a,b}$ with an earlier constructed set $A_{a,b}$.

COROLLARY 6.3. *If the map f satisfies the finiteness condition, then the attractor $D_{a,b}$ has finite rectangular structure, i.e., bounded by nondecreasing step-functions with a finite number of steps.*

THEOREM 6.4. *If the map f satisfies the finiteness condition, then the set $A_{a,b}$ constructed in Theorem 5.5 is the attractor for the map F .*

Proof. We proved in Theorem 5.5 that the set $A_{a,b}$ constructed there is uniquely determined by the prescribed set of y -levels $\mathcal{U}_{a,b} \cup \mathcal{L}_{a,b}$. By Corollary 6.3, the set $D_{a,b}$ has finite rectangular structure with the same set of y -levels. Now we look at the x -levels of the jumps of its boundary step-functions. Take the vertex $(x, b-1)$ of $D_{a,b}$. From the surjectivity of F on $D_{a,b}$, there is a point $z \in D_{a,b}$ such that $F(z) = (x, b-1)$. Then z must be the intersection of the ray at the level b with the boundary of $D_{a,b}$, i.e., $z = (\tilde{x}_b, b)$, hence $x = \tilde{x}_b - 1$. Continue the same argument: look at the vertex at the level $-1/(b-1)$. It must be $F(\tilde{x}_b - 1, b-1)$, etc. Since each y -level of the boundary has a unique “predecessor” in its orbit, all x -levels of the jumps obtained by “transporting” the rays $[-\infty, \tilde{x}_b]$ and $[\tilde{x}_a, \infty]$ over the corresponding cycles, satisfy the same equations that defined the boundary of the set $A_{a,b}$ of Theorem 5.5. Therefore $\tilde{x}_a = x_a$, $\tilde{x}_b = x_b$, the step-functions that define the boundaries are the same, and $D_{a,b} = A_{a,b}$. □

7. REDUCTION THEORY CONJECTURE

Don Zagier conjectured that the Reduction theory properties, stated in the Introduction, hold for every $(a, b) \in \mathcal{P}$. He was motivated by the classical cases and computer experimentations with random parameter values $(a, b) \in \mathcal{P}$ (see Figures 1 and 6 for attractors obtained by iterating random points using Mathematica program).

The following theorem gives a sufficient condition for the Reduction theory conjecture to hold:

THEOREM 7.1. *If both a and b have the strong cycle property, then for every point $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$ there exists $N > 0$ such that $F^N(x, y) \in D_{a,b}$.*

Proof. Every point $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$ is mapped to the trapping region by some iterate F^{N_1} . Since the sets D_n are nested and contain $D_{a,b}$, for large N , $F^N(x, y)$ will be close to the boundary of $D_{a,b}$. By Corollary 5.12, for any boundary component h of $D_{a,b}$ there exists $N_2 > 0$ such that $F^{N_2}(h)$ is inside $D_{a,b}$. Therefore, there exists a large enough $N > 0$ such that $F^N(x, y)$ will be in the interior of $D_{a,b}$. \square

The strong cycle property is not necessary for the Reduction theory conjecture to hold. For example, it holds for the two classical expansions $(-1, 0)$ and $(-1, 1)$ that satisfy only a weak cycle property. In the third classical expansion $(-1/2, 1/2)$ that also satisfies a weak cycle property, property (3) of the conjecture does not hold for some points (x, y) with y equivalent to $r = (3 - \sqrt{5})/2$.

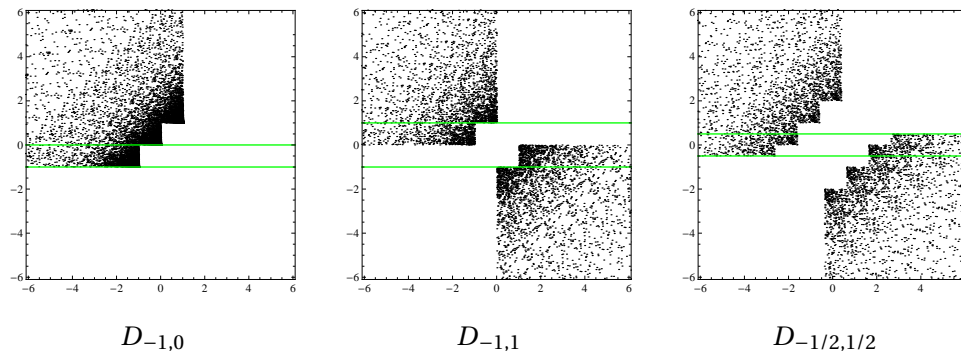


FIGURE 6. Attractors for the classical cases

The next result shows that, under the finiteness condition, almost every point $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$ lands in the attractor $D_{a,b}$ after finitely many iterations.

PROPOSITION 7.2. *If the map $f_{a,b}$ satisfies the finiteness condition, then for almost every point $(x, y) \in \bar{\mathbb{R}}^2 \setminus \Delta$, there exists $N > 0$ such that $F_{a,b}^N(x, y) \in D_{a,b}$.*

Proof. Let $(x, y) \in \mathbb{R}^2$ with y irrational and $y = [n_0, n_1, n_2, \dots]_{a,b}$. In the proof of Theorem 3.1, we showed that there exists $k > 0$ such that

$$(x_{j+1}, y_{j+1}) = ST^{-n_j} \dots ST^{-n_1} ST^{-n_0}(x, y) \in [-1, 1] \times ([-1/a, \infty] \cup [-\infty, -1/b])$$

for all $j \geq k$. The point $F_{a,b}^N(x, y) = (x_{k+1}, y_{k+1})$ is in $A_{a,b}$, if $(x_{k+1}, y_{k+1}) \in [-1, 0] \times [-1/a, \infty]$ or $(x_{k+1}, y_{k+1}) \in [0, 1] \times [-\infty, -1/b]$. Also, $F^{N+1}(x, y) = F(x_{k+1}, y_{k+1})$ is in $A_{a,b}$ if $(x_{k+1}, y_{k+1}) \in [0, 1] \times [-1/a+1, \infty]$ or $(x_{k+1}, y_{k+1}) \in [-1, 0] \times [-\infty, -1/b-1]$. Thus we are left with analyzing the situation when the sequence of iterates

$$(x_{j+1}, y_{j+1}) = ST^{-n_j} \dots ST^{-n_1} ST^{-n_0}(x, y)$$

belongs to $[0, 1] \times [-1/a, -1/a+1]$ for all $j \geq k$ (or $[-1, 0] \times [-1/b, -1/b-1]$ for all $j \geq k$). Assume that we are in the first situation: $y_{j+1} \in [-1/a, -1/a+1]$ for all $j \geq k$. This implies that all digits n_{j+1} , $j \geq k$ are either $\lfloor -1/a \rfloor$ or $\lfloor -1/a \rfloor + 1$. In the second situation, the digits n_{j+1} , $j \geq k$ are either $\lfloor -1/b \rfloor$ or $\lfloor -1/b \rfloor - 1$. Therefore the continued fraction expansion of y is written with only two consecutive digits (starting from a certain position). By using Proposition 2.4 and Remark 2.5 we obtain that the set of all such points has zero Lebesgue measure. This proves our result. \square

REMARK 7.3. In the next section we show that there is a nonempty Cantor-like set $\mathcal{E} \subset \Delta$ belonging to the boundary segment $b = a + 1$ of \mathcal{P} such that for $(a, b) \in \mathcal{E}$ the set $\mathcal{U}_{a,b} \cup \mathcal{L}_{a,b}$ is infinite. Therefore, for $(a, b) \in \mathcal{E}$ either the set D_n^u or D_n^l is disconnected for some $n > 0$, or, by Lemma 6.2(3), the attractor $D_{a,b}$ consists of two connected components whose boundary functions are not step-functions with finitely many steps.

8. SET OF EXCEPTIONS TO THE FINITENESS CONDITION

In this section we study the structure of the set $\mathcal{E} \subset \mathcal{P}$ of exceptions to the finiteness condition. We write $\mathcal{E} = \mathcal{E}_b \cup \mathcal{E}_a$, where \mathcal{E}_b (resp., \mathcal{E}_a) consists of all points $(a, b) \in \mathcal{P}$ for which b (resp., a) does not satisfy the finiteness condition, i.e., either the truncated orbit \mathcal{U}_b or \mathcal{L}_b is infinite (resp., \mathcal{U}_a or \mathcal{L}_a).

We analyze the set \mathcal{E}_b . Recall that, by Proposition 5.3, the set \mathcal{U}_b is infinite if and only if \mathcal{L}_b is infinite, therefore it is sufficient to analyze the condition that the orbit \mathcal{L}_b is not eventually periodic and the values of $\mathcal{L}_b \cap [a, b]$ belong to the interval $(\frac{b}{b+1}, a + 1)$. We point out that this condition also holds for the case of Theorem 4.2(I)(ii) where the lower orbit of b continues with the lower orbit of a ($\mathcal{L}_a \subset \mathcal{L}_b$). For \mathcal{L}_a to be infinite, the values of $\mathcal{L}_a \cap [a, b]$ need to belong to the interval $(\frac{b}{b+1}, a + 1)$, as described by Theorem 4.5.

As before, we restrict our analysis (due to the symmetry considerations) to the parameter subset of \mathcal{P} given by $b \leq -a$ and write $\mathcal{E}_b = \bigcup_{m=3}^{\infty} \mathcal{E}_b^m$ where $b \in \mathcal{E}_b^m$ if $b \in \mathcal{E}_b$ and $T^m Sb \in (\frac{b}{b+1}, a + 1)$. By Theorem 4.2 and its proof, it follows that if $b \in \mathcal{E}_b^m$, then the first digit of the (a, b) -continued fraction expansion of Sb is $-m$ and all the other digits are either $-m$ or $-(m + 1)$.

We describe a recursive construction of the exceptional set \mathcal{E}_b^m . One starts with the ‘triangular’ set

$$\mathcal{F}_b^m = \left\{ (a, b) \in \mathcal{P} \mid \frac{b}{b+1} \leq T^m Sb \leq a + 1 \right\}.$$

The range of possible values of b in \mathcal{F}_b^m is given by the interval $[\underline{b}, \bar{b}]$ where $T^m S \bar{b} = \bar{b}$ and $T^m S \underline{b} = \underline{b}/(\underline{b} + 1)$. Since $\frac{b}{b+1} \leq b$ for all $b \geq 0$, and the function $T^m S b$ is monotone increasing, we obtain that $\underline{b} < \bar{b}$, and \underline{b} is the horizontal boundary of \mathcal{F}_b^m , while \bar{b} is the b -coordinate of its ‘vertex’.

At the next stage we obtain the following regions:

$$\mathcal{F}_b^{m,m} = \left\{ (a, b) \in \mathcal{F}_b^m \mid \frac{b}{b+1} \leq T^m S T^m S b \leq a+1 \right\}$$

$$\mathcal{F}_b^{m,m+1} = \left\{ (a, b) \in \mathcal{F}_b^m \mid \frac{b}{b+1} \leq T^{m+1} S T^m S b \leq a+1 \right\}.$$

By the same argument as above each region is ‘triangular’, *i.e.*, the b -coordinate of its lower (horizontal) boundary is less than the b -coordinate of its vertex. We show that its intersection with the triangular region obtained on the previous step is either empty or has ‘triangular’ shape. The horizontal boundary of $\mathcal{F}_b^{m,m}$ has the b -coordinate given by the relation $T^m S T^m S b = b/(b+1)$ (call it \tilde{b}). We have

$$T^m S T^m S \underline{b} = T^m S \left(\frac{\underline{b}}{\underline{b}+1} \right) = T^m S \underline{b} - 1 = -\frac{1}{\underline{b}+1} < \frac{\underline{b}}{\underline{b}+1},$$

so $\underline{b} < \tilde{b}$. On the other hand,

$$T^m S T^m S \bar{b} = T^m S \bar{b} = \bar{b},$$

which shows that the hyperbola $T^m S T^m S b = b$ intersects the diagonal side $b = a+1$ at the point with b -coordinate \bar{b} . It follows that the region $\mathcal{F}_b^{m,m}$ is triangular and nonempty with $\underline{b} < \tilde{b} < \bar{b}$.

The upper boundary of $\mathcal{F}_b^{m,m+1}$ is given by the hyperbola $T^{m+1} S T^m S b = a+1$. Notice that, if $\underline{a}+1 = T^m S \underline{b}$, then the point $(\underline{a}, \underline{b})$ lies on the curves $T^m S b = a+1$ (obviously) and $T^{m+1} S T^m S b = a+1$ because

$$T^{m+1} S T^m S \underline{b} = T^{m+1} S (b/(\underline{b}+1)) = T^m S \underline{b} = \underline{a}+1.$$

Therefore, the entire horizontal boundary of \mathcal{F}_b^m belongs to that of $\mathcal{F}_b^{m,m+1}$. Moreover, the hyperbola $T^{m+1} S T^m S b = a+1$ intersects the diagonal side $b-a=1$ at the point \hat{b} satisfying $T^{m+1} S T^m S \hat{b} = \hat{b}$. Therefore, $T^m S T^m S \hat{b} = \hat{b} - 1 < \frac{\hat{b}}{\hat{b}+1}$, *i.e.*, $\hat{b} < \tilde{b}$. In this case we have $\underline{b} < \hat{b} < \tilde{b} < \bar{b}$, and the two triangular regions $\mathcal{F}_b^{m,m}$ and $\mathcal{F}_b^{m,m+1}$ are disjoint and nonempty.

The situation becomes more complicated as we proceed recursively. Let $\mathcal{F}_b^{n_1, n_2, \dots, n_k}$ be one of the regions obtained after k steps of this construction, with $n_1 = m$ and $n_i \in \{m, m+1\}$ for $2 \leq i \leq k$. At the next step we get two new sets (possibly empty) (see Figure 7):

$$\mathcal{F}_b^{n_1, n_2, \dots, n_k, m} = \left\{ (a, b) \in \mathcal{F}_b^{n_1, n_2, \dots, n_k} \mid \frac{b}{b+1} \leq T^m S T^{n_k} S \dots T^{n_1} S b \leq a+1 \right\}$$

$$\mathcal{F}_b^{n_1, n_2, \dots, n_k, m+1} = \left\{ (a, b) \in \mathcal{F}_b^{n_1, n_2, \dots, n_k} \mid \frac{b}{b+1} \leq T^{m+1} S T^{n_k} S \dots T^{n_1} S b \leq a+1 \right\}.$$

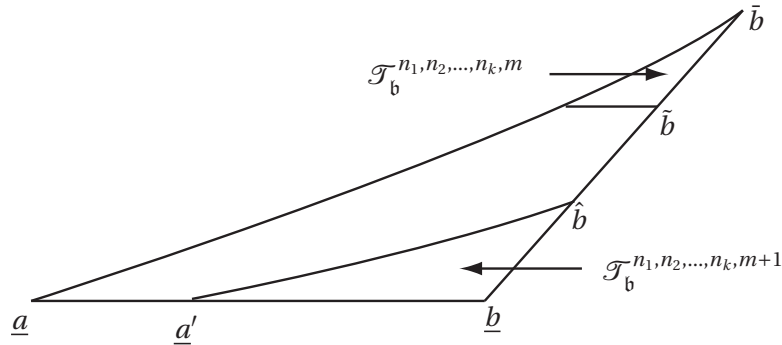


FIGURE 7. Set $\mathcal{T}_b^{n_1, n_2, \dots, n_k}$ and its two subregions

As in the base case, the inequality $T^m S T^{n_k} S \dots T^{n_1} S b \leq a + 1$ of $\mathcal{T}_b^{n_1, n_2, \dots, n_k, m}$ is satisfied by all points of $\mathcal{T}_b^{n_1, n_2, \dots, n_k}$ because of the monotone increasing property of T, S and the fact that $T^{n_k} S \dots T^{n_1} S b \leq a + 1$ implies

$$T^m S T^{n_k} S \dots T^{n_1} S b \leq T^m S(a + 1) \leq T^m S(b) \leq a + 1.$$

Thus the upper boundary of the region $\mathcal{T}_b^{n_1, n_2, \dots, n_k, m}$ (if nonempty) is part of the upper boundary of $\mathcal{T}_b^{n_1, n_2, \dots, n_k}$; it is the lower (horizontal) boundary that changes. In a similar way, the defining inequality $\frac{b}{b+1} \leq T^{m+1} S T^{n_k} S \dots T^{n_1} S b$ of $\mathcal{T}_b^{n_1, n_2, \dots, n_k, m+1}$ is satisfied by all points of $\mathcal{T}_b^{n_1, n_2, \dots, n_k}$ because

$$T^{m+1} S T^{n_k} S \dots T^{n_1} S b \geq T^{m+1} S \frac{b}{b+1} = m - \frac{1}{b} = T^m S b \geq \frac{b}{b+1}.$$

Thus the lower boundary of $\mathcal{T}_b^{n_1, n_2, \dots, n_k, m+1}$ (if nonempty) is part of the lower boundary of $\mathcal{T}_b^{n_1, n_2, \dots, n_k}$. Therefore, we can describe the above sets as

$$\begin{aligned} \mathcal{T}_b^{n_1, n_2, \dots, n_k, m} &= \left\{ (a, b) \in \mathcal{T}_b^{n_1, n_2, \dots, n_k} \mid \frac{b}{b+1} \leq T^m S T^{n_k} S \dots T^{n_1} S b \right\} \\ \mathcal{T}_b^{n_1, n_2, \dots, n_k, m+1} &= \left\{ (a, b) \in \mathcal{T}_b^{n_1, n_2, \dots, n_k} \mid T^{m+1} S T^{n_k} S \dots T^{n_1} S b \leq a + 1 \right\}. \end{aligned}$$

By the same reason as in the base case, the regions $\mathcal{T}_b^{n_1, \dots, n_k, m}$ and $\mathcal{T}_b^{n_1, \dots, n_k, m+1}$ do not overlap.

The set \mathcal{E}_b^m is now obtained as the union of all sets of type

$$\mathcal{E}_b^{(n_i)} = \bigcap_{k=1}^{\infty} \mathcal{T}_b^{n_1, n_2, \dots, n_k}$$

where $n_1 = m$, $n_i \in \{m, m+1\}$ if $i \geq 2$, and the sequence (n_i) is not eventually periodic. If such a set $\mathcal{E}_b^{(n_i)}$ is nonempty and (a, b) belongs to it, then b is uniquely determined from the (a, b) -expansion of $Sb = [-n_1, -n_2, \dots]$.

First we need some additional lemmas.

LEMMA 8.1.

- (i) A point $b \in [0, 1]$ satisfying $T^{n_k} S \cdots T^{n_1} S b = b$ with $|n_i| \geq 2$ can be written formally using a periodic “-” continued fraction expansion

$$(8.1) \quad b = -1/(\overline{-n_1, -n_2, \dots, -n_k}) = (0, \overline{-n_1, -n_2, \dots, -n_k}).$$

If b is in $\mathcal{F}_b^{n_1, n_2, \dots, n_k}$, then Sb has the (a, b) -continued fraction expansion

$$[Sb]_{a,b} = [\overline{-n_1, -n_2, \dots, -n_k}].$$

- (ii) A point b in $[0, 1]$ satisfying $T^{n_k} S \cdots T^{n_1} S b = b/(b+1)$ can be written formally using the periodic “-” continued fraction expansion

$$(8.2) \quad b = (0, \overline{-n_1, -n_2, \dots, -n_k, -(m+1)}).$$

If the point $b \in \mathcal{F}_b^{n_1, n_2, \dots, n_k}$, then $[Sb]_{a,b} = [\overline{-n_1, -n_2, \dots, -n_k, -(m+1)}]$.

Proof. One can verify directly that the point b given by (8.1) is the fixed point of the hyperbolic transformation $T^{n_k} S \cdots T^{n_1} S$ and $b \in [0, 1]$ (see also [11, Proposition 1.3]).

The equation in part (ii) can be written as $STST^{n_k} S \cdots T^{n_1} S b = b$ and one verifies directly that the value b given by (8.2) is the fixed point of that hyperbolic transformation and $b \in [0, 1]$. \square

Notice that the relation $(0, -n_1, -n_2, \dots) = -(0, n_1, n_2, \dots)$ is satisfied, assuming that the formal “-” continued fraction expansions are convergent (from the proof of Theorem 2.1, the convergence property holds if $|n_i| \geq 2$ for all $i \geq 1$).

DEFINITION 8.2. We say that two sequences (finite or infinite) $\sigma_1 = (n_i)$ and $\sigma_2 = (p_j)$ of positive integers are in lexicographic order, $\sigma_1 < \sigma_2$, if on the first position k where the two sequences differ one has $n_k < p_k$, or if the finite sequence (n_i) is a starting subsequence of (p_j) .

The following property follows from the monotonicity of T and S .

LEMMA 8.3. Given two infinite sequences $\sigma_1 = (n_i)$ and $\sigma_2 = (p_j)$ of integers $n_i \geq 2$ and $p_j \geq 2$ such that $\sigma_1 < \sigma_2$, then $(0, n_1, n_2, \dots) < (0, p_1, p_2, \dots)$.

The next lemma provides necessary conditions for a set $\mathcal{E}_b^{(n_i)}$ to be nonempty. Denote by l_m the length of the initial block of m 's and by l_{m+1} the length of the first block of $(m+1)$'s in (n_i) .

LEMMA 8.4.

- (i) If a set $\mathcal{E}_b^{(n_i)}$ in the upper region $\mathcal{F}_b^{m,m}$ is nonempty then the sequence (n_i) contains no consecutive $(m+1)$'s and the length of any block of m 's is equal to l_m or $l_m - 1$.
- (ii) If a set $\mathcal{E}_b^{(n_i)}$ in the lower region $\mathcal{F}_b^{m,m+1}$ is nonempty then the sequence (n_i) contains no consecutive m 's and the length of any block of $(m+1)$'s is equal to l_{m+1} or $l_{m+1} + 1$.

Proof. (i). Assume that the sequence (n_i) contains two consecutive $(m + 1)$'s. Then some $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m+1, m+1}$ (with $n_1 = n_2 = n_k = m$) is nonempty. The upper vertex of such a triangular set satisfies the inequality

$$\bar{b} \leq -(\overline{0, n_1, n_2, \dots, n_k, m+1, m+1}) = -(\overline{0, m, m, \dots, m, m+1, \boxed{m+1}, \dots}),$$

while the lower (horizontal) boundary satisfies

$$\underline{b} \geq -(\overline{0, n_1, n_2, \dots, n_k, m+1}) = -(\overline{0, m, m, \dots, m, m+1, \boxed{m}, \dots}).$$

This implies that $\underline{b} > \bar{b}$ because the entries of the corresponding continued fractions with positive entries are in lexicographic order (they coincide on the first $k + 1$ places, and on the $(k + 2)$ th position the first continued fraction has digit $m + 1$ while the second one has digit m), *i.e.*, the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m+1, m+1}$ is empty.

Now assume that there exists a nonempty set $\mathcal{F}^{n_1, n_2, \dots, n_k, m, m, \dots, m}$ ($n_k = m + 1$) with the final block of m 's of length greater than l_m . The upper vertex of this set is given by

$$\begin{aligned} \bar{b} &\leq -(\overline{0, n_1, n_2, \dots, n_k}) = -(\overline{0, \underbrace{m, m, \dots, m}_{l_m}, m+1, \dots, n_k}) \\ &= -(\overline{0, \underbrace{m, m, \dots, m}_{l_m}, m+1, \dots, n_k, \underbrace{m, m, \dots, m}_{l_m}, m+1, \dots}), \end{aligned}$$

while the lower horizontal segment is given by

$$\underline{b} \geq -(\overline{0, n_1, n_2, \dots, n_k, \underbrace{m, m, \dots, m}_q, m+1}).$$

If $l_m < q$ then the two continued fractions coincide on the first $k + p$ entries. Looking at the $k + p + 1$ entry, we get that $\bar{b} < \underline{b}$, hence the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m, \dots, m}$ would be empty.

Assume now that there exists a nonempty set of type $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m, \dots, m, m+1}$ ($n_k = m + 1$) with the last block of m 's of length q strictly less than $l_m - 1$. Because $n_k = m + 1$, $n_{k-1} = m$, and $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m, \dots, m, m+1} \subset \mathcal{F}_b^{n_1, n_2, \dots, n_k}$ the lower limit of the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m, \dots, m, m+1}$ satisfies the relation

$$\begin{aligned} \underline{b} &\geq -(\overline{0, n_1, n_2, \dots, n_{k-1}, m+1}) = -(\overline{0, n_1, n_2, \dots, n_{k-1}, n_k}) \\ &= -(\overline{0, \underbrace{m, m, \dots, m}_{l_m}, m+1, \dots, n_k, \underbrace{m, \dots, m}_{l_m-1}, m+1, \dots}), \end{aligned}$$

while the upper limit of the same set satisfies the relation

$$\bar{b} \leq -(\overline{0, n_1, n_2, \dots, n_k, \underbrace{m, m, \dots, m}_q, m+1}).$$

This implies that $\bar{b} < \underline{b}$ because the two continued fractions coincide on their first $k + q$ entries, and the $k + q + 1$ entries are m , and $m + 1$ respectively. Therefore the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m, \dots, m, m+1}$ is empty.

(ii). Assume that a set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m}$ with $n_1 = m, n_2 = m + 1$, and $n_k = m + 1$ is nonempty. The upper vertex of such a set satisfies the inequality

$$\bar{b} \leq -(0, \overline{n_1, n_2, \dots, n_k}) = -(0, m, m + 1, \dots, n_k, m, \boxed{m + 1}, \dots),$$

while the lower horizontal segment satisfies the relation

$$\underline{b} \geq -(0, n_1, \overline{n_2, \dots, n_k, m, m, m + 1}) = -(0, m, m + 1, \dots, n_k, m, \boxed{m}, m + 1, \dots).$$

Then $\underline{b} > \bar{b}$ because the sequences of the corresponding continued fractions with positive entries are in lexicographic order, *i.e.*, the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m, m}$ is empty.

Now assume that there exists a nonempty set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m + 1, m + 1, \dots, m + 1}$ ($n_k = m$) with the final block of $(m + 1)$'s of length q greater than $l_{m+1} + 1$. The upper vertex of this set satisfies

$$\bar{b} \leq -(0, m, \underbrace{m + 1, \dots, m + 1}_{l_{m+1}}, m, \dots, n_k, \underbrace{m + 1, \dots, m + 1}_q),$$

while the lower horizontal segment satisfies the relation

$$\begin{aligned} \underline{b} &\geq -(0, n_1, \overline{n_2, \dots, n_k, m + 1}) \\ &= -(0, m, \underbrace{m + 1, \dots, m + 1}_{l_{m+1}}, m, \dots, n_k, \underbrace{m + 1, \dots, m + 1}_{l_{m+1} + 1}, m, \dots). \end{aligned}$$

Since the two continued fraction expansions with positive entries coincide on the first $k + l_{m+1} + 1$ entries and their $k + l_{m+1} + 2$ entries are $m + 1$ and m , respectively, we obtain $\bar{b} < \underline{b}$, *i.e.*, the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m + 1, m + 1, \dots, m + 1}$ is empty.

Finally, suppose that there exists a nonempty set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m + 1, m + 1, \dots, m + 1}$ ($n_k = m$) with the final block of $(m + 1)$'s of length q less than l_{m+1} . The upper vertex of this set satisfies

$$\bar{b} \leq -(0, m, \underbrace{m + 1, \dots, m + 1}_{l_{m+1}}, m, \dots, n_k, \underbrace{m + 1, \dots, m + 1}_{l_{m+1}}),$$

while the lower horizontal segment satisfies the relation

$$\begin{aligned} \underline{b} &\geq -(0, n_1, \overline{n_2, \dots, n_k, \underbrace{m + 1, \dots, m + 1}_q, m, m + 1}) \\ &= -(0, m, \underbrace{m + 1, \dots, m + 1}_{l_{m+1}}, m, \dots, n_k, \underbrace{m + 1, \dots, m + 1}_q, m, \dots). \end{aligned}$$

Since the two continued fraction expansions with positive entries coincide on the first $k + l_{m+1}$ entries and their $(k + l_{m+1} + 1)$ th entries are $(m + 1)$ and m , respectively, we obtain $\bar{b} < \underline{b}$, *i.e.*, the set $\mathcal{F}_b^{n_1, n_2, \dots, n_k, m + 1, m + 1, \dots, m + 1}$ is empty. □

In what follows, we describe in an explicit manner the symbolic properties of a sequence (n_i) for which $\mathcal{E}_b^{(n_i)} \neq \emptyset$. Notice that in both cases of Lemma 8.4,

there are two admissible blocks that can be used to express the admissible sequence (n_i) :

$$A^{(1)} = (\underbrace{m, \dots, m}_{l_m}, m+1) \quad \text{and} \quad B^{(1)} = (\underbrace{m, \dots, m}_{l_m-1}, m+1) \quad \text{with } l_m \geq 2;$$

$$A^{(1)} = (m, \underbrace{m+1, \dots, m+1}_{l_{m+1}}) \quad \text{and} \quad B^{(1)} = (m, \underbrace{m+1, \dots, m+1}_{l_{m+1}+1}) \quad \text{with } l_{m+1} \geq 1.$$

In both situations, $A^{(1)} < B^{(1)}$. One could think of $A^{(1)}$ as being the new ‘ m ’ and $B^{(1)}$ the new ‘ $m+1$ ’, and treat the original sequence of m ’s and $m+1$ ’s as a sequence of $A^{(1)}$ ’s and $B^{(1)}$ ’s. Furthermore, the next lemma shows that such a substitution process can be continued recursively to construct blocks $A^{(n)}$ and $B^{(n)}$ for any $n \geq 1$, so the original sequence (n_i) may be considered to be a sequence of $A^{(n)}$ ’s and $B^{(n)}$ ’s. Moreover, only particular blocks of $A^{(n)}$ ’s and $B^{(n)}$ ’s warrant nonempty triangular regions of the next generation.

Let us also introduce the notations $A^{(0)} = m$ and $B^{(0)} = m+1$. Assume that $\mathcal{E}_b^{(n_i)}$ is a nonempty set.

LEMMA 8.5. *For every $n \geq 0$, there exist integers $l_{A^{(n)}} \geq 2$, $l_{B^{(n)}} \geq 1$ such that the sequence (n_i) can be written as a concatenation of blocks*

$$(8.3) \quad A^{(n+1)} = (\underbrace{A^{(n)}, \dots, A^{(n)}}_{l_{A^{(n)}}}, B^{(n)}), \quad B^{(n+1)} = (\underbrace{A^{(n)}, \dots, A^{(n)}}_{l_{A^{(n)}}-1}, B^{(n)})$$

or

$$(8.4) \quad A^{(n+1)} = (A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}}), \quad B^{(n+1)} = (A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}+1}).$$

Proof. Notice that Lemma 8.4 proves the above result for $n = 0$ with $l_{A^{(0)}} = l_m$, $l_{B^{(0)}} = l_{m+1}$. We show inductively that

$$(8.5) \quad A^{(n)} < B^{(n)}$$

and if a finite sequence σ starts with an $A^{(n)}$ block and ends with a $B^{(n)}$ block, $\sigma = (A^{(n)}, \tau, B^{(n)})$, then the lower boundary $\underline{b}(\sigma)$ of \mathcal{T}_b^σ (if nonempty) satisfies

$$(8.6) \quad \underline{b}(\sigma) \geq -(0, A^{(n)}, \overline{\tau, B^{(n)}}).$$

Relation (8.5) is obviously true for $n = 0$; (8.6) is also satisfied if $n = 0$, since one applies Lemma 8.1(ii) to the sequence $\tilde{\sigma} = (A^{(0)}, \tau)$ where $\mathcal{T}_b^{\tilde{\sigma}} \supset \mathcal{T}_b^\sigma$. We point out that by applying Lemma 8.1(i) to the region \mathcal{T}^σ we have

$$(8.7) \quad \bar{b}(\sigma) \leq -(0, \bar{\sigma}) = -(0, \overline{A^{(n)}, \tau, B^{(n)}}).$$

To prove the inductive step, suppose that for some $n \geq 1$, we can rewrite the sequence (n_i) using blocks $A^{(n+1)}$ and $B^{(n+1)}$ as in case (8.3) or (8.4).

Case I: Assume $A^{(n+1)}$ and $B^{(n+1)}$ are given by (8.3). It follows immediately that $A^{(n+1)} < B^{(n+1)}$ since $A^{(n)} < B^{(n)}$. Also, if a sequence σ starts with an $A^{(n+1)}$

block and ends with a $B^{(n+1)}$ block (thus, implicitly, σ starts with an $A^{(n)}$ block and ends with a $B^{(n)}$ block),

$$\sigma = (A^{(n+1)}, \tau, B^{(n+1)}) = \underbrace{(A^{(n)}, \dots, A^{(n)})}_{l_{A^{(n)}}}, B^{(n)}, \tau, \underbrace{(A^{(n)}, \dots, A^{(n)})}_{l_{A^{(n)}}-1}, B^{(n)}$$

then, by applying (8.6) to

$$\tilde{\sigma} = \underbrace{(A^{(n)}, \dots, A^{(n)})}_{l_{A^{(n)}}}, B^{(n)}, \tau = (A^{(n)}, B^{(n+1)}, \tau),$$

which starts with $A^{(n)}$ and ends with $B^{(n)}$, we get

$$\underline{b}(\sigma) \geq b(\tilde{\sigma}) \geq -(0, A^{(n)}, \overline{B^{(n+1)}, \tau}) = -(0, A^{(n)}, B^{(n+1)}, \overline{\tau, B^{(n+1)}}).$$

Therefore, (8.6) holds for $n + 1$, since $(A^{(n)}, B^{(n+1)}) = A^{(n+1)}$.

Now assume that (n_i) starts with a block of $A^{(n+1)}$'s of length $l_{A^{(n+1)}} > 1$. We prove that the sequence (n_i) cannot have two consecutive $B^{(n+1)}$'s and any sequence of consecutive blocks $A^{(n+1)}$ has length $l_{A^{(n+1)}}$ or $l_{A^{(n+1)}}-1$. Suppose the sequence (n_i) contains two consecutive blocks of type $B^{(n+1)}$:

$$(n_i) = (A^{(n+1)}, A^{(n+1)}, \dots, A^{(n+1)}, B^{(n+1)}, B^{(n+1)}, \dots).$$

We look at the set

$$\mathcal{F}^{A^{(n+1)}A^{(n+1)}\dots A^{(n+1)}B^{(n+1)}B^{(n+1)}}$$

and remark that the upper boundary satisfies (from (8.7))

$$(8.8) \quad \bar{b} \leq -(0, \overline{A^{(n+1)}, A^{(n+1)}, \dots, A^{(n+1)}, B^{(n+1)}, B^{(n+1)}})$$

and the lower boundary satisfies (from (8.6))

$$(8.9) \quad \underline{b} \geq -(0, A^{(n+1)}, \overline{A^{(n+1)}, \dots, A^{(n+1)}, B^{(n+1)}}).$$

But (8.8) and (8.9) imply that $\underline{b} > \bar{b}$, because the two corresponding continued fractions with positive entries are in lexicographic order. Thus, there cannot be two consecutive $B^{(n+1)}$ blocks in the sequence (n_i) .

Now, let us check that the sequence (n_i) cannot have a block of $A^{(n+1)}$'s of length $q > l_{A^{(n+1)}}$. Assume the contrary,

$$(n_i) = \underbrace{(A^{(n+1)}, \dots, A^{(n+1)})}_{l_{A^{(n+1)}}}, B^{(n+1)}, \tau, B^{(n+1)}, \underbrace{(A^{(n+1)}, \dots, A^{(n+1)})}_q, B^{(n+1)}, \dots.$$

Then the set $\mathcal{F}_b^{(n_i)}$ has the upper bound \bar{b} satisfying

$$\bar{b} \leq -(0, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{l_{A^{(n+1)}}}, B^{(n+1)}, \tau, B^{(n+1)})$$

while the lower bound \underline{b} satisfies by (8.6)

$$\underline{b} \geq -(0, A^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{l_{A^{(n+1)}}-1}, B^{(n+1)}, \tau, B^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_q, B^{(n+1)}).$$

Comparing the two continued fractions, we get that $\bar{b} < \underline{b}$, since $A^{(n+1)} < B^{(n+1)}$ and $q > l_{A^{(n+1)}}$.

Now assume that (n_i) starts with $A^{(n+1)}$ and then continues with a block of $B^{(n+1)}$'s of length $l_{B^{(n+1)}} \geq 1$. We prove that the sequence (n_i) cannot have two consecutive $A^{(n+1)}$'s and any sequence of consecutive blocks $B^{(n+1)}$ has length $l_{B^{(n+1)}}$ or $l_{B^{(n+1)}} + 1$. Suppose the sequence (n_i) contains two (or more) consecutive blocks of type $A^{(n+1)}$:

$$(n_i) = (A^{(n+1)}, B^{(n+1)}, \tau, B^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{q \geq 2}, B^{(n+1)}, \dots).$$

We study the region $\mathcal{F}^{A^{(n+1)}, B^{(n+1)}, \tau, B^{(n+1)}, A^{(n+1)}, \dots, A^{(n+1)}, B^{(n+1)}}$ and remark that its upper boundary satisfies (from (8.7))

$$(8.10) \quad \bar{b} \leq -\overline{(0, A^{(n+1)}, B^{(n+1)}, \tau, B^{(n+1)})}$$

and the lower boundary satisfies (from (8.6))

$$(8.11) \quad \underline{b} \geq -\overline{(0, A^{(n+1)}, B^{(n+1)}, \tau, B^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{q \geq 2}, B^{(n+1)})}.$$

But (8.10) and (8.11) imply that $\underline{b} > \bar{b}$ because the two corresponding continued fractions with positive entries are in lexicographic order. Thus, there cannot be two consecutive $A^{(n+1)}$ blocks in the sequence (n_i) .

Now, let us check that the sequence (n_i) cannot have a block of $B^{(n+1)}$'s of length $q > l_{B^{(n+1)}} + 1$. Assume the contrary,

$$(n_i) = (A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}}, A^{(n+1)}, \tau, A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_q, A^{(n+1)}, \dots).$$

Then the set $\mathcal{F}^{(n_i)}$ has the upper bound \bar{b} satisfying

$$\bar{b} \leq -\overline{(0, A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}}, A^{(n+1)}, \tau, A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_q, A^{(n+1)})}$$

while the lower bound \underline{b} satisfies by (8.6)

$$\underline{b} \geq -\overline{(0, A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}}, A^{(n+1)}, \tau, A^{(n+1)}, B^{(n+1)})}.$$

Comparing the two continued fractions, we get that $\bar{b} < \underline{b}$.

Case 2: Assume $A^{(n+1)}$ and $B^{(n+1)}$ are given by (8.4). It follows that $A^{(n+1)} < B^{(n+1)}$ since $A^{(n+1)}$ is the beginning block of $B^{(n+1)}$. Also, if a sequence σ starts with an $A^{(n+1)}$ block and ends with a $B^{(n+1)}$ block (thus, implicitly, σ starts with an $A^{(n)}$ block and ends with a $B^{(n)}$ block),

$$\sigma = (A^{(n+1)}, \tau, B^{(n+1)}) = (A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}}, \tau, A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}+1}).$$

Then by applying (8.6) to

$$\tilde{\sigma} = (A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}}, \tau, A^{(n)}, B^{(n)}),$$

which starts with $A^{(n)}$ and ends with $B^{(n)}$, we get

$$\begin{aligned} \underline{b}(\sigma) &\geq \underline{b}(\tilde{\sigma}) \geq -(0, A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)}}}, \tau, A^{(n)}, B^{(n)}) \\ &= -(0, A^{(n+1)}, \tau, A^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{l_{B^{(n)+1}}}). \end{aligned}$$

Thus, (8.6) holds for $n + 1$.

Assume that (n_i) starts with a sequence of $A^{(n+1)}$'s of length $l_{A^{(n+1)}} > 1$. Similar to the analysis of the first case, one proves that the sequence (n_i) cannot have two consecutive $B^{(n+1)}$'s and any sequence of consecutive blocks $A^{(n+1)}$ has length $l_{A^{(n+1)}}$ or $l_{A^{(n+1)}} - 1$.

If the sequence (n_i) starts with $A^{(n+1)}$ and then continues with a sequence of $B^{(n+1)}$'s of length $l_{B^{(n+1)}} \geq 1$, one can prove that the sequence (n_i) cannot have two consecutive $A^{(n+1)}$'s and any sequence of consecutive blocks $B^{(n+1)}$ has length $l_{B^{(n+1)}}$ or $l_{B^{(n+1)}} + 1$. \square

Additionally, we prove the following.

LEMMA 8.6. *If the block $\tau_1 = (n_i, \dots, n_l)$ is a tail of $A^{(n)}$ and $\tau_2 = (p_j, \dots, p_h)$ is a tail of $B^{(n)}$, then $A^{(n)} < \tau_1$ and $B^{(n)} < \tau_2$.*

Proof. The statement is obviously true if $n = 1$. Assume it is true for some n both for $A^{(n)}$ and $B^{(n)}$. We analyze the case of $A^{(n+1)}$ being given by (8.3), $A^{(n+1)} = (\underbrace{A^{(n)}, \dots, A^{(n)}}_{l_{A^{(n)}}}, B^{(n)})$. Consider an arbitrary tail τ of $A^{(n+1)}$; τ could start

with a block $A^{(n)}$ or a tail of $A^{(n)}$ or τ coincides with $B^{(n)}$ or a tail of $B^{(n)}$. In all situations, the inductive hypothesis and the fact that $A^{(n)} < B^{(n)}$ prove that $A^{(n+1)} < \tau$. The case of $A^{(n+1)}$ given by (8.4) is treated similarly. \square

REMARK 8.7. Using the relations (8.6) and (8.7), notice that a set $\mathcal{F}_b^{A^{(n+1)}}$ (if nonempty) has the upper vertex satisfying

$$(8.12) \quad \bar{b}_{n+1} \leq -(0, \overline{A^{(n+1)}})$$

and a lower horizontal boundary that satisfies

$$(8.13) \quad \underline{b}_{n+1} \geq -(0, A^{(n+1)}, \overline{B^{(n+1)}})$$

if $A^{(n+1)}$ is given by the substitution rule (8.3), and

$$(8.14) \quad \underline{b}_{n+1} \geq -(0, A^{(n)}, \overline{B^{(n)}})$$

if $A^{(n+1)}$ is given by (8.4).

We will prove that the above inequalities are actually equality relations. For that we construct a starting subsequence of $A^{(n+1)}$ defined inductively as

$$\sigma^{(1)} = \begin{cases} \underbrace{(m, \dots, m)}_{l_m} & \text{if } A^{(1)} = \underbrace{(m, \dots, m, m+1)}_{l_m} \\ (m) & \text{if } A^{(1)} = \underbrace{(m, m+1, \dots, m+1)}_{l_{m+1}} \end{cases}$$

Case 1: If $A^{(n)}$ is given by a relation of type (8.3), i.e., if we have $A^{(n)} = (A^{(n-1)}, \dots, A^{(n-1)}, B^{(n-1)})$, then

$$(8.15) \quad \sigma^{(n+1)} = \begin{cases} \underbrace{(A^{(n)}, \dots, A^{(n)}, \sigma^{(n)})}_{l_A^{(n)}-1} & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, \dots, A^{(n)}, B^{(n)})}_{l_{A^{(n)}}} \\ \sigma^{(n)} & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, B^{(n)}, \dots, B^{(n)})}_{l_{B^{(n)}}} \end{cases}$$

Case 2: If $A^{(n)}$ is given by a relation of type (8.4), i.e., if we have $A^{(n)} = (A^{(n-1)}, B^{(n-1)}, \dots, B^{(n-1)})$, then

$$(8.16) \quad \sigma^{(n+1)} = \begin{cases} \underbrace{(A^{(n)}, \dots, A^{(n)}, \sigma^{(n)})}_{l_A^{(n)}} & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, \dots, A^{(n)}, B^{(n)})}_{l_{A^{(n)}}} \\ (A^{(n)}, \sigma^{(n)}) & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, B^{(n)}, \dots, B^{(n)})}_{l_{B^{(n)}}} \end{cases}$$

We introduce the notation f^σ to denote the transformation $T^{n_k} S \dots T^{n_1} S$ if $\sigma = (n_1, \dots, n_k)$.

LEMMA 8.8. *Let $\sigma^{(n+1)}$ be the starting block of $A^{(n+1)}$ defined as above. Then the equation*

$$f^{\sigma^{(n+1)}} b = \frac{b}{b+1}$$

has a unique solution $b \in [0, 1]$ given by

$$(8.17) \quad b_{n+1} = \begin{cases} -(0, A^{(n+1)}, \overline{B^{(n+1)}}) & \text{if } A^{(n+1)} \text{ is given by (8.3),} \\ -(0, A^{(n)}, \overline{B^{(n)}}) & \text{if } A^{(n+1)} \text{ is given by (8.4).} \end{cases}$$

Proof. We proceed with an inductive proof, and as part of it we also show that

$$(8.18) \quad (\sigma^{(n+1)}, m+1, \tilde{A}^{(n)}) = \begin{cases} A^{(n+1)} & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, \dots, A^{(n)}, B^{(n)})}_{l_{A^{(n)}}} \\ (A^{(n)}, B^{(n)}) & \text{if } A^{(n+1)} = \underbrace{(A^{(n)}, B^{(n)}, \dots, B^{(n)})}_{l_{B^{(n)}}} \end{cases}$$

where $A^{(n)} = (m, \tilde{A}^{(n)})$. The relation (8.17) is true for $n = 0$ due to Lemma 8.1(ii). Also, (8.18) follows immediately. Suppose now that the inductive relations hold for some n . We analyze the solution of $f^{\sigma^{(n+2)}} b = \frac{b}{b+1}$.

Assume that $A^{(n+1)} = \underbrace{(A^{(n)}, \dots, A^{(n)})}_{l_{A^{(n)}}}, B^{(n)}$. We look at the two possible cases:

Case 1: If $A^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}}, B^{(n+1)})$, $\sigma^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-1}, \sigma^{(n+1)})$.

Using Lemma 8.1(ii), the solution to $f^{\sigma^{(n+2)}} b = \frac{b}{b+1}$ is given by

$$\begin{aligned} b_{n+2} &= -(\overline{(0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-2}, \sigma^{(n+1)}, m+1)}) \\ &= -(\overline{(0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-2}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n+1)})}) \\ &= -(\overline{(0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-2}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n)}, B^{(n+1)})}) \\ &= -(\overline{(0, A^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-2}, A^{(n+1)}, B^{(n+1)})}) \\ &= -(\overline{(0, A^{(n+1)}, B^{(n+2)})}) = -(\overline{(0, A^{(n+2)}, B^{(n+2)})}). \end{aligned}$$

Also,

$$\begin{aligned} (\sigma^{(n+2)}, m+1, \tilde{A}^{(n+1)}) &= (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-1}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n)}, B^{(n+1)}) \\ &= (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}-1}, A^{(n+1)}, B^{(n+1)}) = A^{(n+2)}. \end{aligned}$$

Case 2: If $A^{(n+2)} = (A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{I_{B^{(n+1)}}})$, then $\sigma^{(n+2)} = \sigma^{(n+1)}$, and the in-

duction step gives us the solution of $f^{\sigma^{(n+2)}} b = \frac{b}{b+1}$ as $b_{n+2} = -(\overline{(0, A^{(n+1)}, B^{(n+1)})})$. Also,

$$(\sigma^{(n+2)}, m+1, \tilde{A}^{(n+1)}) = (\sigma^{(n+1)}, m+1, \tilde{A}^{(n)}, B^{(n+1)}) = (A^{(n+1)}, B^{(n+1)}).$$

Now assume that $A^{(n+1)} = (A^n, \underbrace{B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}})$. We look again at the two possible cases:

Case 1: If $A^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}}, B^{(n+1)})$, $\sigma^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}}, \sigma^{(n+1)})$.

Using Lemma 8.1(ii), the solution to $f^{\sigma^{(n+2)}} b = \frac{b}{b+1}$ is given by

$$\begin{aligned} b_{n+2} &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}-1}}}_{I_{A^{(n+1)}-1}}, \sigma^{(n+1)}, m+1) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}-1}}}_{I_{A^{(n+1)}-1}}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n+1)}) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}-1}}}_{I_{A^{(n+1)}-1}}, \sigma^{(n+1)}, m+1, \underbrace{\tilde{A}^{(n)}, B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}-1}}}_{I_{A^{(n+1)}-1}}, \underbrace{A^{(n)}, B^{(n)}, B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}) \\ &= -(\underbrace{0, A^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}-1}}}_{I_{A^{(n+1)}-1}}, \underbrace{B^{(n+1)}}_{I_{B^{(n)}}}) = -(\underbrace{0, A^{(n+2)}, \underbrace{B^{(n+2)}}_{I_{B^{(n)}}}}_{I_{A^{(n+1)}-1}}). \end{aligned}$$

A similar approach gives us that $(\sigma^{(n+2)}, m+1, \tilde{A}^{(n+1)}) = A^{(n+2)}$.

Case 2: If $A^{(n+2)} = (A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{I_{B^{(n+1)}}})$, then $\sigma^{(n+2)} = (A^{(n+1)}, \sigma^{(n+1)})$. Using Lemma 8.1(ii), the solution to $f^{\sigma^{(n+2)}} b = \frac{b}{b+1}$ is given by

$$\begin{aligned} b_{n+2} &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \sigma^{(n+1)}, m+1}_{I_{B^{(n+1)}}}) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n+1)}}_{I_{B^{(n+1)}}}) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}}_{I_{B^{(n+1)}}}) \\ &= -(\underbrace{0, m, \tilde{A}^{(n+1)}, \sigma^{(n+1)}, \underbrace{A^{(n)}, B^{(n)}, B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}}_{I_{B^{(n+1)}}}) \\ &= -(\underbrace{0, A^{(n+1)}, \underbrace{B^{(n+1)}}_{I_{B^{(n)}}}}_{I_{B^{(n+1)}}}). \end{aligned}$$

Also,

$$\begin{aligned} (\sigma^{(n+2)}, m+1, \tilde{A}^{(n+1)}) &= (A^{(n+1)}, \sigma^{(n+1)}, m+1, \tilde{A}^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}) \\ &= (A^{(n+1)}, A^n, B^{(n)}, \underbrace{B^{(n)}, \dots, B^{(n)}}_{I_{B^{(n)}}}) = (A^{(n+1)}, B^{(n+1)}). \quad \square \end{aligned}$$

THEOREM 8.9. Any sequence (n_i) constructed recursively using relations (8.3) and (8.4) provides a nonempty set $\mathcal{E}_b^{(n_i)}$.

Proof. We prove inductively that any set $\mathcal{T}_b^{A^{(n+1)}}$ is nonempty and the relations (8.12) and (8.13) or (8.14) are actual equalities, *i.e.*,

$$(8.19) \quad \bar{b}_{n+1} = -(0, \overline{A^{(n+1)}})$$

and a lower horizontal boundary that satisfies

$$(8.20) \quad \underline{b}_{n+1} = -(0, A^{(n+1)}, \overline{B^{(n+1)}})$$

if $A^{(n+1)}$ is given by the substitution rule (8.3) or $\underline{b}_{n+1} = -(0, A^{(n)}, \overline{B^{(n)}})$ if $A^{(n+1)}$ is given by (8.4). As part of the inductive proof, we also show that any tail block τ of $A^{(n+1)}$, $\tau \neq \tau^{(n+1)}$ satisfies $\tau < \tau^{(n+1)}$, where $\tau^{(n+1)}$ denotes the tail block of $A^{(n+1)}$ obtained by eliminating the starting block $\sigma^{(n+1)}$ defined by (8.15) or (8.16).

Indeed for $n = 0$, one can check directly that the sets $\mathcal{T}_b^{m,m,\dots,m,m+1}$ and $\mathcal{T}_b^{m,m+1,\dots,m+1}$ satisfy the above equalities using the fact that an “ m ” digit does not change the position of the upper vertex, while an “ $m + 1$ ” digit does not change the position of the horizontal segment of such a triangular set. Also, for any tail $\tau \neq \tau^{(1)}$ of $A^{(1)}$, $\tau < \tau^{(1)}$.

Now, let us assume that $\mathcal{T}_b^{A^{(n+1)}}$ obtained from $A^{(n+1)} = (\underbrace{A^{(n)}, \dots, A^{(n)}}_{I^{(n)}}, B^{(n)})$ is nonempty and satisfies (8.19) and (8.20). For $\mathcal{T}_b^{A^{(n+2)}}$ we look at the two possible cases:

Case 1: $A^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)}}}, B^{(n+1)})$. By Remark 8.7,

$$\bar{b}_{n+2} \leq -(0, \overline{A^{(n+2)}}) = -(0, \overbrace{A^{(n+1)}, \dots, A^{(n+1)}}^{I_{A^{(n+1)}}}, B^{(n+1)}) =: \hat{b}$$

and

$$\underline{b}_{n+2} \geq -(0, A^{(n+2)}, \overline{B^{(n+2)}}) = -(0, A^{(n+1)}, \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)-1}}, B^{(n+1)}) =: \tilde{b}$$

where \tilde{b} was obtained by applying Lemma 8.1(ii) to the starting block

$$\sigma^{(n+2)} = (\underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_{I_{A^{(n+1)-1}}, \sigma^{(n+1)})$$

of $A^{(n+2)}$.

We prove first the other inductive step: any tail block τ of $A^{(n+2)}$, $\tau \neq \tau^{(n+2)}$ satisfies $\tau < \tau^{(n+2)}$. Notice that $\tau^{(n+2)} = (\tau^{(n+1)}, B^{(n+1)})$. There exists τ' a tail block of $A^{(n+1)}$ with the property that

$$\tau = (\tau', \underbrace{A^{(n+1)}, \dots, A^{(n+1)}}_l, B^{(n+1)}), \quad 0 \leq l \leq I_{A^{(n+1)-1}},$$

or $\tau = \tau'$. The latter case holds when τ is just a tail of $B^{(n+1)}$ (which itself is a tail of $A^{(n+1)}$). It is possible that $\tau' = \emptyset$, but in this case $\tau < \tau^{(n+2)}$ because

$A^{(n+1)} < \tau^{(n+1)}$ by Lemma 8.6. If $\tau' \neq \emptyset$, we also get that $\tau < \tau^{(n+2)}$ by using the inductive hypothesis relation $\tau' < \tau^{(n+1)}$.

Now we show that the points $(\tilde{b}-1, \tilde{b})$ and $(\hat{b}-1, \hat{b})$ belong to the set $\mathcal{F}_b^{A^{(n+2)}}$. The point $(\hat{b}-1, \hat{b})$ belongs to $\mathcal{F}_b^{A^{(n+1)}}$ so $f^{A^{(n+1)}} \hat{b} \leq \hat{b}$. If σ is an intermediate block between $A^{(n+1)}$ and $A^{(n+2)}$, $A^{(n+1)} \subset \sigma \subset A^{(n+2)}$, then

$$f^\sigma(\hat{b}) = -(0, \tau, \overline{A^{(n+2)}}) \leq -(0, \overline{A^{(n+2)}}) = \hat{b}.$$

The inequality is due to the fact that τ is a tail block of $A^{(n+2)}$ obtained by eliminating σ , so $A^{(n+2)} < \tau$.

Now we show that $f^\sigma(\tilde{b}) \geq \tilde{b}/(\tilde{b}+1)$ for any intermediate block σ between $A^{(n+1)}$ and $A^{(n+2)}$. We have that $f^{\sigma^{(n+2)}}(\tilde{b}) = \tilde{b}/(\tilde{b}+1)$ by Lemma 8.8, and

$$f^{\sigma^{(n+2)}}(\tilde{b}) = -(0, \tau^{(n+2)}, \overline{B^{(n+2)}}),$$

where $\tau^{(n+2)} = (\tau^{(n+1)}, B^{(n+1)})$. Also $f^\sigma(\tilde{b}) = -(0, \tau, \overline{B^{(n+2)}})$ with τ being the tail block of $A^{(n+2)}$ obtained by eliminating σ . But $\tau < \tau^{(n+2)}$ as we have just proved, hence $f^\sigma(\tilde{b}) \geq f^{\sigma^{(n+2)}}(\tilde{b})$.

In conclusion, any intermediate block σ between $A^{(n+1)}$ and $A^{(n+2)}$ satisfies

$$\tilde{b}/(\tilde{b}+1) \leq f^\sigma(\tilde{b}) \leq f^\sigma(\hat{b}) \leq \hat{b}.$$

Therefore the points $(\tilde{b}-1, \tilde{b})$ and $(\hat{b}-1, \hat{b})$ belong to the intermediate set \mathcal{F}_b^σ . This proves the induction step for $\mathcal{F}_b^{A^{(n+2)}}$.

Case 2: $A^{(n+2)} = (A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}})$. By Remark 8.7,

$$\tilde{b}_{n+2} \leq -(0, \overline{A^{(n+2)}}) = -(0, A^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}}) =: \hat{b}$$

and $\underline{b}_{n+2} \geq -(0, A^{(n+1)}, \overline{B^{(n+1)}}) =: \tilde{b}$, where \tilde{b} was obtained by applying Lemma 8.1(ii) to the starting block $\sigma^{(n+2)} = \sigma^{(n+1)}$ of $A^{(n+2)}$.

We prove first the other inductive step: any tail block τ of $A^{(n+2)}$, $\tau \neq \tau^{(n+2)}$, satisfies $\tau < \tau^{(n+2)}$. There exists τ' a tail block of $A^{(n+1)}$ with the property that

$$\tau = (\tau', \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_l), \quad 0 \leq l \leq l_{B^{(n+1)}}$$

(again, using the fact that $B^{(n+1)}$ is a tail block of $A^{(n+1)}$). Since

$$\tau^{(n+2)} = (\tau^{(n+1)}, \underbrace{B^{(n+1)}, \dots, B^{(n+1)}}_{l_{B^{(n+1)}}}),$$

we get that $\tau < \tau^{(n+2)}$ by using the inductive hypothesis $\tau' < \tau^{(n+1)}$.

Now we show that the points $(\tilde{b}-1, \tilde{b})$ and $(\hat{b}-1, \hat{b})$ belong to the set $\mathcal{F}_b^{A^{(n+2)}}$. The point $(\hat{b}-1, \hat{b})$ belongs to $\mathcal{F}_b^{A^{(n+1)}}$ so $f^{A^{(n+1)}} \hat{b} \leq \hat{b}$. If σ is an intermediate block between between $A^{(n+1)}$ and $A^{(n+2)}$ then

$$f^\sigma(\hat{b}) = -(0, \tau, \overline{A^{(n+2)}}) \leq -(0, \overline{A^{(n+2)}}) = \hat{b}$$

because τ is a tail block of $A^{(n+2)}$ obtained by eliminating σ , so $A^{(n+2)} < \tau$.

Now we show that $f^\sigma(\tilde{b}) \geq \tilde{b}/(\tilde{b} + 1)$. We have that $f^{\sigma^{(n+2)}}(\tilde{b}) = \tilde{b}/(\tilde{b} + 1)$ by Lemma 8.8, and

$$f^{\sigma^{(n+2)}}(\tilde{b}) = -(0, \tau^{(n+1)}, \overline{B^{(n+1)}}), \quad f^\sigma(\tilde{b}) = -(0, \tau, \overline{B^{(n+1)}}),$$

with τ being the end block of $A^{(n+2)}$ obtained by eliminating σ . But $\tau < \tau^{(n+2)}$ as we have just proved, hence $f^\sigma(\tilde{b}) \geq f^{\sigma^{(n+2)}}(\tilde{b})$. In conclusion, any intermediate sequence σ between $A^{(n+1)}$ and $A^{(n+2)}$ satisfies

$$\tilde{b}/(\tilde{b} + 1) \leq f^\sigma(\tilde{b}) \leq f^\sigma(\hat{b}) \leq \hat{b}.$$

Therefore the points $(\tilde{b} - 1, \tilde{b})$ and $(\hat{b} - 1, \hat{b})$ belong to the intermediate set \mathcal{T}_b^σ .

We proved the induction step for $\mathcal{T}_b^{A^{(n+2)}}$, when $A^{(n+1)}$ is given by (8.3). A similar argument can be provided for the case when $A^{(n+1)}$ is given by (8.4), so the conclusion of the theorem is true. \square

We prove now that each set nonempty set $\mathcal{E}^{(n_i)}$ with (n_i) not eventually aperiodic sequence is actually a singleton.

THEOREM 8.10. *Assume that (n_i) is a not eventually periodic sequence such that the set $\mathcal{E}_b^{(n_i)}$ is nonempty. Then, $\mathcal{E}_b^{(n_i)}$ is a point on the line segment $b - a = 1$.*

Proof. The sequence (n_i) satisfies the recursive relations (8.3) or (8.4). We look at the set $\mathcal{T}_b^{A^{(n+1)}}$ and estimate the length of its lower base. In case (8.3) its upper vertex is given by (8.19) and its lower base satisfies (8.20). The lower base is a segment whose right end coordinate is

$$\underline{a}_{n+1}^r = -(0, A^{(n+1)}, \overline{B^{(n+1)}}) - 1$$

and left end coordinate is

$$\underline{a}_{n+1}^l = f^{A^{(n+1)}}(-(0, A^{(n+1)}, \overline{B^{(n+1)}})) - 1 = -(0, \overline{B^{(n+1)}}) - 1.$$

Hence the length of the lower base is given by

$$L_{n+1} = \underline{a}_n^r - \underline{a}_{n+1}^l = (0, \overline{B^{(n+1)}}) - (0, A^{(n+1)}, \overline{B^{(n+1)}}).$$

In case (8.4), the lower base is a segment whose right end coordinate is

$$\underline{a}_{n+1}^r = -(0, A^{(n)}, \overline{B^{(n)}}) - 1$$

and the left end coordinate is given by

$$\underline{a}_{n+1}^l = f^{A^{(n+1)}}(-(0, A^{(n)}, \overline{B^{(n)}})) - 1 = -(0, \overline{B^{(n)}}) - 1.$$

Hence the length of the lower base is given by

$$L_{n+1} = \underline{a}_{n+1}^r - \underline{a}_{n+1}^l = (0, \overline{B^{(n)}}) - (0, A^{(n)}, \overline{B^{(n)}}).$$

Notice that in the first case the two continued fraction expansions have in common at least the block $A^{(n)}$, while in the second case they have in common at least the block $A^{(n-1)}$. This implies that in both cases $L_{n+1} \rightarrow 0$ as $n \rightarrow \infty$. Moreover, the bases of the sets $\mathcal{T}_b^{n_1, \dots, n_k}$ have nonincreasing length and we have

found a subsequence of these bases whose lengths converge to zero. Therefore the set $\mathcal{E}_b^{(n_i)}$ consists of only one point $(b-1, b)$, where $b = -(0, n_1, n_2, \dots)$. \square

The above result gives us a complete description of the set of exceptions \mathcal{E}_b to the finiteness condition. It is a subset of the boundary segment $b = a + 1$ of \mathcal{P} . Moreover, each set \mathcal{E}_b^m is uncountable because the recursive construction of a nonempty set $\mathcal{E}_b^{(n_i)}$ allows for an arbitrary number of successive blocks $A^{(k)}$ at step $(k+1)$. Formally, one constructs a surjective map $j: \mathcal{E}_b^m \rightarrow \mathbb{N}^{\mathbb{N}}$ by associating to a singleton set $\mathcal{E}_b^{(n_i)}$ a sequence of positive integers defined as

$$j(\mathcal{E}_b^{(n_i)})(k) = \# \text{ of consecutive } A^{(k)}\text{-blocks at the beginning of } (n_i).$$

The set \mathcal{E}_b has one-dimensional Lebesgue measure 0. The reason is that all associated formal continued fractions expansions of $b = -(0, n_1, n_2, \dots)$ have only two consecutive digits; such formal expansions $(0, n_1, n_2, \dots)$ are valid $(-1, 0)$ -continued fractions. Hence the set of such b 's has measure zero by Proposition 2.4. Analogous conclusions hold for \mathcal{E}_a . Thus we have the following.

THEOREM 8.11. *For any $(a, b) \in \mathcal{P}$, $b \neq a + 1$, the finiteness condition holds. The set of exceptions \mathcal{E} to the finiteness condition is an uncountable set of one-dimensional Lebesgue measure 0 that lies on the boundary $b = a + 1$ of \mathcal{P} .*

Now we are able to provide the last ingredient in the proof of part (ii) of the Main Result:

PROPOSITION 8.12. *The strong cycle property is an open and dense condition.*

Proof. It follows from Theorems 4.2 and 4.5 that the condition is open. Theorem 8.11 asserts that for all $(a, b) \in \mathcal{P}$, $b \neq a + 1$ the finiteness condition holds, i.e., all we need to show is that if b has the weak cycle property or the (a, b) -expansions of Sb and $T^{-1}b$ are eventually periodic, then in any neighborhood of it there is a b with the strong cycle property. For, if b has the weak cycle property, it is a rational number obtained from the equation $\hat{f}^n T^m Sb = 0$, and any small perturbation of it will have the strong cycle property. Similarly, if the (a, b) -expansions of Sb and $T^{-1}b$ are eventually periodic, then b is a quadratic irrationality (see Remark 2.3), and for any neighborhood of b will contain values satisfying the strong cycle property. A similar argument holds for Sa and Ta . \square

9. INVARIANT MEASURES AND ERGODIC PROPERTIES

Based on the finite rectangular geometric structure of the domain $D_{a,b}$, one can study the measure-theoretic properties of the Gauss-type map $\hat{f}_{a,b}: [a, b) \rightarrow [a, b)$

$$\hat{f}_{a,b}(x) = -1/x - [-1/x]_{a,b} \quad \hat{f}_{a,b}(0) = 0$$

and its associated natural extension map $\hat{F}_{a,b}: \hat{D}_{a,b} \rightarrow \hat{D}_{a,b}$

$$(9.1) \quad \hat{F}_{a,b}(x, y) = \left(\hat{f}_{a,b}(x), -\frac{1}{y - [-1/x]_{a,b}} \right).$$

We remark that $\hat{F}_{a,b}$ is obtained from the map $F_{a,b}$ induced on the set $D_{a,b} \cap \{(x, y) \mid a \leq y < b\}$ by a change of coordinates $x' = y$, $y' = -1/x$. Therefore the domain $\hat{D}_{a,b}$ is easily identified knowing $D_{a,b}$ and may be considered its “compactification”.

We present the simple case when $1 \leq -1/a \leq b+1$ and $a-1 \leq -1/b \leq -1$. The general theory is the subject our paper in preparation [14]. The truncated orbits of a and b are

$$\begin{aligned} \mathcal{L}_a &= \{a+1, -1/(a+1)\}, & \mathcal{U}_a &= \{-1/a, -(a+1)/a\}, \\ \mathcal{L}_b &= \{-1/b, (b-1)/b\}, & \mathcal{U}_b &= \{b-1, -1/(b-1)\}, \end{aligned}$$

and the end-points of the cycles are $c_a = a/(a+1)$ and $c_b = b/(1-b)$, respectively.

THEOREM 9.1. *If $1 \leq -1/a \leq b+1$ and $a-1 \leq -1/b \leq -1$, then the domain $\hat{D}_{a,b}$ of $\hat{F}_{a,b}$ is given by*

$$\begin{aligned} \hat{D}_{a,b} &= [a, 1-1/b] \times [-1, 0] \cup [1-1/b, a+1] \times [-1/2, 0] \\ &\quad \cup [b-1, -1-1/a] \times [0, 1/2] \cup [-1-1/a, b] \times [0, 1], \end{aligned}$$

and $\hat{F}_{a,b}$ preserves the Lebesgue equivalent probability measure

$$d\nu_{a,b} = \frac{1}{\log[(1+b)(1-a)]} \frac{dx dy}{(1+xy)^2}.$$

Proof. The description of $\hat{D}_{a,b}$ follows directly from the cycle relations and the finite rectangular structure. It is a standard computation that the measure $\frac{dx dy}{(y-x)^2}$ is preserved by Möbius transformations applied to real variables x and y , hence by $F_{a,b}$. It follows that the measure $\frac{dx dy}{(1+xy)^2}$ obtained by the coordinate changes $x' = y$, $y' = -1/x$ is preserved by the transformation $\hat{F}_{a,b}$ given in (9.1). By considering $\hat{F}_{a,b}$ as a reduction map acting on geodesics (see Remark 1.1), this can be derived more elegantly by using the geodesic flow on the modular surface and the Poincaré cross-section maps, as explained in [14]. Moreover, the density $1/(1+xy)^2$ is bounded away from zero on $\hat{D}_{a,b}$ and

$$\int_{\hat{D}_{a,b}} \frac{dx dy}{(1+xy)^2} = \log[(b+1)(1-a)] < \infty.$$

Hence the last part of the theorem is true. \square

The Gauss-type map $\hat{f}_{a,b}$ is a factor of $\hat{F}_{a,b}$ (projecting on the x -coordinate), so one can obtain its smooth invariant measure $d\mu_{a,b}$ by integrating $d\nu_{a,b}$ over $\hat{D}_{a,b}$ with respect to the y -coordinate as explained in [2]. Thus, if we know the exact shape of the set $D_{a,b}$, we can calculate the invariant measure precisely.

The measure $d\mu_{a,b}$ is ergodic and the measure-theoretic entropy of $\hat{f}_{a,b}$ can be computed explicitly using Rokhlin’s formula.

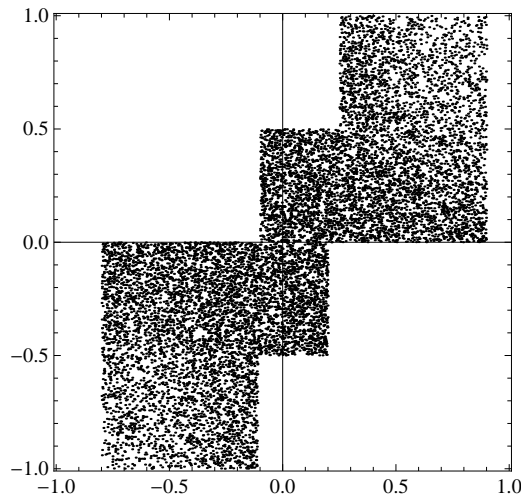


FIGURE 8. Typical domain $\hat{D}_{a,b}$ for the case studied

THEOREM 9.2. *The map $\hat{f}_{a,b}: [a, b] \rightarrow [a, b]$ is ergodic with respect to Lebesgue equivalent invariant probability measure*

$$d\mu_{a,b} = \frac{1}{C_{a,b}} \left(\frac{\chi_{(a,1-1/b)}}{1-x} + \frac{\chi_{(1-1/b,a+1)}}{2-x} + \frac{\chi_{(b-1,-1-1/a)}}{x+2} + \frac{\chi_{(-1-1/a,b)}}{x+1} \right) dx$$

where $C_{a,b} = \log[(1+b)(1-a)]$. The measure-theoretic entropy of $\hat{f}_{a,b}$ is given by

$$(9.2) \quad h_{\mu_{a,b}}(\hat{f}_{a,b}) = \frac{\pi^2}{3\log[(1-a)(1+b)]}.$$

Proof. The measure $d\mu_{a,b}$ is obtained by integrating $dv_{a,b}$ over $\hat{D}_{a,b}$. Ergodicity follows from a more general result concerning one-dimensional expanding maps (see [2, 24]). To compute the entropy, we use Rokhlin’s formula

$$\begin{aligned} h_{\mu_{a,b}}(\hat{f}_{a,b}) &= \int_a^b \log|\hat{f}'_{a,b}| d\mu_{a,b} = -2 \int_a^b \log|x| d\mu_{a,b} \\ &= \frac{-2}{C_{a,b}} \left(\int_a^{1-1/b} \frac{\log|x|}{1-x} dx + \int_{1-1/b}^{a+1} \frac{\log|x|}{2-x} dx \right. \\ &\quad \left. + \int_{b-1}^{-1-1/a} \frac{\log|x|}{x+2} dx + \int_{-1-1/a}^b \frac{\log|x|}{x+1} dx \right). \end{aligned}$$

Let $I(a, b)$ denote the sum of the four integrals. The function depends smoothly on a, b , hence we can compute the partial derivatives $\partial I/\partial a$ and $\partial I/\partial b$. We get that both partial derivatives are zero, hence $I(a, b)$ is constant. Using $a = -1$, $b = 1$, we get

$$I(a, b) = I(-1, 1) = 2 \int_0^1 \frac{\log|x|}{1+x} dx = -\pi^2/6,$$

and the entropy formula (9.2). □

REFERENCES

- [1] R. Adler and L. Flatto, *The backward continued fraction map and geodesic flow*, Ergod. Th. & Dynam. Sys., **4** (1984), 487–492.
- [2] R. Adler and L. Flatto, *Geodesic flows, interval maps, and symbolic dynamics*, Bull. Amer. Math. Soc., **25** (1991), 229–334.
- [3] E. Artin, *Ein mechanisches system mit quasiergodischen Bahnen*, Abh. Math. Sem. Univ. Hamburg, **3** (1924), 170–175.
- [4] J. Bourdon, B. Daireaux and B. Vallée, *Dynamical analysis of α -Euclidean algorithms*, J. Algorithms, **44** (2002), 246–285.
- [5] C. Carminati and G. Tiozzo, *A canonical thickening of \mathbb{Q} and the dynamics of continued fractions*, preprint [arXiv:1004.3790v1](https://arxiv.org/abs/1004.3790v1).
- [6] G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers," sixth edition, Revised by D. R. Heath-Brown and J. H. Silverman, Oxford University Press, Oxford, 2008.
- [7] A. Hurwitz, *Über eine besondere Art der Kettenbruch-Entwicklung reeler Grössen*, (German), Acta Math., **12** (1889), 367–405.
- [8] S. Katok, "Fuchsian Groups," Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1992.
- [9] S. Katok, *Coding of closed geodesics after Gauss and Morse*, Geom. Dedicata, **63** (1996), 123–145.
- [10] S. Katok and I. Ugarcovici, *Arithmetic coding of geodesics on the modular surface via continued fractions*, European women in mathematics–Marseille 2003, 59–77, CWI Tract, 135, Centrum Wisk. Inform., Amsterdam, (2005).
- [11] S. Katok, I. Ugarcovici, *Geometrically Markov geodesics on the modular surface*, Moscow Math. J. **5** (2005), 135–151.
- [12] S. Katok and I. Ugarcovici, *Symbolic dynamics for the modular surface and beyond*, Bull. Amer. Math. Soc., **44** (2007), 87–132.
- [13] S. Katok and I. Ugarcovici, *Theory of (a, b) -continued fraction transformations and applications*, Electron. Res. Announc. Math. Sci., **17** (2010), 20–33.
- [14] S. Katok and I. Ugarcovici, *Applications of (a, b) -continued fraction transformations*, in preparation.
- [15] C. Kraaikamp, *A new class of continued fraction expansions*, Acta Arith., **57** (1991), 1–39.
- [16] L. Luzzi and S. Marmi, *On the entropy of Japanese continued fractions*, Discrete Cont. Dyn. Syst., **20** (2008), 673–711.
- [17] P. Moussa, A. Cassa and S. Marmi, *Continued fractions and Brjuno functions*, Continued fractions and geometric function theory (CONFUN) (Trondheim, 1997), J. Comput. Appl. Math., **105** (1999), 403–415.
- [18] H. Nakada, *Metric theory for a class of continued fraction transformations and their natural extensions*, Tokyo J. Math., **4** (1981), 399–426.
- [19] H. Nakada and R. Natsui, *Some metric properties of α -continued fractions*, Journal of Number Theory, **97** (2002), 287–300.
- [20] H. Nakada and R. Natsui *The non-monotonicity of the entropy of α -continued fraction transformations*, Nonlinearity, **21** (2008) 1207–1225.
- [21] C. Series, *On coding geodesics with continued fractions*, Ergodic theory (Sem., Les Plans-sur-Bex, 1980) (French), 67–76, Monograph. Enseign. Math., 29, Univ. Genève, Geneva, (1981).
- [22] F. Schweiger, "Ergodic Theory of Fibred Systems and Metric Number Theory," Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1995.
- [23] D. Zagier, "Zetafunktionen und Quadratische Körper: Eine Einführung in die höhere Zahlentheorie," Springer-Verlag, 1981.
- [24] R. Zweimüller, *Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points*, Nonlinearity, **11** (1998), 1263–1276.

SVETLANA KATOK <katok_s@math.psu.edu>: Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

ILIE UGARCOVICI <iugarcov@depaul.edu>: Department of Mathematical Sciences, DePaul University, Chicago, IL 60614, USA