



UNIVERSITÀ
DEGLI STUDI
DI UDINE

Università degli studi di Udine

Slow continued fractions, transducers, and the Serret theorem

Original

Availability:

This version is available <http://hdl.handle.net/11390/1118527> since 2021-03-15T19:45:33Z

Publisher:

Published

DOI:10.1016/j.jnt.2017.08.034

Terms of use:

The institutional repository of the University of Udine (<http://air.uniud.it>) is provided by ARIC services. The aim is to enable open access to all the world.

Publisher copyright

(Article begins on next page)

SLOW CONTINUED FRACTIONS, TRANSDUCERS, AND THE SERRET THEOREM

GIOVANNI PANTI

ABSTRACT. A basic result in the elementary theory of continued fractions says that two real numbers share the same tail in their continued fraction expansions iff they belong to the same orbit under the projective action of $\mathrm{PGL}_2 \mathbb{Z}$. This result was first formulated in Serret's *Cours d'algèbre supérieure*, so we'll refer to it as to the Serret theorem.

Notwithstanding the abundance of continued fraction algorithms in the literature, a uniform treatment of the Serret result seems missing. In this paper we show that there are finitely many possibilities for the groups $\Sigma \leq \mathrm{PGL}_2 \mathbb{Z}$ generated by the branches of the Gauss maps in a large family of algorithms, and that each Σ -equivalence class of reals is partitioned in finitely many tail-equivalence classes, whose number we bound. Our approach is through the finite-state transducers that relate Gauss maps to each other. They constitute opfibrations of the Schreier graphs of the groups, and their synchronizability—which may or may not hold—assures the a.e. validity of the Serret theorem.

1. INTRODUCTION

Let α, β be irrational numbers, with infinite regular continued fraction expansions $[0, a_1, a_2, \dots]$, $[0, b_1, b_2, \dots]$, respectively. It is a classical fact that these expansions have the same tail (i.e., there exists $t_1, t_2 \geq 0$ such that $a_{t_1+n} = b_{t_2+n}$ for every $n \geq 1$) if and only if α and β are conjugated by an element of the extended modular group $\mathrm{PGL}_2 \mathbb{Z}$. This result first appeared as §16 of the third edition (1866) of Serret's *Cours d'algèbre supérieure* [22], the second edition (1854) making no mention of continued fractions; easily accessible modern references are [9, §10.11], [14, §9.6], [5, §2.7]. An equivalent reformulation is that α, β (without loss of generality in the real unit interval) are in the same $\mathrm{PGL}_2 \mathbb{Z}$ -orbit iff they have the same eventual orbit under the Gauss map $T : x \mapsto x^{-1} - \lfloor x^{-1} \rfloor$ (see Figure 1 right). The key point here is that the c. f. expansion of, say, α is nothing else than its T -symbolic orbit: $T^t(\alpha) \in [(a_{t+1} + 1)^{-1}, a_{t+1}^{-1})$ for every $t \geq 0$. We refer to [6, Chapter 7], [7, Chapter 3] and references therein for the interpretation of continued fractions in terms of dynamical systems.

Besides the regular “Floor” one, a great number of continued fraction algorithms appear in the literature, the complex of them forming a large passacaglia on the theme of the euclidean algorithm. As a definitely incomplete list we cite the Ceiling, Nearest Integer, Even, Odd, Farey fractions [2], the α -fractions [17], [1], and the (a, b) -fractions [13], not to mention algorithms with coefficients in rings of algebraic integers and multidimensional continued fractions. Asking for the status of the Serret result for these systems is then quite natural.

Key words and phrases. Continued fractions, Gauss maps, tail property, extended modular group, transducers.

2010 Math. Subj. Class.: 11A55; 37A45.

In this paper we give a fairly complete answer for the algorithms in a certain specific class, namely the class of accelerations of Gauss-type maps arising from finite unimodular partitions of a unimodular interval. After setting notation and stating a few well known facts, we introduce our class in §2; we provide various explicit examples, showing that our class, albeit nonexhaustive, contains many important and much studied algorithms. It is a fortunate fact that the validity—or lack of it—of the Serret property is untouched by the acceleration process, so that we can restrict to “slow” algorithms. In §3 we associate a graph \mathcal{G}_T to each such algorithm T , and show that \mathcal{G}_T is an opfibration of the Schreier graph of the group Σ_T generated by the branches of T . The rôle of Σ_T is clearly crucial; indeed, if α, β have the same eventual T -orbit then they must necessarily be Σ_T -equivalent. Thus, the key question becomes “In how many tail-equivalence classes is partitioned a given Σ_T -equivalence class?”, the Serret property amounting to the constant answer “Precisely one”. In §4 we show that the index of each Σ_T in $\mathrm{PGL}_2 \mathbb{Z}$ is at most 8, so that there are finitely many possibilities for these groups. In §5 we introduce finite-state transducers, and employ them in two ways: in Lemma 5.1 to relate different algorithms to each other, and in Lemma 5.5 to compute the expansion of a rational function of α directly from the expansion of α ; neither use is new, see [8, §3.5] for the first and [19], [15] for the second. In Theorem 5.3 we answer the question cited above: every Σ_T -equivalence class is partitioned in finitely many tail-equivalence classes, whose number is bounded by the *defect* of the algorithm. We also give an explicit criterion, Corollary 5.6, for deciding the validity of the Serret property for a given T . In the final §6 we relate the synchronizability of the graph \mathcal{G}_T to the almost-everywhere (w.r.t. the Lebesgue measure) validity of the Serret theorem.

Before defining our class, we fix notation and recall a few well-known facts. We denote the group $\mathrm{PGL}_2 \mathbb{Z}$ and its index-2 subgroup $\mathrm{PSL}_2 \mathbb{Z}$ by Π and Γ , respectively. We set names for certain elements of Π , using square brackets to emphasize that matrices are taken up to multiplication by ± 1 :

$$L = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}, N = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix}, S = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}, R = \begin{bmatrix} 1 & -1 \\ & 1 \end{bmatrix}, F = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}.$$

The following facts are well known:

- (1) $\Gamma = \langle L, N \rangle = \langle S, R \rangle$; it has the presentation $\langle s, r | s^2 = r^3 = 1 \rangle$ and hence is isomorphic to $Z_2 * Z_3$.
- (2) The set of elements of Γ having nonnegative entries is the free monoid \mathcal{M} on the two free generators L, N .
- (3) $\Pi = \langle S, R, F \rangle$ with the presentation $\langle s, r, f | s^2 = r^3 = f^2 = 1, fs = sf, fr = r^2f \rangle$; it is isomorphic to $\langle S, F \rangle *_{\langle F \rangle} \langle R, F \rangle \simeq D_2 *_{Z_2} D_3$.
- (4) The automorphism group of Γ is Π , acting by conjugation. The outer automorphism group of Π has order 2, and is generated by the involution

$$\alpha : \begin{cases} F \leftrightarrow F, \\ S \leftrightarrow \begin{bmatrix} -1 & \\ & 1 \end{bmatrix}, \\ R \leftrightarrow R^2. \end{cases}$$

Note that Γ is not α -invariant. Moreover, the usual distinction of elements of Γ in elliptic, parabolic, and hyperbolic (according to the absolute value

of the trace being less than, equal, or greater than 2) is destroyed: for example, α exchanges the parabolic N^2 with the hyperbolic $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$.

- (5) Π acts on $P^1 \mathbb{R}$ in the standard projective way: if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $A * \alpha = (a\alpha + b)/(c\alpha + d)$. If Σ is a subgroup of Π and α, β are in the same Σ -orbit, then we say that α and β are Σ -equivalent.

2. GAUSS-TYPE MAPS AND THE GROUPS THEY GENERATE

We identify (actually, we define) a continued fraction algorithm with the corresponding Gauss-type map; in our setting, the latter are defined as follows. Let us say that an interval with rational endpoints $[p/q, p'/q']$ is *unimodular* if $\det \begin{pmatrix} p & p' \\ q & q' \end{pmatrix} = -1$. A *unimodular partition* of the base unimodular interval $\Delta = [0, \infty] = [0/1, 1/0]$ is a finite family $\{\Delta_0, \dots, \Delta_{n-1}\}$ (of cardinality at least 2) of unimodular intervals such that $\bigcup_a \Delta_a = \Delta$ and distinct Δ_a 's intersect at most in a common endpoint; we always assume that $\Delta_0, \dots, \Delta_{n-1}$ are listed in consecutive order, with $0 \in \Delta_0$ and $\infty \in \Delta_{n-1}$. For every index a , we choose arbitrarily $e_a \in \{+1, -1\}$.

Definition 2.1. The *slow continued fraction algorithm* determined by the family of pairs $\{(\Delta_a, e_a)\}$ as above is the map $T : \Delta \rightarrow \Delta$ which is induced on $\Delta_a = [p/q, p'/q']$ by the matrix

$$A_a^{-1} := \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}^{(e_a+1)/2} \begin{bmatrix} p & p' \\ q & q' \end{bmatrix}^{-1}.$$

If Δ_a and Δ_b are consecutive and $e_a = e_b$, then the above definition is ambiguous in their common vertex p'/q' . In this case we consider T as a multivalued map, admitting both 0 and ∞ as T -images of p'/q' ; this ambiguity may occur at most once along the T -orbit of a point, necessarily rational.

We usually specify T by providing the finite set $\{A_0, \dots, A_{n-1}\}$ of matrices in Π with nonnegative entries whose inverses determine T . Note that the defining intervals Δ_a are recovered by $\Delta_a = A_a * \Delta$. We identify matrices with the maps they induce, and we say that the A_a 's are the *inverse branches* of T .

At least one of the extreme points $0, \infty$ is inevitably an indifferent fixed point either for T or for T^2 (*indifferent* means that the one-sided derivative at the point is ± 1). One removes such points by accelerating the algorithm, as follows.

Definition 2.2. Let T be as above, let $0 \leq i \leq j \leq n-1$, and let $E = E(i, j) = \bigcup\{\Delta_a : i \leq a \leq j\}$; we allow the possibility of removing one or both endpoints from the interval E . The *first-return map*, or *accelerated continued fraction algorithm*, is the map $T_E : E \rightarrow E$ defined as follows. Given $x \in E$, let $r(x) = \min\{t \geq 1 : T^t(x) \in E\} \in [1, 2, 3, \dots, +\infty]$, and set

$$T_E(x) = \begin{cases} T^{r(x)}(x), & \text{if } r(x) \neq +\infty; \\ \text{undefined,} & \text{otherwise.} \end{cases}$$

Example 2.3. The set $\{LF, N\}$ (equivalently, the set $\{([0, 1], -1), ([1, \infty], +1)\}$) determines the map whose graph is in Figure 1 left. Since we want ∞ to appear as an ordinary point, we draw graphs by conjugating E to $[0, 1]$ via an appropriate projective transformation, in this case $L : E = \Delta \rightarrow [0, 1]$. The fixed point ∞ is indifferent, and by inducing on $E = [0, 1)$ we get the usual Gauss map of Figure 1 right, whose inverse branches are $\{LFN^a : a \geq 0\} = \{\begin{bmatrix} 1 & 1 \\ 1 & a \end{bmatrix} : a \geq 1\}$.

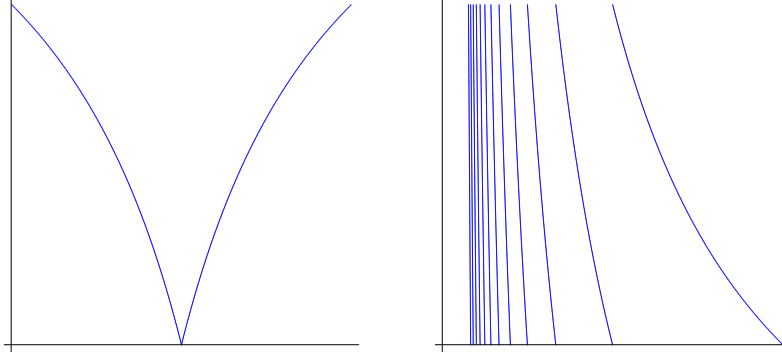


FIGURE 1.

Example 2.4. Conjugating the slow algorithm of Example 2.3 by F we get the Farey map $\{L, NF\}$ on $[0, \infty]$; see, e.g., [10] and references therein. Inducing on $(1, \infty]$ we get a conjugated Gauss map with inverse branches $\left\{\begin{bmatrix} a & 1 \\ 1 & 1 \end{bmatrix} : a \geq 1\right\}$, whose finite products give the classical matrices $\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}$.

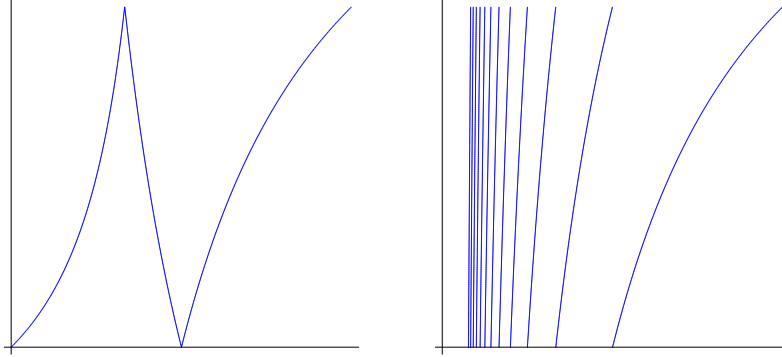


FIGURE 2.

Example 2.5. Starting from $\{L, N\}$ and inducing on $[1, \infty]$ we get Zagier's ceiling algorithm (Figure 2 right). Inducing on $[0, 1]$ we get Rényi's algorithm; the two are of course conjugated by F . See [24], [18], [11] and references therein.

Example 2.6. The map $\{LL, LNF, N\}$ in Figure 2 left enumerates all pythagorean triples [20].

Example 2.7. The maps $\{LL, LNF, N\}$ and $\{LL, LN, NF\}$ are, up to conjugation by L , the even and odd continued fractions in [3].

If T is a slow algorithm and T_E one of its accelerations, then we say that T_E is a *c. f. algorithm*, or *Gauss-type map*. Writing $E = E(i, j)$ as in Definition 2.2, it is easy to see that the inverse branches of T_E are the matrices of the form $A_a B$ with $i \leq a \leq j$ and B in the monoid generated by $\{A_b : b < i \text{ or } j < b\}$. This set of inverse branches is countable (unless $E = \Delta$), and it is clear that the group it generates equals the subgroup Σ_T of Π generated by A_0, \dots, A_{n-1} . If $\alpha, \beta \in E$

are such that $T_E^{t_1}(\alpha) = T_E^{t_2}(\beta)$ for some $t_1, t_2 \geq 0$, then obviously α, β are Σ_T -equivalent; this paper deals with the reverse implication.

Definition 2.8. We say that T_E has the tail property, or that the Serret theorem holds for T_E , if for every $\alpha, \beta \in E \setminus \mathbb{Q}$ that are Σ_T -equivalent, and whose forward T_E -orbit is never undefined, there exist $t_1, t_2 \geq 0$ such that $T_E^{t_1}(\alpha) = T_E^{t_2}(\beta)$.

We discuss the above definition in the following remarks.

Remark 2.9. For every slow algorithm T and every rational $\alpha = p/q \in \Delta$, the T -orbit of α ends up either in the fixed point 0, or in the fixed point ∞ , or in the 2-cycle $\{0, \infty\}$. This is readily proved by observing that whenever p/q is in the topological interior of one of the intervals Δ_a , then $T(\alpha) = p'/q'$ satisfies $0 < p' + q' < p + q$. As a consequence, the T_E -orbit of any rational number either ends up in $\{0, \infty\}$, or is eventually undefined (this surely happens if $0, \infty \notin E$).

On the other hand, it is easy to construct accelerated maps that are undefined in points not Π -equivalent. For example, define T by the inverse branches $\{L, NL, N^2F\}$; it is increasing on $\Delta_0 = [0, 1]$, increasing on $\Delta_1 = [1, 2]$ with fixed point $(1 + \sqrt{5})/2$, and decreasing on $\Delta_2 = [2, \infty]$ with fixed point $1 + \sqrt{2}$. If we induce on $E = [0, 1]$, then the points $\alpha = L * ((1 + \sqrt{5})/2)$, $\beta = L * (1 + \sqrt{2}) \in E$ lie in different quadratic fields, and hence are not Π -equivalent. However, T_E is undefined in both α and β , since their T -images are T -fixed points external to E . It is therefore safe to discard rational and eventually undefined points from consideration.

Remark 2.10. Let $T_{[0,1]}$ denote the classical Gauss map; its inverse branches generate Π . As noted in §1, the irrational numbers α, β have the same eventual $T_{[0,1]}$ -orbit (namely, they satisfy the condition in Definition 2.8) iff their regular c. f. expansions have the same tail. Thus the formulation in Definition 2.8 is equivalent to the classical one.

Remark 2.11. If T has the tail property, then so does any of its accelerations. Conversely, if T_E has the tail property and $\alpha, \beta \in \Delta \setminus \mathbb{Q}$ are Σ_T -equivalent and such that their T -orbits enter E infinitely often, then $T^{t_1}(\alpha) = T^{t_1}(\beta)$ for some t_1, t_2 .

Example 2.12. Here is a simple example of an algorithm for which the tail property fails; we'll construct a more elaborate one in Example 5.4. Let T have inverse

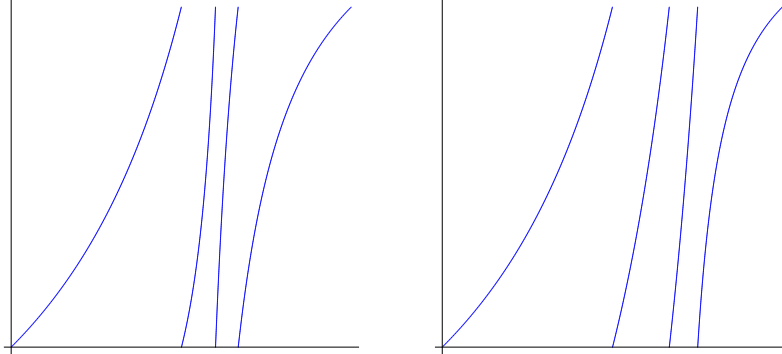


FIGURE 3.

branches $\{L, NLL, NLN, NN\}$, and let T' have $\{L, NL,>NNL, NNN\}$. Their graphs, shown in Figure 3, are similar, and obviously $\Sigma_T = \Sigma_{T'} = \Gamma$ (see also Corollary 4.3). We will show in Corollary 5.6 that T' has the tail property. On the other hand, consider the third inverse branch $A_2 = NLN$ of T , and let $\alpha = A_2 * \alpha = \sqrt{3}$ be the fixed point of T in Δ_2 . Of course α has T -symbolic orbit $\bar{2}$. However $NA_2N^{-1} = A_3A_0$, and thus $\beta = N * \alpha = \sqrt{3} + 1$ has T -symbolic orbit $\overline{30}$; hence the tail property fails.

3. THE GRAPH OF AN ALGORITHM

As noted in Remark 2.11, the validity of the Serret theorem for a slow algorithm is equivalent to its validity for any acceleration; accordingly, for the rest of this paper we will only consider slow algorithms. In this section we show that there are finitely many possibilities for the groups Σ_T ; as a matter of fact, all such groups have index at most 8 in Π .

We recall that the *Schreier graph* $\mathcal{S} = \mathcal{S}(\Sigma, \Pi, \{S, R, F\})$ of $\Sigma \leq \Pi$ w.r.t. the generating set $\{S, R, F\}$ is defined as follows:

- the vertices of \mathcal{S} are the right cosets $\Sigma B \in \Sigma \backslash \Pi$, and there is a distinguished vertex 1 (called the *root*), corresponding to Σ ;
- each edge is directed and labelled by one of S, R, F ;
- there is an S -edge from ΣB to ΣC iff $\Sigma C = \Sigma BS$, and analogously for R - and F -edges.

We'll also need the Schreier graph $\mathcal{S}(\Sigma, \Pi, \{L, N, F\})$ taken w.r.t. the generating set $\{L, N, F\}$. We drop reference to Π and to the generating set whenever possible; if the generating set is relevant we'll write $SR\text{-}\mathcal{S}(\Sigma)$ and $LN\text{-}\mathcal{S}(\Sigma)$.

Schreier graphs are instances of rooted directed connected edge-labelled graphs. A homomorphism $\varphi : \mathcal{G}' \rightarrow \mathcal{G}$ of such graphs is a function that maps vertices to vertices, edges to edges, and preserves all the structure (root, labelling, edge source and target). The homomorphism φ is a *covering* if it is surjective on vertices and edges, and *locally trivial*: for each vertex y of \mathcal{G}' , φ gives a bijection from the set of edges leaving and entering y to the set of edges leaving and entering $\varphi(y)$. If in the above definition we drop the words “and entering” we get the definition of *opfibration* of graphs (the name originates from category theory [4, Definition 2.2], an alternative name being *right-covering* [16, Definition 8.2.1]).

If $\Sigma' \leq \Sigma$ then $\Sigma'B \mapsto \Sigma B$ gives a covering from $\mathcal{S}(\Sigma')$ to $\mathcal{S}(\Sigma)$; in particular, the *Cayley graph* $\mathcal{S}(1)$ covers any Schreier graph for Π . The simple form of the relations involving S, R, F makes the drawing of $SR\text{-}\mathcal{S}(1)$ easy: we represent the “upper part” of it —namely, the Cayley graph of Γ — in Figure 4. The pattern in Figure 4 extends to infinity in all directions, and R -edges are represented as plain arrows. As $S^2 = 1$, S -edges appear in pairs going in opposite directions; we represent such a pair by a single unoriented dashed edge. We'll discuss later the vertices marked by a black circle. We obtain the full Cayley graph of Π by attaching a twin copy of Figure 4 “under the page”. Each upper vertex is connected to its lower twin by a pair of F -edges going in opposite directions, and again represented by a single dotted unoriented edge; S -edges are faithfully copied from the upper level to the lower one. R -edges are copied as well but, due to the relation $FRF = R^{-1}$, the clockwise 3-cycles become counterclockwise ones. Drawing Schreier graphs w.r.t. the generating set $\{L, N, F\}$ is messier, because the arrows tend to intersect; in Example 3.3 and later on we'll use dashed arrows for L -edges and plain arrows

for N -edges, while F -edges remain dotted and unoriented. One easily commutes between the two generating sets via the relations $L = R^2S$, $N = RS$, $S = NL^{-1}N$, $R = LN^{-1}$.

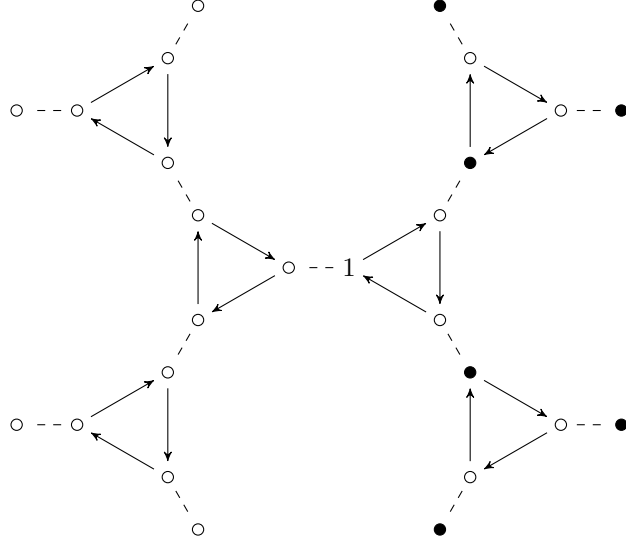


FIGURE 4.

The free monoid \mathcal{M} mentioned in §1(2) determines a proper full subgraph of the LN -Cayley graph of Γ , and this subgraph is an infinite binary tree; the root 1 and the black circles in Figure 4 are precisely the vertices of this tree. The F -twins “under the page” of these vertices correspond to the matrices in Π with nonnegative entries and determinant -1 . They form a twin binary tree, whose L, N labelling is flipped from the upper one, due to the relation $FLF = N$. The complex of these two trees, together with the vertical F -edges connecting each vertex to its twin, constitutes a subgraph \mathcal{TT} of the LN -Cayley graph of Π .

Before returning to our slow algorithms and stating Definition 3.1 we need an observation about unimodular partitions. Again by §1(2), each unimodular partition of $\Delta = [0, \infty]$ can be obtained (not uniquely) from the trivial partition $\{\Delta\}$ in finitely many steps, each step consisting in choosing an interval $\Delta' = B * \Delta$ of the current partition (B being some element of \mathcal{M}), and replacing it with its two *Farey splittings* $\Delta'' = BL * \Delta$ and $\Delta''' = BN * \Delta$. Let now T be a slow algorithm; recall that we specify it by providing a finite set $\{A_0, \dots, A_{n-1}\}$ of matrices in Π with nonnegative entries, which we number according to the left-to-right enumeration of the intervals $\Delta_a = A_a * [0, \infty]$. Such a set is uniquely determined by T , and every A_a is uniquely factorizable as $A_a = B_a F^{e(a)}$, with $B_a \in \mathcal{M}$ and $e(a) \in \{0, 1\}$. If we describe the unimodular partition associated to T in terms of successive Farey splittings as above, then the set of B ’s encountered during the process is precisely the set of left factors in \mathcal{M} of B_0, \dots, B_{n-1} . In particular, given any B in this set of left factors, precisely one of the following holds:

- (a) neither BL nor BN are in the set (this holds iff B is one of B_0, \dots, B_{n-1});
- (b) both BL and BN are in the set.

Definition 3.1. Let T , $A_a = B_a F^{e(a)}$ be as above. We define the *graph of T* , denoted by \mathcal{G}_T , as follows. We start from the double binary tree \mathcal{TT} described in the penultimate paragraph. We delete from \mathcal{TT} all vertices (and all edges incident to them) except those of the form B and BF , where B is a left factor in \mathcal{M} of one of B_0, \dots, B_{n-1} . This leaves us with two copies of a finite binary tree, connected by vertical F -edges. By the previous observation each vertex either has an L -child and an N -child, which are distinct, or is a leaf and has no children; moreover, the set of leaves is precisely $\{B_a : a \in \{0, \dots, n-1\}\} \cup \{B_a F : a \in \{0, \dots, n-1\}\}$. Now, for each a , if $e(a) = 0$ then we glue B_a with the root 1, and $B_a F$ with the vertex F -connected to 1, which we'll always denote 2. Conversely, if $e(a) = 1$ we glue B_a with 2 and $B_a F$ with the root. This leaves us with a connected rooted graph such that precisely one L -, N -, and F -edge stems from every vertex, while the number of L - and N -edges entering in a given vertex may be 0 or greater than 1.

Theorem 3.2. *There exists an opfibration from \mathcal{G}_T to the Schreier graph $LN\text{-}\mathcal{S}(\Sigma_T)$; in particular, since \mathcal{G}_T is finite, Σ_T has finite index in Π .*

Example 3.3. Let T be determined by $A_0 = L^3$, $A_1 = L^2 NF$, $A_2 = LNL$, $A_3 = LN^2 F$, $A_4 = N$. Labelling the vertices for clarity, and omitting the vertical F -edges, we obtain the graph \mathcal{G}_T shown in Figure 5.

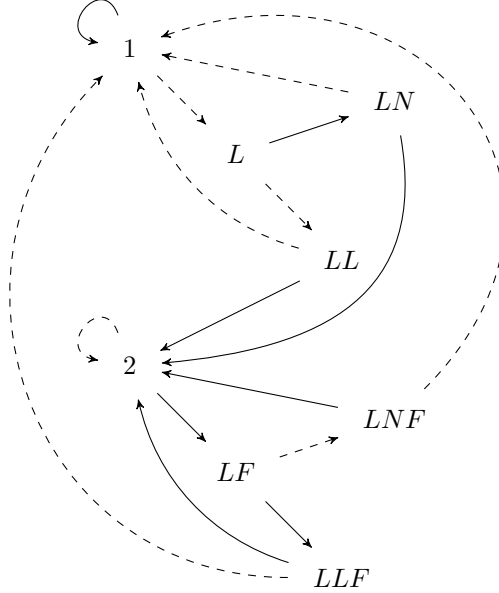


FIGURE 5.

Example 3.4. The minimum number of inverse branches for T is 2, and in this case there are 4 possibilities for \mathcal{G}_T , shown in Figure 6 (if two distinct vertices are connected by a pair of arrows with the same labelling and going in opposite directions, we draw a single shaft without arrowheads). In the first and third case we directly obtain the Schreier graph for Σ_T , that turns out to be Γ (this is

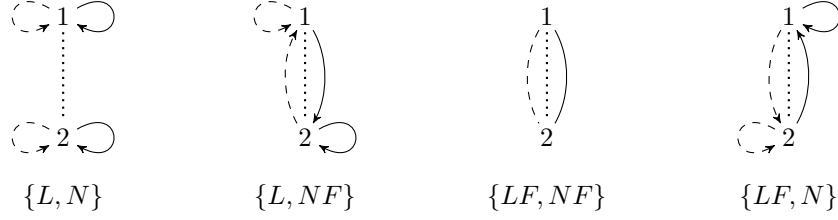


FIGURE 6.

Example 2.5) and $\alpha\Gamma$, respectively. In both cases flipping the two vertices is a graph automorphism, corresponding to the fact that T is invariant under conjugation by F . In the second and fourth case we have a proper opfibration of the trivial Schreier graph, so $\Sigma_T = \Pi$. Flipping the vertices exchanges the second graph with the fourth and indeed, as remarked in Example 2.4, the map in Example 2.3 is F -conjugated to the Farey map.

Proof of Theorem 3.2. By construction, all closed circuits in \mathcal{G}_T starting and ending at the root determine a product of $L^{\pm 1}, N^{\pm 1}, F$ that belongs to Σ_T . If no vertex is the target of two distinct edges with the same labelling, then the fact that every vertex is the source of precisely one edge for each edge label implies that it is also the target of precisely one edge for each label. Therefore \mathcal{G}_T is already a Schreier graph, necessarily —by the previous remark— of Σ_T . If the vertex v is the target of two distinct edges with the same labelling, one originating from u and the other from w , then we glue u and w and take the quotient graph. The new root-based closed circuits originating from this process are still in Σ_T , because the gluing is just a consequence of the cancellation property in groups. We repeat the process, which must eventually terminate, leaving us with the LN -Schreier graph of Σ_T . \square

Continuing with Example 3.3, the vertices LN, LL, LNF, LLF of Figure 5 must all be glued together, because any of them is the source of an L -edge to the root (as well as of an N -edge to F). The resulting vertex is now the target of an N -edge from L , and another one from LF . This forces the gluing of L and LF , and we are left with the Schreier graph in Figure 7. Since the root is the unique vertex carrying an N -loop, $\mathcal{S}(\Sigma_T)$ has a trivial automorphism group; hence Σ_T is an index-4 subgroup of Π that equals its own normalizer.

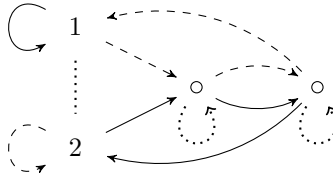


FIGURE 7.

4. INDEX AT MOST 8

Since Π is finitely generated, the following theorem implies that the list of possible Σ_T 's is finite.

Theorem 4.1. *Let T be a slow continued fraction algorithm. Then the group Σ_T generated by the inverse branches of T has index at most 8 in Π . All indices from 1 to 8 are realized by some Σ_T , possibly in non-isomorphic ways.*

Proof. The proof is based on two observations and a lemma.

- (A) Every Σ_T must contain at least one of $SRS, SRSF, SR^2SF$. Indeed, the observation preceding Definition 3.1 implies that at least one pair of consecutive intervals Δ_a, Δ_{a+1} in the unimodular partition associated to T arises from the Farey splitting of the interval $\Delta_a \cup \Delta_{a+1} = B * \Delta$, with $B \in \mathcal{M}$ such that $A_a = BLF^{e(a)}$ and $A_{a+1} = BNF^{e(a+1)}$. Therefore $\Sigma_T \ni A_{a+1}^{-1}A_a = F^{e(a+1)}N^{-1}LF^{e(a)} = F^{e(a+1)}SRSF^{e(a)}$. If $e(a), e(a+1)$ have the same parity, then SRS belongs to Σ_T , while if they have different parity either $SRSF$ or its F -conjugate SR^2SF belongs to Σ_T .
- (B) The Schreier graph of Σ_T cannot contain an infinite path that runs along the positively oriented L - and N -edges and avoids forever both 1 and 2 (which may coincide). Indeed, if such a path existed then it would be liftable to \mathcal{G}_T by the opfibration property established in Theorem 3.2 (as a matter of fact, the liftability of positively oriented paths characterizes opfibrations [4, §2.2]). But, by the very construction of \mathcal{G}_T , such an infinite path cannot exist.

Conjugating by F we assume without loss of generality that our fixed $\Sigma = \Sigma_T$ contains at least one of $SRS, SRSF$. We will establish Theorem 4.1 by showing that $\mathcal{S}(\Sigma)$ cannot have 9 or more vertices. We work with SR -graphs, and we repeat our drawing conventions: S -, R -, and F -edges are, respectively, dashed, plain, and dotted. Each pair of S - or F -edges connecting two distinct vertices and going in opposite directions are drawn as a single edge with no arrowheads.

We remark that whenever (u, v, w) is an R -cycle in $\mathcal{S}(\Sigma)$ (R -cycles are always assumed to be nontrivial, otherwise we speak of R -loops), then u is F -connected to v iff w carries an F -loop. This fact follows from the identity $F = RFR$ and will be used several times.

Lemma 4.2. *Suppose we have constructed a subgraph \mathcal{G} of $\mathcal{S}(\Sigma)$. Let $1, 2, \dots, n-1, n$ be the vertices of \mathcal{G} (still keeping our standard notation of 1, 2 for the root and the vertex F -connected to it). Assume that $n-1$ and n are S -connected (possibly $n-1 = n$, i.e., there is an S -loop at $n-1$), and that \mathcal{G} is almost complete, meaning that the mere addition of an R -loop at n gives a graph \mathcal{G}^* which is the Schreier graph of some subgroup of Π . Then $\mathcal{S}(\Sigma)$ is indeed \mathcal{G}^* .*

Proof. Suppose not. Then $\mathcal{S}(\Sigma)$ must contain vertices $n+1, n+2, n+3, n+4$ (and possibly others), all of them not in \mathcal{G} , with $(n, n+1, n+2)$ forming an R -cycle, $n+1$ S -connected to $n+3$, and $n+2$ S -connected to $n+4$ ($n+3$ and $n+4$ are not necessarily distinct from $n+1$ and $n+2$). But then we can find in $\mathcal{S}(\Sigma)$ an infinite positively oriented LN -path touching only vertices not in \mathcal{G} . Indeed, we start from either $n+3$ or $n+4$ and move along $L = R^2S$ and $N = RS$ edges arbitrarily. Since \mathcal{G} is almost complete, there is no risk of touching a vertex in \mathcal{G} , unless we land at $n+1$ or at $n+2$. If we land at $n+1$ we move to $n+4$ via N , and if we land at $n+2$ we move to $n+3$ via L , and continue our errand. The existence of such a path contradicts (B). \square

We now argue by cases.

Case 1: $R \in \Sigma$. If $SRS \in \Sigma$ then surely the index of Σ is ≤ 4 . Indeed every element of Π factors (uniquely) as a word in S and R , possibly followed by a single occurrence of F . One easily deduces that each right coset must be of the form ΣB , with $B \in \{1, S, F, SF\}$.

If $SRS \notin \Sigma$ then $SRSF \in \Sigma$ and $F \notin \Sigma$. This implies that the Schreier graph of Σ in a neighborhood of the root must be of the form given in Figure 8. Indeed,

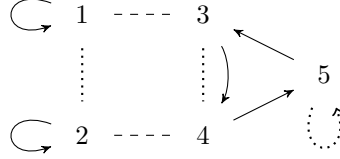


FIGURE 8.

the R -loop in 2 appears by $R = FR^2F$. The two distinct vertices 3, 4 appear since 1 cannot be S -connected with either 1 or 2 (otherwise SRS would belong to Σ). The vertex 3 must be R -connected to 4 since $SRSF \in \Sigma$, and this creates an R -cycle involving a fresh vertex 5. The vertices 3 and 4 must be F -connected because 1 and 2 so are, and $F = SFS$. Finally, the F -loop in 5 appears by the remark preceding Lemma 4.2.

Claim. The only way of completing the graph in Figure 8 to $\mathcal{S}(\Sigma)$ is either by adding an S -loop to 5, or by S -connecting 5 to a new vertex 6 and adding to 6 an R - and an F -loop.

Proof of Claim. Adding an S -loop at 5 completes the graph and we are done. If we do not do so, we are forced to S -connect 5 to a new vertex 6, that must carry an F -loop because so does 5. The resulting graph is almost complete, so our statement follows from Lemma 4.2.

By the claim, Σ must have index 5 or 6, and this concludes the analysis of Case 1.

We assume now $R \notin \Sigma$, so 1 belongs to an R -cycle $(1, 3, 5)$ and we have 4 more cases: $2 = 1$, $2 = 3$, $2 = 5$, and $2 \notin \{1, 3, 5\}$.

Case 2: $2 = 1$, i.e., there is an F -loop at 1. This simply cannot happen. Indeed, by (A) and $F \in \Sigma$ we have $SRS \in \Sigma$. Now, 1 cannot be S -connected to any of 1, 3, 5, as in that case starting from 1 and following the path SRS does not bring us back to 1. Therefore 1 must be S -connected to a new vertex 4, carrying both an F - and an R -loop. The graph \mathcal{G} containing the two S -connected vertices 1, 4, an F -loop at both, and an R -loop at 4, is almost complete. Therefore, by Lemma 4.2, there could not be any R -cycle at 1.

Case 3: $2 = 3$. Then there must be an F -loop at 5. Also, 5 cannot be S -connected to either 1 or 2, since otherwise 1 and 2 would collapse. The vertices 1 and 2 cannot be S -connected, otherwise (A) is violated. Therefore, in a neighborhood of the root the Schreier graph of Σ is of the form in Figure 9; it might be $6 = 1$, which is equivalent to $4 = 2$.

Subcase 3.1: $6 = 1$. Then there is an S -loop both at 1 and at 2. If there is an S -loop at 5 then the graph is complete, while if 5 is S -connected to a new vertex then the latter must carry an F -loop. This leaves us with an almost complete graph and Lemma 4.2 applies. Hence Σ has index 3 or 4.

Subcase 3.2: $6 \neq 1$, and 6 carries an R -loop. Due to $R = FR^2F$, 4 carries an R -loop as well. By the same argument as in Subcase 3.1, Σ has index 5 or 6.

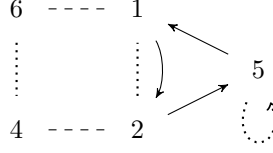


FIGURE 9.

Subcase 3.3: $6 \neq 1$, and 6 carries an R -cycle. By (A), 6 must be R -connected to 4, so we are left with the situation in Figure 10, where possibly $8 = 5$ and/or $9 = 7$.

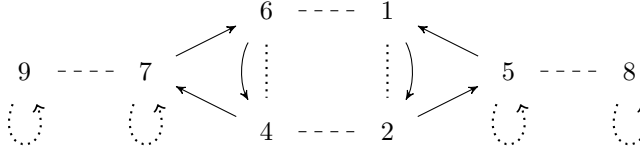


FIGURE 10.

- If $8 = 5$ (i.e., 5 carries an S -loop), then Lemma 4.2 assures us that Σ has index 6 (if $9 = 7$) or 7 (if $9 \neq 7$). Analogously if $9 = 7$.
- If $8 \neq 5$, $9 \neq 7$, and either 8 or 9 carries an R -loop, then the other one must carry an R -loop as well, again by Lemma 4.2, so Σ has index 8.

We claim that there are no other possibilities, i.e., that any Schreier graph \mathcal{S} extending Figure 10 (with $8 \neq 5$ and $9 \neq 7$) and with strictly more vertices contradicts (B). Indeed, by the discussion above, \mathcal{S} should contain two R -cycles $(8, 10, 11)$ and $(9, 12, 13)$. The presence of F -loops at 8 and 9 forces $\{8, 10, 11\} \cap \{9, 12, 13\} = \emptyset$, again by the remark preceding Lemma 4.2. It follows that the set $D = \{\text{vertices of } \mathcal{S}\} \setminus \{1, 2(= 3), 4, 5, 6, 7, 8, 9\}$ contains at least the four distinct vertices 10, 11, 12, 13, as well as the vertices $10', 11', 12', 13'$ S -connected to them. But then we can construct an infinite LN -path all contained in D , just taking care that whenever we land in, say, 10, we apply N and go to $11'$ (while L would lead us to $5 \notin D$). This contradicts (B), establishes our claim and concludes the analysis of Case 3.

Case 4: $2 = 5$. This is analogous to Case 3, except for the possibility, at the beginning of the discussion, that 1 and $2(= 5)$ be S -connected. This is now possible (and yields $SRSF \in \Sigma$), while it was not in Case 2. However, if 1 and 2 are S -connected it is easy to see, by looking at the S -edge leaving 3 and using Lemma 4.2, that Σ must then have index 3 or 4. The rest of the proof is completely analogous to that in Case 3.

Case 5: $2 \notin \{1, 3, 5\}$. Again, this cannot happen. Indeed, one easily sees that the hypotheses yield the existence of another R -cycle $\{2, 4, 6\}$ with F -edges connecting 3 with 6 and 5 with 4. As in Case 2, condition (A) precludes 1 to be S -connected to any of 1, 2, 3, 4, 5, 6, and this forces two new vertices $7 \neq 8$, which are connected by S - and F -edges as in Figure 11. In order to satisfy (A), precisely two cases are possible.

Subcase 5.1: *There is an R -loop at 7.* Since $R = FR^2F$, this forces an R -loop at 8 as well. Then, as in Subcase 3.3, we can construct an infinite LN -path that avoids the set $\{1, 2, 7, 8\}$ forever; this subcase is therefore impossible.

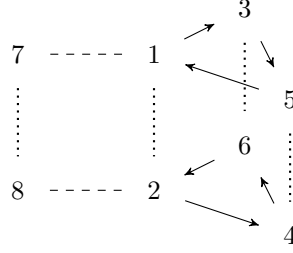


FIGURE 11.

Subcase 5.2: There is an R -edge from 7 to 8. This creates an R -cycle $(7, 8, 9)$, with an F -loop at 9. This F -loop precludes the possibility of S -connecting 9 with any of 3, 4, 5, 6 (because none of them carries an F -loop). There are now precisely three mutually exclusive possibilities, namely

- (i) 9 carries an S -loop;
- (ii) 9 is S -connected to a new vertex 10 that carries both an F - and an R -loop;
- (iii) 9 is S -connected to a new vertex 10 that carries an F -loop and an R -cycle $(10, 11, 12)$.

In each of these cases we can again construct an infinite LN -path, avoiding forever $\{1, 2, 7, 8, 9\}$ (in case (i)), or $\{1, 2, 7, 8, 9, 10\}$ (in cases (ii) and (iii)). This subcase is then impossible as well, and the analysis of Case 5 is completed.

As all indices from 1 to 8 have been realized in several nonisomorphic ways during the previous analysis, Theorem 4.1 is proved. \square

Corollary 4.3. *If all inverse branches of T have positive determinant (i.e., T is increasing on each interval), then Σ_T equals either Γ or its unique index-2 subgroup $\langle R, SRS \rangle$. It equals Γ iff at least one of the matrices A_a factors in \mathcal{M} as an L - N product of odd length.*

Proof. In the proof of Theorem 4.1 we discussed the five possible cases for Σ_T , cases 2 and 5 being void. Case 3 implies $RF \in \Sigma_T$, and Case 4 implies $R^2F \in \Sigma_T$; both are impossible here, since Σ_T cannot contain elements of determinant -1 . So we are left with Case 1; taking into account (A), we conclude that Σ_T extends $\langle R, SRS \rangle$. Expressing L, N in terms of S, R , one sees that $LL, LN, NL, NN \in \langle R, SRS \rangle$, while $L, N \notin \langle R, SRS \rangle$; our statements follow immediately. \square

5. THE TAIL PROPERTY

We briefly recall the definition of a finite-state transducer. Let an *input alphabet* \mathcal{Z} and an *output alphabet* \mathcal{A} be given, both finite. As usual, \mathcal{Z}^* denotes the set of all finite words over \mathcal{Z} (including the empty word), while \mathcal{Z}^ω is the set of all one-sided infinite sequences $\mathbf{Z} = Z_0Z_1Z_2\cdots$. A *finite-state transducer* is a finite directed graph, whose edges are labelled by *transition rules* of the form $Z|w$, with $Z \in \mathcal{Z}$ and $w \in \mathcal{A}^*$; we only consider *deterministic* transducers, i.e., transducers such that, for each vertex and each Z , at most one edge labelled $Z|w$ leaves that vertex (in this context vertices are usually called *states*). Given the input \mathbf{Z} and a vertex v in a given set of initial states, the transducer acts in the obvious way: it

first checks if an edge labelled $Z_0|w_0$ starts from v . If so, it moves to the target vertex, checks if an edge labelled $Z_1|w_1$ starts from it, and goes on. The process stops and fails, producing no output, if a vertex is reached from which no appropriate edge starts; if this never happens the computation succeeds, yielding the output $\mathbf{a} = w_0w_1w_2 \cdots \in \mathcal{A}^\omega$.

For every slow algorithm T , the graph \mathcal{G}_T yields a canonical transducer as follows. We take $\mathcal{Z} = \{L, N\}$ and $\mathcal{A} = \{0, \dots, n-1\}$ as our alphabets; it is expedient to equip \mathcal{Z}^* and \mathcal{Z}^ω with the involution $'$ that exchanges L with N component-wise. We first remove from \mathcal{G}_T all “vertical” F -edges, retaining however the handy notation $vX_0 \cdots X_{r-1}$ for the vertex reached from v by following the edges labelled $X_0, \dots, X_{r-1} \in \{L, N, F\}$, in this order. For $Z \in \mathcal{Z}$, we relabel each Z -edge that does not end either at 1 or at $2 = 1F$ by $Z|$. We now examine the edges terminating at 1 or at 2. Each A_a induces precisely two such edges; namely, as in Definition 3.1, we write uniquely each A_a as $A_a = B_a F^{e(a)} = C_a Z F^{e(a)}$, with $C_a \in \mathcal{M}$. Then, by construction, \mathcal{G}_T contains a Z -edge from C_a to either 1 (if $e(a) = 0$) or 2 (if $e(a) = 1$), as well as a Z' -edge from $C_a F$ to either 2 (if $e(a) = 0$) or 1 (if $e(a) = 1$). We relabel these two edges by $Z|a$ and $Z'|a$, respectively; repeating this relabeling for each A_a , and taking $\{1\}$ as the set of initial states, leaves us with a transducer, again denoted by \mathcal{G}_T . Since for each vertex of \mathcal{G}_T and each Z precisely one edge labelled $Z|w$ starts from that vertex, this transducer is deterministic and succeeds at every input. As an example, Figure 12 shows “one half” (see §6) of the transducer determined by the map T in Example 2.12.

For $\mathbf{a} = a_0a_1 \cdots \in \mathcal{A}^\omega$ and $\alpha \in [0, \infty]$, we write $\sigma(\mathbf{a}) = \alpha$ if \mathbf{a} is a T -symbolic orbit for α , i.e., $T^t(\alpha) \in \Delta_{a_t}$ for every $t \geq 0$. Every irrational number has precisely one symbolic orbit, while every rational has at most two orbits. For the slow map $\{L, N\}$ of Example 2.5, which has a distinguished status, we’ll use $\pi : \mathcal{Z}^\omega \rightarrow [0, \infty]$ instead of σ . By the description in [21], writing $\pi(\mathbf{Z}) = \alpha$ is equivalent to saying that starting from an arbitrary point β in the imaginary axis $\mathbb{R}_{>0}i$ of the hyperbolic upper-half plane and moving along the geodesic arc connecting β to α , the resulting cutting sequence of the Farey tessellation is \mathbf{Z} (note that our N, L are Series’s L, R). We abuse language by writing π also for the map that associates to a word $W \in \mathcal{Z}^*$ the matrix in Γ resulting from reading W as an ordinary matrix product. This abuse is justified by the identity $\pi(W\mathbf{Z}) = \pi(W) * \pi(\mathbf{Z})$, that is immediately proved by induction on the length of W .

Lemma 5.1. *Let $\pi(\mathbf{Z}) = \alpha$. Then the T -symbolic orbit of α is the output of \mathcal{G}_T at \mathbf{Z} .*

Proof. The proof is straightforward, but we use it to introduce a formalism and an alternative description of \mathcal{G}_T that will be used later on. We start at the root of \mathcal{G}_T and follow the path determined by \mathbf{Z} , producing no output but taking note of our visits to 1 and 2 by inserting 1 and 2 as indices along \mathbf{Z} . We call the times $0 = t_0 < t_1 < t_2 < \cdots$ at which these visits take place the *hitting times*. For example, let $\mathbf{Z} = NLLNNLNNL\bar{L}N\bar{L}$, whose π -image is $\alpha = (1335 + \sqrt{3})/939$. Applying the transducer in Figure 5 we obtain the annotated sequence

$${}_1N_1LLN_2NLN_2NNL_1LNL_1LNL_1L \cdots . \quad (*)$$

For $W \in \mathcal{Z}^*$, $\mathbf{Y} \in \mathcal{Z}^\omega$, $i, j \in \{1, 2\}$, we call ${}_iW_j$ and ${}_i\mathbf{Y}$ a *marked word* and a *marked sequence*, respectively. We extend π to marked words and sequences by

defining

$$\begin{aligned}\pi({}_1W_1) &= \pi(W), & \pi({}_1W_2) &= \pi(W)F, \\ \pi({}_2W_1) &= F\pi(W) = \pi(W')F, & \pi({}_2W_2) &= \pi(W'), \\ \pi({}_1\mathbf{Y}) &= \pi(\mathbf{Y}), & \pi({}_2\mathbf{Y}) &= F * \pi(\mathbf{Y}) = 1/\pi(\mathbf{Y}).\end{aligned}$$

These definitions are concocted so that the identity $\pi({}_iW_j) * \pi({}_j\mathbf{Y}) = \pi({}_iW\mathbf{Y})$ holds, as can be easily verified.

Now, each marked word ${}_iW_j$ appearing in a sequence such as $(*)$ corresponds to a path from i to j making no intermediate visits to either 1 or 2; we call such a path *primitive*. Also, by construction, upon completing a primitive path the transducer outputs the symbol a determined by $\pi({}_iW_j) = A_a$. We can then move along the annotated sequence and read the output $\mathbf{a} = a_0a_1\cdots$ by computing π on the successive marked words. For example, the input \mathbf{Z} above yields the output $\mathbf{a} = 4121\bar{2}$, and indeed $\pi({}_1N_1) = N = A_4$, $\pi({}_1LLN_2) = LLNF = A_1$, $\pi({}_2NLN_2) = LNL = A_2$, and so on.

The statement of the lemma follows now by induction on the set of hitting times. Suppose that at the hitting time t_k ($k \geq 0$) the transducer produced the correct output $a_0 \cdots a_{k-1}$, is in state i , and is to read the sequence ${}_iW_j\mathbf{Y}$ whose π -image is $T^k(\alpha)$. Let $\pi({}_iW_j) = A_a$; then $T^k(\alpha) = A_a * \pi({}_j\mathbf{Y}) \in \Delta_a$. It follows that the next symbol in the T -symbolic sequence of α is $a_k = a$, which agrees with the transducer output in its moving from i to j along W . Also, $T^{k+1}(\alpha) = A_a^{-1} * T^k(\alpha) = \pi({}_j\mathbf{Y})$, so that at the next hitting time t_{k+1} the transducer is in state j and is to read the sequence ${}_j\mathbf{Y}$, as required. \square

The map σ^{-1} from the Baire space $\mathbb{R}_{>0} \setminus \mathbb{Q}$ to the Cantor space \mathcal{A}^ω is injective and bicontinuous, hence a homeomorphism from its domain to its image \mathcal{A}^\sharp . Writing Σ for Σ_T , on \mathcal{A}^\sharp we have two equivalence relations:

- \mathbf{a} and \mathbf{b} *have the same tail*, written $\mathbf{a} \equiv_t \mathbf{b}$, if there exist $h, k \geq 0$ such that, for every $l \geq 0$, $a_{h+l} = b_{k+l}$;
- \mathbf{a} and \mathbf{b} *are Σ -equivalent*, written $\mathbf{a} \equiv_\Sigma \mathbf{b}$, if $\sigma(\mathbf{a})$ and $\sigma(\mathbf{b})$ are Σ -equivalent according to §1(5).

As remarked in §1 and before Definition 2.8, \equiv_Σ is coarser than \equiv_t : each Σ -equivalence class is partitioned into tail-equivalence classes, and the Serret theorem holds for T precisely when these two relations agree.

Definition 5.2. In the proof of Lemma 5.1 we noted that the primitive paths in \mathcal{G}_T correspond to marked words ${}_iW_j$ ($i, j \in \{1, 2\}$). If $W = Z_0 \cdots Z_{l-1}$, then the path has *length* l and contains the vertices $i, {}_iZ_0, {}_iZ_0Z_1, \dots, {}_iW = j$, which are all distinct, except for the possibility $i = j$. The F -twins of these vertices constitute another primitive path, namely ${}_iFW'_{jF}$, and $\pi({}_iW_j) = \pi({}_iFW'_{jF}) = A_a$ for some a . Each A_a corresponds as above to precisely these two primitive paths; let D_a be the set of vertices along them, of cardinality $2l$. Let $\varphi^{-1}\{1\}$ be the counterimage of 1 under the opfibration $\varphi : \mathcal{G}_T \rightarrow \mathcal{S}_T$. We define the *defect* of T to be the maximum, say d , of the cardinalities of $D_a \cap \varphi^{-1}\{1\}$, as a varies in $\{0, \dots, n-1\}$.

Theorem 5.3. *Let $\mathbf{a} \in \mathcal{A}^\sharp$. Then the Σ -equivalence class of \mathbf{a} is partitioned in at most d tail-equivalence classes.*

As an immediate corollary, the condition $\varphi^{-1}\{1\} = \{1\}$ is sufficient for the validity of the Serret theorem. This condition holds, e.g., for the pythagorean

map of Example 2.6; by explicit computation one checks that Σ has index 3, while $\varphi^{-1}\{1\}$ is trivial. It is not a necessary condition, as witnessed by the map T' of Example 2.12; we'll provide a complete characterization in Corollary 5.6.

Example 5.4. Let T be defined by

$$\begin{array}{llll} A_0 = LLL, & A_1 = LLN, & A_2 = LNL F, & A_3 = LNN, \\ A_4 = NLL, & A_5 = NLN, & A_6 =>NNL, & A_7 = NNN. \end{array}$$

Computing the opfibration $\varphi : \mathcal{G}_T \rightarrow \mathcal{S}(\Sigma)$ one sees that $\mathcal{S}(\Sigma)$ is trivial, so $\Sigma = \Pi$ and $\varphi^{-1}\{1\} =$ (all vertices of \mathcal{G}_T). Each primitive path in \mathcal{G}_T has length 3, therefore T has defect 6.

Let n_1, n_2, \dots be any sequence in $\mathbb{Z}_{>0}$, let

$$\mathbf{Z} = (LLNNLL)^{n_1} LLL (LLNNLL)^{n_2} LLL \dots,$$

and let $\alpha = \pi(\mathbf{Z})$. Chasing paths along \mathcal{G}_T , one checks that:

α	has T -symbolic orbit	$(14)^{n_1} 0 (14)^{n_2} 0 \dots$,
$F * \alpha$	has T -symbolic orbit	$(63)^{n_1} 7 (63)^{n_2} 7 \dots$,
$L^{-1} * \alpha$	has T -symbolic orbit	$(30)^{n_1} 0 (30)^{n_2} 0 \dots$,
$FL^{-1} * \alpha$	has T -symbolic orbit	$(47)^{n_1} 7 (47)^{n_2} 7 \dots$,
$L^{-2} * \alpha$	has T -symbolic orbit	$(60)^{n_1} 0 (60)^{n_2} 0 \dots$,
$FL^{-2} * \alpha$	has T -symbolic orbit	$(17)^{n_1} 7 (17)^{n_2} 7 \dots$.

Here the Serret theorem fails as badly as possible: there are 2^{\aleph_0} counterexamples, each of them breaking its Σ -equivalence class in the maximum available number of tail-equivalence classes.

The proof of Theorem 5.3 proceeds in three stages. We first note that, having started \mathcal{G}_T on input \mathbf{Z} from the root, we'll arrive infinitely often to a vertex v such that either $\varphi(v) = 1$ or $\varphi(vF) = 1$. Every time we hit such a v , we restart the transducer either from the root (if $\varphi(v) = 1$) or from 2 (if $\varphi(vF) = 1$), feeding it with the remaining input. The resulting outputs are then the T -symbolic orbits of numbers β Σ -equivalent to α . In the second part of the proof we will show that every γ Σ -equivalent to α is tail-equivalent to a β of the form given by the first part. Finally, we will prove that the set of such β 's is partitioned in at most d tail-equivalence classes.

Proof of Theorem 5.3. The first part is easy: let $\mathbf{Z} = W\mathbf{Y}$ be a factorization of \mathbf{Z} such that $\pi({}_1W_i) \in \Sigma$ for a certain $i \in \{1, 2\}$. Then $\beta = \pi({}_i\mathbf{Y}) \equiv_{\Sigma} \alpha$, and the proof of Lemma 5.1 shows that the T -symbolic orbit of β is the output of \mathcal{G}_T when restarted in state i and fed with \mathbf{Y} .

For the second part, let $\gamma = M * \alpha \in [0, \infty]$ for some $M \in \Sigma$, and let $W(t) = Z_0 \dots Z_{t-1}$ be the initial segment of \mathbf{Z} of length t . Then $\pi(W(t)) * [0, \infty]$ is a unimodular interval that shrinks, as t increases, to α , so that $M\pi(W(t)) * [0, \infty]$ is a unimodular interval shrinking to γ . Since $\gamma \in [0, \infty]$, we can choose t (fixed from now on) so large that $[p/q, p'/q'] = M\pi(W(t)) * [0, \infty]$ is a subinterval of $[0, \infty]$; its extrema $p/q, p'/q'$ are the $M\pi(W(t))$ -images of $0, \infty$, in this or the other order according to M being of determinant $+1$ or -1 . By §1(2), there exists a unique $U \in \mathcal{Z}^*$ such that $\pi(U) = \begin{bmatrix} p' & p \\ q' & q \end{bmatrix}$. Therefore $M\pi(W(t)) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ equals

$\pi(U)F$ or $\pi(U)$, according to the positive or negative sign of $\det(M)$; in short, $M\pi(W(t)) = \pi(U)F^e$, with $e = 0$ if $\det(M) = 1$ and $e = 1$ otherwise. As the sequence $\mathbf{X} \in \mathcal{Z}^\omega$ such that $\pi(\mathbf{X}) = \gamma$ is unique, we must have:

- if $e = 0$, then $\mathbf{X} = UZ_t Z_{t+1} \cdots$,
- if $e = 1$, then $\mathbf{X} = UZ'_t Z'_{t+1} \cdots$.

Therefore there exists $q \geq t$ and $i \in \{1, 2\}$ such that, writing \mathcal{L} for the monoid generated by A_0, \dots, A_{n-1} , we have:

- if $e = 0$, then $\pi(1(UZ_t \cdots Z_{q-1})_i) \in \mathcal{L}$ and $\gamma \equiv_t \pi(i(Z_q Z_{q+1} \cdots))$,
- if $e = 1$, then $\pi(1(UZ'_t \cdots Z'_{q-1})_i) \in \mathcal{L}$ and $\gamma \equiv_t \pi(i(Z'_q Z'_{q+1} \cdots)) = \pi(i_F(Z_q Z_{q+1} \cdots))$.

Now:

- if $e = 0$, then

$$\begin{aligned} \pi(1(Z_0 \cdots Z_{q-1})_i) &= \pi(W(t)) \pi(1(Z_t \cdots Z_{q-1})_i) \\ &= M^{-1} \pi(U) \pi(1(Z_t \cdots Z_{q-1})_i) = M^{-1} \pi(1(UZ_t \cdots Z_{q-1})_i) \end{aligned}$$

belongs to Σ ,

- if $e = 1$, then

$$\begin{aligned} \pi(1(Z_0 \cdots Z_{q-1})_{iF}) &= \pi(W(t)) \pi(1(Z_t \cdots Z_{q-1})_{iF}) \\ &= M^{-1} \pi(U) F \pi(1(Z_t \cdots Z_{q-1})_{iF}) = M^{-1} \pi(1(UZ'_t \cdots Z'_{q-1})_i) \end{aligned}$$

belongs to Σ too.

We conclude that γ is tail-equivalent to a β of the form given in the first part of the proof, namely $\beta = \pi(i(Z_q Z_{q+1} \cdots))$ if $e = 0$, and $\beta = \pi(i_F(Z_q Z_{q+1} \cdots))$ if $e = 1$.

We finally show that any set of $d+1$ β 's obtained as above must contain two tail-equivalent elements. Let us say that this set has been obtained by stopping \mathcal{G}_T on input \mathbf{Z} at times $t_0 < t_1 < \cdots < t_d$, the transducer being in state v_0, \dots, v_d , respectively. For each $0 \leq r \leq d$, let $i_r := 1$ if $\varphi(v_r) = 1$ and $i_r := 2$ if $\varphi(v_r F) = 1$. We refer to the run of \mathcal{G}_T when restarted from state i_r and fed with $Z_{t_r} Z_{t_r+1} \cdots$ as the r th run. Let m_r be the limsup of the lengths of the primitive paths occurring during the r th run; without loss of generality $m = m_0 \geq m_1, \dots, m_d$. Denoting the vertex $1Z_0 \cdots Z_{q-1}$ by $v(q)$, we choose some $q_0 > t_d$ such that $\{v(q_0), v(q_0 + 1), \dots, v(q_0 + m)\}$ is a primitive path of length m for the 0th run, while for every other run it is a path containing 1 or 2 at least once; this choice is possible due to our assumptions on m . Let

$$f : \{q_0, \dots, q_0 + m - 1\} \times \{1, 2\} \rightarrow (\text{vertices of } \mathcal{G}_T)$$

be defined by

$$f(q, j) = \begin{cases} v(q), & \text{if } j = 1; \\ v(q)F, & \text{if } j = 2. \end{cases}$$

Then f is injective, and its image is the set D of all vertices and their F -twins in the primitive path of length m for the 0th run referred to above. Note that the two final vertices $\{v(q_0 + m), v(q_0 + m)F\}$ are not missing from D , since they appear as the initial ones $\{v(q_0), v(q_0)F\}$. Due to the assumptions on q_0 , for each $0 \leq r \leq d$ there exists a pair $(q_r, j_r) \in \{q_0, \dots, q_0 + m - 1\} \times \{1, 2\}$ such that $j_r = i_r Z_{t_r} \cdots Z_{q_r-1}$. We claim that $f(q_r, j_r) \in \varphi^{-1}\{1\}$ for every r . Indeed, $\pi(1(Z_0 \cdots Z_{t_r-1})_{i_r}) \in \Sigma$ and $\pi(i_r(Z_{t_r} \cdots Z_{q_r-1})_{j_r}) \in \mathcal{L}$, so that $\pi(1(Z_0 \cdots Z_{q_r-1})_{j_r}) \in \Sigma$, as claimed. Since

the cardinality of $D \cap \varphi^{-1}\{1\}$ is less than or equal to the defect d of \mathcal{G}_T , by the pigeonhole principle there must be $r \neq s \in \{0, \dots, d\}$ such that $(q_r, j_r) = (q_s, j_s)$. This implies that the r th and the s th run yield tail-equivalent outputs. \square

Theorem 5.3 (or, rather, its proof) yields a characterization of the T 's satisfying the tail property. It is expedient to introduce another transducer, \mathcal{F}_T , whose set of states is $\varphi^{-1}\{1\} = \{1 = P_1, P_2, \dots, P_m\} \subseteq \{\text{states of } \mathcal{G}_T\}$. By the construction in Definition 3.1, each P_r is of the form B or BF , where B is a left factor of one of A_0, \dots, A_{n-1} . For each P_r and each A_a we have a unique commutator relation $P_r A_a = A_{b_1} \cdots A_{b_q} P_s$ (the product $A_{b_1} \cdots A_{b_q}$ may be empty), and we add to \mathcal{F}_T an edge from P_r to P_s labelled with the transition rule $a|b_1 \cdots b_q$.

Lemma 5.5. *Let α have T -symbolic orbit \mathbf{a} . Then the T -symbolic orbit of $P_r * \alpha$ is the output of \mathcal{F}_T , when starting from P_r on input \mathbf{a} .*

Proof. Straightforward, using the facts that $\alpha = \lim_{t \rightarrow \infty} A_{a_0} A_{a_1} \cdots A_{a_{t-1}} * \infty$ and that $P_r * -$ is continuous. \square

We save space by removing from \mathcal{F}_T the state P_1 and all edges entering it; since the edges leaving P_1 are loops labelled $a|a$, they are automatically removed. Call \mathcal{F}_T^* the resulting graph, which may be empty if $\varphi^{-1}\{1\} = \{1\}$; this is the trivial case cited after the statement of Theorem 5.3.

Corollary 5.6. *Let T, \mathcal{F}_T^* be as above. Then the Serret theorem holds for T iff, for every input sequence $\mathbf{a} \in \mathcal{A}^\omega$ and every state P_r of \mathcal{F}_T^* ,*

- either \mathcal{F}_T^* eventually stops;
- or \mathcal{F}_T^* runs forever, producing an output \mathbf{b} tail-equivalent to \mathbf{a} .

Proof. The left-to-right direction is clear from Lemma 5.5, noting that the stopping of \mathcal{F}_T^* amounts to \mathcal{F}_T entering state P_1 . For the reverse direction, we assume that the Serret theorem fails for T and construct α and P_r such that α and $\beta := P_r * \alpha$ have different T -tails (this implies $P_r \neq P_1$, so that P_r is a state of \mathcal{F}_T^*). By the proof of Theorem 5.3, there exist an irrational number $\gamma = \pi(\mathbf{Z})$, an integer $q > 0$, and an exponent $e \in \{0, 1\}$ such that:

- $\varphi(1Z_0 \cdots Z_{q-1}F^e) = 1$;
- the tail of $\alpha := \pi({}_1 F^e Z_q Z_{q+1} \cdots)$ is different from the tail of γ .

Let $t \geq 0$ be the greatest integer $< q$ such that $i := 1Z_0 \cdots Z_{t-1}$ belongs to $\{1, 2\}$. Then $\pi({}_i(Z_t \cdots Z_{q-1})_1 F^e) \in \Sigma$, because $\pi({}_1(Z_0 \cdots Z_{q-1})_1 F^e) \in \Sigma$ and $\pi({}_1(Z_0 \cdots Z_{t-1})_i) \in \mathcal{L}$. Therefore the state $iZ_t \cdots Z_{q-1}F^e$ belongs to $\varphi^{-1}\{1\}$; say it is equal to P_r . Now, $\beta := P_r * \alpha = \pi({}_i Z_t Z_{t+1} \cdots)$ has the same tail as γ (because $\pi({}_1(Z_0 \cdots Z_{t-1})_i) \in \mathcal{L}$), and thus has tail different from α . \square

It is easy to check that the graph $\mathcal{F}_{T'}^*$, determined by the algorithm T' of Example 2.12 contains just two states, $P_2 = N$ and $P_3 = NN$, each carrying a loop labelled $3|3$. The condition in Corollary 5.6 clearly holds, thus T' has the tail property.

6. SYNCHRONIZING WORDS

Some algorithms fail the tail property in a very fragile way, in the sense that for Lebesgue almost every input α the tail- and Σ -equivalence classes of $\sigma^{-1}(\alpha)$ coincide. This is surely the case when the graph \mathcal{G}_T is *synchronizing*, i.e., admits

a synchronizing word. A *synchronizing word* for a deterministic transducer is an input word W that resets the transducer, that is, leaves it in the same state, no matter which state we started with: $vW = uW$ for every two states v, u [23].

Theorem 6.1. *Let \mathcal{G}_T^1 be the connected component of 1 in \mathcal{G}_T ; it equals all of \mathcal{G}_T iff $\Sigma_T \not\leq \Gamma$. Assume that \mathcal{G}_T^1 is synchronizing. Then the set of $\alpha \in \mathbb{R}_{>0} \setminus \mathbb{Q}$ whose Σ_T -equivalence class is partitioned in more than one tail-equivalence classes has Lebesgue measure 0.*

Proof. Let $R : \mathbb{R}_{>0} \setminus \mathbb{Q} \rightarrow \mathbb{R}_{>0} \setminus \mathbb{Q}$ be the slow map of Example 2.5, explicitly given by

$$R(x) = \begin{cases} x/(1-x), & \text{if } x \in \Delta_0 = [0, 1]; \\ x-1, & \text{if } x \in \Delta_1 = [1, \infty]. \end{cases}$$

Then R preserves the σ -finite, infinite measure $d\mu = dx/x$, and is conservative and ergodic w.r.t. it (see, e.g., [12]). Let $W = Z_0 \cdots Z_{s-1}$ be a synchronizing word for \mathcal{G}_T^1 , let $a(t)$ be 0 or 1 according whether Z_t is L or N , and let $B = \bigcap_{t=0}^{s-1} R^{-t} \Delta_{a(t)}$. The conservativity and ergodicity of R easily imply (this is really a version of the Poincaré recurrence theorem) that μ -all points enter any set of positive μ -measure infinitely often. Since:

- (i) $\mu(B) > 0$,
- (ii) μ and the Lebesgue measure have the same nullsets,
- (iii) $\pi : (\mathcal{Z}^\omega, \text{shift}) \rightarrow (\mathbb{R}_{>0} \setminus \mathbb{Q}, R)$ is a measurable conjugacy,

we conclude that for Lebesgue-all $\alpha = \pi(\mathbf{Z})$ the input \mathbf{Z} to \mathcal{G}_T contains W infinitely often. Our statement then follows from the proof of Theorem 5.3. \square

We refer to [13, Corollary 4.3] for a result in the same vein, albeit stated in a different context and proved with different means.

Theorem 6.1 applies, e.g., to the map T in Example 2.12, whose \mathcal{G}_T^1 is in Figure 12. The word LL is synchronizing for this graph, resetting it to state 1. Note

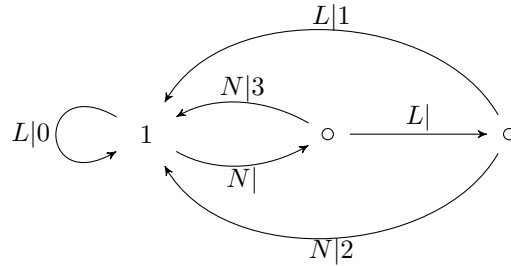


FIGURE 12.

that the exceptional points $\alpha = \sqrt{3}$, $\beta = \sqrt{3} + 1$ are the π -images of the sequences NLN , NNL , which avoid LL , as well as any other synchronizing word.

Again, Example 5.4 provides a much more robust counterexample to the Serret theorem. Indeed, explicit computation (which is not trivial, since \mathcal{G}_T has 14 vertices, so that the synchronizability criterion in [23, Proposition 1] involves a graph with 105 vertices) shows that the associated transducer does not admit any synchronizing word.

REFERENCES

- [1] P. Arnoux and T. A. Schmidt. Cross sections for geodesic flows and α -continued fractions. *Nonlinearity*, 26(3):711–726, 2013.
- [2] V. Baladi and B. Vallée. Euclidean algorithms are Gaussian. *J. Number Theory*, 110(2):331–386, 2005.
- [3] F. P. Boca and C. Linden. On Minkowski type question mark functions associated with even or odd continued fractions. <https://arxiv.org/abs/1705.01238>, 2017.
- [4] P. Boldi and S. Vigna. Fibrations of graphs. *Discrete Math.*, 243(1-3):21–66, 2002.
- [5] J. Borwein, A. van der Poorten, J. Shallit, and W. Zudilin. *Neverending fractions*, volume 23 of *Australian Mathematical Society Lecture Series*. Cambridge University Press, 2014.
- [6] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1982.
- [7] M. Einsiedler and T. Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer, 2011.
- [8] D. J. Grabiner and J. C. Lagarias. Cutting sequences for geodesic flow on the modular surface and continued fractions. *Monatsh. Math.*, 133(4):295–339, 2001.
- [9] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, 5th edition, 1985.
- [10] B. Heersink. An effective estimate for the Lebesgue measure of preimages of iterates of the Farey map. *Adv. Math.*, 291:621–634, 2016.
- [11] G. Iommi. Multifractal analysis of the Lyapunov exponent for the backward continued fraction map. *Ergodic Theory Dynam. Systems*, 30(1):211–232, 2010.
- [12] S. Isola. From infinite ergodic theory to number theory (and possibly back). *Chaos Solitons Fractals*, 44(7):467–479, 2011.
- [13] S. Katok and I. Ugarcovici. Applications of (a, b) -continued fraction transformations. *Ergodic Theory Dynam. Systems*, 32(2):755–777, 2012.
- [14] W. J. LeVeque. *Fundamentals of number theory*. Dover, 1996. Reprint of the 1977 original.
- [15] P. Liardet and P. Stambul. Algebraic computations with continued fractions. *J. Number Theory*, 73(1):92–121, 1998.
- [16] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, 1995.
- [17] H. Nakada. Metrical theory for a class of continued fraction transformations and their natural extensions. *Tokyo J. Math.*, 4(2):399–426, 1981.
- [18] C. G. Pinner. More on inhomogeneous Diophantine approximation. *J. Théor. Nombres Bordeaux*, 13(2):539–557, 2001.
- [19] G. N. Raney. On continued fractions and finite automata. *Math. Ann.*, 206:265–283, 1973.
- [20] D. Romik. The dynamics of Pythagorean triples. *Trans. Amer. Math. Soc.*, 360(11):6045–6064, 2008.
- [21] C. Series. The modular surface and continued fractions. *J. London Math. Soc. (2)*, 31(1):69–80, 1985.
- [22] J. A. Serret. *Cours d’algèbre supérieure*. Gauthier-Villars, 3rd edition, 1866.
- [23] M. V. Volkov. Synchronizing automata and the Černý conjecture. In Martn-Vide C., Otto F., and Fernau H., editors, *Language and Automata Theory and Applications*, volume 5196 of *Lecture Notes in Computer Science*, pages 11–27. Springer, 2008.
- [24] D. Zagier. Nombres de classes et fractions continues. *Astérisque*, 24-25:81–97, 1975.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UDINE, VIA DELLE SCIENZE 206, 33100 UDINE, ITALY

E-mail address: giovanni.panti@uniud.it