The entropy of α-continued fractions: numerical results

2010 Nonlinearity 23 2429

(http://iopscience.iop.org/0951-7715/23/10/005)

Nonlinearity **23** (2010) 2429–2456 doi:10.1088/0951-7715/23/10/005

The entropy of α-continued fractions: numerical results

Carlo Carminati[1], Stefano Marmi[2], Alessandro Profeti[2] and Giulio Tiozzo[3]

[1] Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
[2] Scuola Normale Superiore, Piazza dei Cavalieri 7, 56123 Pisa, Italy
[3] Department of Mathematics, Harvard University, 1 Oxford St, Cambridge MA 02138, USA

E-mail: carminat@dm.unipi.it, s.marmi@sns.it, a.profeti@sns.it, tiozzo@math.harvard.edu

Received 8 January 2010, in final form 12 July 2010
Published 11 August 2010
Online at stacks.iop.org/Non/23/2429

Recommended by K M Khanin

Abstract
We consider the one-parameter family of interval maps arising from generalized continued fraction expansions known as α-continued fractions. For such maps, we perform a numerical study of the behaviour of metric entropy as a function of the parameter. The behaviour of entropy is known to be quite regular for parameters for which a *matching condition* on the orbits of the endpoints holds. We give a detailed description of the set $\mathcal{M}$ where this condition is met: it consists of a countable union of open intervals, corresponding to different combinatorial data, which appear to be arranged in a hierarchical structure. Our experimental data suggest that the complement of $\mathcal{M}$ is a proper subset of the set of bounded-type numbers, hence it has measure zero. Furthermore, we give evidence that the entropy on matching intervals is smooth; on the other hand, we can construct points outside of $\mathcal{M}$ on which it is not even locally monotone.

Mathematics Subject Classification: 11K50, 37A10, 37A35, 37E05

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Let $\alpha \in [0, 1]$. We will study the one-parameter family of one-dimensional maps of the interval

$$T_\alpha : [\alpha - 1, \alpha] \to [\alpha - 1, \alpha],$$

$$T_\alpha(x) = \begin{cases} \dfrac{1}{|x|} - \left\lfloor \dfrac{1}{|x|} + 1 - \alpha \right\rfloor & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

0951-7715/10/102429+28\$30.00

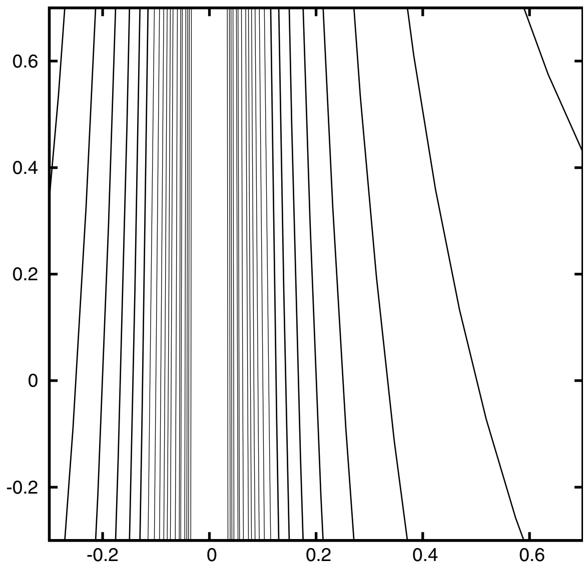

Figure 1. Graph of T_α for $\alpha = 0.7$.

A graph of T_α is shown in figure 1. If we let

$$x_{n,\alpha} = T_\alpha^n(x), \qquad a_{n,\alpha} = \left\lfloor \frac{1}{|x_{n-1,\alpha}|} + 1 - \alpha \right\rfloor, \qquad \epsilon_{n,\alpha} = \mathrm{Sign}\ (x_{n-1,\alpha}),$$

then for every $x \in [\alpha - 1, \alpha]$ we get the expansion

$$x = \cfrac{\epsilon_{1,\alpha}}{a_{1,\alpha} + \cfrac{\epsilon_{2,\alpha}}{a_{2,\alpha} + \ddots}}$$

with $a_{i,\alpha} \in \mathbb{N}, \epsilon_{i,\alpha} \in \{\pm 1\}$ which we call the *α-continued fraction* of x. These systems were introduced by Nakada [9] and are also known in the literature as *Japanese continued fractions* [13].

The algorithm, analogous to the Gauss map in the classical case (the case $\alpha = 1$), provides rational approximations of real numbers. The convergents $\frac{p_{n,\alpha}}{q_{n,\alpha}}$ are given by

$$\begin{cases} p_{-1,\alpha} = 1 & p_{0,\alpha} = 0 \quad p_{n+1,\alpha} = \epsilon_{n+1,\alpha} p_{n-1,\alpha} + a_{n+1,\alpha} p_{n,\alpha}, \\ q_{-1,\alpha} = 0 & q_{0,\alpha} = 1 \quad q_{n+1,\alpha} = \epsilon_{n+1,\alpha} q_{n-1,\alpha} + a_{n+1,\alpha} q_{n,\alpha}. \end{cases}$$

It is known (see [7]) that for each $\alpha \in (0, 1]$ there exists a unique invariant measure $\mu_\alpha(\mathrm{d}x) = \rho_\alpha(x)\mathrm{d}x$ absolutely continuous w.r.t. Lebesgue measure, even though no explicit form of the density is known for $\alpha < \sqrt{2} - 1$.

In this paper we will focus on the metric entropy of the T_αs, which by Rohlin's formula [11] is given by

$$h(T_\alpha) = -2 \int_{\alpha-1}^{\alpha} \log|x| \rho_\alpha(x)\, \mathrm{d}x.$$

Equivalently, entropy can be thought of as the average exponential growth rate of the denominators of convergents: for μ_α-a.e. $x \in [\alpha - 1, \alpha]$,

$$h(T_\alpha) = 2 \lim_{n\to\infty} \frac{1}{n} \log q_{n,\alpha}(x).$$

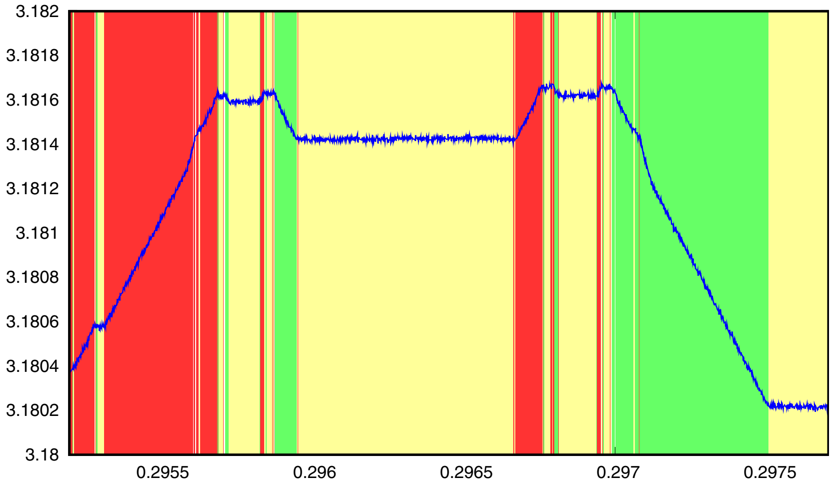

Figure 2. Numerical data versus matching intervals.

The exact value of $h(T_\alpha)$ has been computed for $\alpha \geqslant \frac{1}{2}$ by Nakada [9] and for $\sqrt{2}-1 \leqslant \alpha \leqslant \frac{1}{2}$ by Cassa *et al* [8]:

$$h(T_\alpha) = \begin{cases} \dfrac{\pi^2}{6\log(1+\alpha)} & \text{for } \dfrac{\sqrt{5}-1}{2} < \alpha \leqslant 1, \\ \dfrac{\pi^2}{6\log\dfrac{\sqrt{5}+1}{2}} & \text{for } \sqrt{2}-1 \leqslant \alpha \leqslant \dfrac{\sqrt{5}-1}{2}. \end{cases}$$

In [7], Luzzi and Marmi computed numerically the entropy for $\alpha \leqslant \sqrt{2}-1$ by approximating the integral in Rohlin's formula with Birkhoff averages

$$h(\alpha, N, x) = -\frac{2}{N}\sum_{j=0}^{N-1} \log|T_\alpha^j(x)|$$

for a large number M of starting points $x \in (\alpha-1, \alpha)$ and then averaging over the samples:

$$h(\alpha, N, M) = \frac{1}{M}\sum_{k=1}^{M} h(\alpha, N, x_k).$$

Their computations reveal a rich structure for the behaviour of the entropy as a function of α; it seems that the function $\alpha \mapsto h(T_\alpha)$ is piecewise regular and changes monotonicity on different intervals of regularity.

These features have been confirmed by some results by Nakada and Natsui [10, theorem 2] yielding a *matching condition* on the orbits of α and $\alpha - 1$;

$$T_\alpha^{k_1}(\alpha) = T_\alpha^{k_2}(\alpha-1) \qquad \text{for some } k_1, k_2 \in \mathbb{N}$$

which allows one to find countable families of intervals where the entropy is increasing, decreasing or constant (see section 3). It is not difficult to check that the numerical data computed *via* the Birkhoff theorem fit extremely well with the matching intervals of [10] (see figure 2).

In this paper we will study the matching condition in great detail. First of all, we analyse the mechanism which produces it from a group-theoretical point of view and find an algorithm to relate the α-continued fraction expansion of α and $\alpha - 1$ when a matching occurs. This allows us to understand the combinatorics behind the matchings in a uniform way, without having to resort to specific matrix identities. As an example, we will explicitly construct a family of matching intervals which accumulate on a point different from 0. In fact we also have numerical evidence that there exist positive values, such as $[0; 3, \overline{1}]$, which are cluster points for intervals of all the three matching types: with $k_1 < k_2$, $k_1 = k_2$ and $k_1 > k_2$.

We then describe an algorithm to produce a huge quantity of matching intervals, whose exact endpoints can be found computationally, and we analyse the data thus obtained. These data show that matching intervals are organized in a hierarchical structure, and we will describe a *bisection algorithm* which produces such a structure.

Let now $\mathcal{M}$ be the union of all matching intervals. Nakada and Natsui conjectured [10, section 4, p 1213] that $\mathcal{M}$ is an open, dense set of full Lebesgue measure. In fact, the correctness of our bisection scheme implies the following stronger

Claim 1.1. *For any n, all elements of* $(\frac{1}{n+1}, \frac{1}{n}] \setminus \mathcal{M}$ *have regular continued fraction expansion with partial quotients bounded by* n.

Since the set of numbers with bounded partial quotients has Lebesgue measure zero, this clearly implies the conjecture of Nakada and Natsui[4]

We will then discuss some consequences of these matchings on the shape of the entropy function, coming from a formula in [10]. This formula allows us to recover the behaviour of entropy in a neighbourhood of points where a matching condition is present. First of all, we will use it to prove that entropy has one-sided derivatives at every point belonging to some matching interval, and also to recover the exact value of $h(T_\alpha)$ for $\alpha \geqslant 2/5$. In general, though, to reconstruct the entropy one also has to know the invariant density at one point.

As an example, we shall examine the entropy on an interval J on which (by previous experiments, see [7, section 3] it was thought to be linearly increasing: we numerically compute the invariant density for a single value of $\alpha \in J$ and use it to predict the analytical form of the entropy on J, which in fact happens to be not linear. The data produced with this extrapolation method agree with high precision, and much better than any linear fit, with the values of $h(T_\alpha)$ computed via Birkhoff averages.

The paper is structured as follows: in section 2 we will discuss numerical simulations of the entropy and provide some theoretical framework to justify the results; in section 3 we shall analyse the mechanisms which produce the matching intervals and in section 4 we will numerically produce them and study their hierarchical structure; in section 5 we will see how these matching conditions affect the entropy function.

2. Numerical computation of the entropy

Let us examine more closely the algorithm used in [7] to compute the entropy. A numerical problem in evaluating Birkhoff averages arises from the fact that the orbit of a point can fall very close to the origin: the computer will not distinguish a very small value from zero. In this case we neglect this point, and complete the (pseudo)orbit restarting from a new random

[4] Claim 1.1 has now been proved by two of the authors (see [2]).

seed[5]. As a matter of fact this algorithm produces an approximate value of

$$h_\epsilon(\alpha) := \int_{I_\alpha} f_\epsilon(x)\, \mathrm{d}\mu_\alpha(x) \quad \text{with} \quad f_\epsilon(x) := \begin{cases} 0 & |x| \leqslant \epsilon, \\ -2\log|x| & |x| > \epsilon, \end{cases}$$

where $\epsilon = 10^{-16}$; of course $h_\epsilon(\alpha)$ is an excellent approximation of the entropy $h(\alpha)$, since the difference is of order $\epsilon \log \epsilon^{-1}$. To calculate $h_\epsilon(\alpha)$ we use the Birkhoff sums

$$h_\epsilon(\alpha, N, x) := \frac{1}{N} \sum_{j=0}^{N-1} f_\epsilon(T_\alpha^j(x)) \tag{1}$$

and in [14] the fourth author proves that for large N the random variable $h_\epsilon(\alpha, N, \cdot)$ is distributed around its mean $h_\epsilon(\alpha)$ approximately with normal law and standard deviation $\sigma_\epsilon(\alpha)/\sqrt{N}$ where

$$\sigma_\epsilon^2(\alpha) := \lim_{n\to+\infty} \int_{I_\alpha} \left(\frac{S_n f_\epsilon - n \int f_\epsilon \, \mathrm{d}\mu_\alpha}{\sqrt{n}} \right)^2 \mathrm{d}\mu_\alpha$$

as numerically observed by Luzzi and Marmi (see [7, figure 3]).

One of our goals is to study the function $\alpha \mapsto \sigma_\epsilon^2(\alpha)$, in particular we ask whether it displays some regularity like continuity or semicontinuity. To this aim we pushed the same scheme as in [7] to get higher precision:

1. We take a sample of values α chosen in a particular subinterval $J \subset [0, 1]$.
2. For each value α we choose a random sample $\{x_1, \ldots, x_M\}$ in I_α (the cardinality M of this sample is usually 10^6 or 10^7).
3. For each $x_i \in I_\alpha$ ($i = 1, \ldots, M$) we evaluate $h_\epsilon(\alpha, N, x_i)$ as given in (1) (the number of iterates N will be 10^4).
4. Finally, we evaluate the (approximate) entropy and determine the standard deviation as well:

$$\hat{h}_\epsilon(\alpha, N, M) := \frac{1}{M} \sum_{i=1}^{M} h_\epsilon(\alpha, N, x_i),$$

$$\hat{\sigma}_\epsilon(\alpha) := \sqrt{\frac{1}{M} \sum_{i=1}^{M} [h_\epsilon(\alpha, N, x_i) - \hat{h}_\epsilon(\alpha, N, M)]^2}.$$

2.1. *Central limit theorem*

Let us recall some convergence results for Birkhoff sums of α-transformations (see [14]). Let us denote by $BV(I_\alpha)$ the space of real-valued, μ_α-integrable, bounded variation functions of the interval I_α. We will denote by $S_n f$ the Birkhoff sum

$$S_n f = \sum_{j=0}^{n-1} f \circ T_\alpha^j.$$

[5] Another choice is to throw away the whole orbit and restart; it seems there is not much difference on the final result.

Lemma 2.1 (see [1]). *Let $\alpha \in (0, 1]$ and f be an element of $BV(I_\alpha)$. Then the sequence*

$$M_n = \int_{I_\alpha} \left(\frac{S_n f - n \int f \, \mathrm{d}\mu_\alpha}{\sqrt{n}} \right)^2 \mathrm{d}\mu_\alpha$$

converges to a real non-negative value, which will be denoted by σ^2. Moreover, $\sigma^2 = 0$ if and only if there exists $u \in L^2(\mu_\alpha)$ such that $u\rho_\alpha \in BV(I_\alpha)$ and

$$f - \int_{I_\alpha} f \, \mathrm{d}\mu_\alpha = u - u \circ T_\alpha. \tag{2}$$

The condition given by (2) is the same as in the proof of the central limit theorem for the Gauss map, and it is known as the *cohomological equation*. The main result is the following [14, theorem 2.1]:

Theorem 2.2. *Let $\alpha \in (0, 1]$ and f be an element of $BV(I_\alpha)$ such that (2) has no solutions. Then, for every $v \in \mathbb{R}$ we have*

$$\lim_{n\to\infty} \mu_\alpha \left(\frac{S_n f - n \int_I f \, \mathrm{d}\mu_\alpha}{\sigma \sqrt{n}} \leqslant v \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} \mathrm{e}^{-x^2/2} \, \mathrm{d}x.$$

By using the fact that the invariant density ρ_α is bounded from below by a positive constant [16], one can show [14, proposition 2.6]:

Proposition 2.3. *For every real-valued non-constant $f \in BV(I_\alpha)$, equation (2) has no solutions. Hence, the central limit theorem holds.*

Now, for every $\epsilon > 0$ the function f_ϵ defined in the previous section is of bounded variation, hence the central limit theorem holds and the distribution of the approximate entropy $h_\epsilon(\alpha, N, \cdot)$ approaches a Gaussian when $N \to \infty$.

As a corollary, for the standard deviation of Birkhoff averages

$$\mathrm{Std}\left[\frac{S_n f_\epsilon}{n}\right] = \mathbb{E}\left[\left(\frac{S_n f_\epsilon}{n} - \int_{I_\alpha} f_\epsilon \, \mathrm{d}\mu_\alpha\right)^2\right]^{1/2} = \frac{\sigma}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right).$$

Theorem 2.2 follows from a spectral decomposition method as presented in [1]. For the use of bounded variation techniques in the treatment of piecewise expanding maps, see also [12, 15].

2.2. *Speed of convergence*

In terms of numerical simulations it is of primary importance to estimate the difference between the sum computed at the nth step and the asymptotic value: a semi-explicit bound is given by the following

Theorem 2.4. *For every non-constant real-valued $f \in BV(I_\alpha)$, there exists $C > 0$ such that*

$$\sup_{v\in\mathbb{R}} \left[\mu_\alpha \left(\frac{S_n f - n \int_{I_\alpha} f \, \mathrm{d}\mu_\alpha}{\sigma \sqrt{n}} \leqslant v \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x \right] \leqslant \frac{C}{\sqrt{n}}.$$

Proof. It follows from a Berry–Esséen type of inequality. For details see [1, theorem 8.1]. □

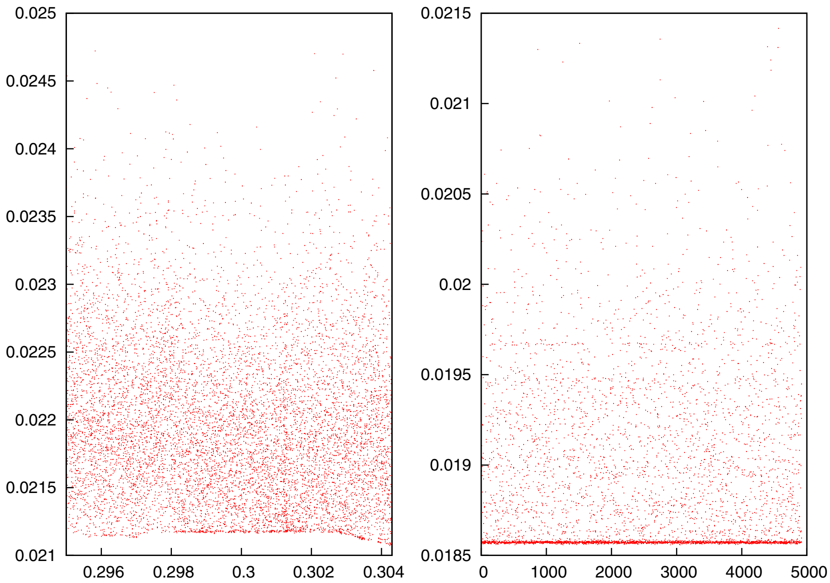

Figure 3. Standard deviation relative to the interval $J = [0.295, 0.3043]$ (left); standard deviation of the different runs on the Gauss map (right).

2.3. Dependence of standard deviation on α

Given these convergence results for the entropy, it is natural to ask how the standard deviation varies with α. In this case not a single exact value of $\sigma_\epsilon(\alpha)$ is known; using the fact that natural extensions of T_α are conjugate [5, 10], it is straightforward to prove the following:

Lemma 2.5. *The map $\alpha \mapsto \sigma(\alpha)$ is constant for $\alpha \in [\sqrt{2} - 1, \frac{\sqrt{5}-1}{2}]$.*

Proof. See proposition A.2 in the appendix. □

The numerical study of this quantity is pretty interesting. We first considered the window $J = [0.295, 0.3043]$, where the entropy is non-monotone. On this interval the standard deviation shows quite a strange behaviour: the values we have recorded do not form a cloud clustering around a continuous line (like for the entropy) but they cluster all above it (see figure 3, left).

One might guess that this is due to the fact that the map $\alpha \mapsto \sigma(\alpha)$ is only semicontinuous, but the same kind of asymmetry takes place also on the interval $J = [0.616, 0.618]$, where σ^2 is constant. Indeed, we can observe the same phenomenon also while evaluating $\hat{\sigma}_\epsilon(\alpha)$ for a fixed value α but taking several different sample sets (see figure 3, right).

On the other hand this strange behaviour cannot be detected for other maps, like the logistic map, and could yet not be explained. Nevertheless, we point out that if you only consider C^1 observables, the standard deviation of Birkhoff sums can be proved to be continuous, at least for $\alpha \in (0.056, 2/3)$; see [14].

3. Matching conditions

In [10], Nakada and Natsui find a condition on the orbits of α and $\alpha - 1$ which allows one to predict the behaviour of the entropy more precisely. Let us denote for any $\alpha \in [0, 1]$, $x \in I_\alpha$,

$n \geqslant 1$ by $M_{\alpha,x,n}$ the matrix such that $T_\alpha^n(x) = M_{\alpha,x,n}^{-1}(x)$, i.e.

$$M_{\alpha,x,n} = \begin{pmatrix} 0 & \epsilon_{\alpha,1} \\ 1 & c_{\alpha,1} \end{pmatrix} \begin{pmatrix} 0 & \epsilon_{\alpha,2} \\ 1 & c_{\alpha,2} \end{pmatrix} \ldots \begin{pmatrix} 0 & \epsilon_{\alpha,n} \\ 1 & c_{\alpha,n} \end{pmatrix}.$$

They proved the following:

Theorem 3.1 ([10, theorem 2]). *Let us suppose that there exist positive integers k_1 and k_2 such that*

(I) $\{T_\alpha^n(\alpha) : 0 \leqslant n < k_1\} \cap \{T_\alpha^m(\alpha - 1) : 0 \leqslant m < k_2\} = \emptyset$,
(II) $M_{\alpha,\alpha,k_1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} M_{\alpha,\alpha-1,k_2}$ $(\Rightarrow T_\alpha^{k_1}(\alpha) = T_\alpha^{k_2}(\alpha - 1))$,
(III) $T_\alpha^{k_1}(\alpha)(= T_\alpha^{k_2}(\alpha - 1)) \notin \{\alpha, \alpha - 1\}$.

Then there exists $\eta > 0$ such that, on $(\alpha - \eta, \alpha + \eta)$, $h(T_\alpha)$ is:

(i) *strictly increasing if $k_1 < k_2$,*
(ii) *constant if $k_1 = k_2$,*
(iii) *strictly decreasing if $k_1 > k_2$.*

It turns out that conditions (I)–(II)–(III) define a collection of open intervals (called *matching intervals*); they also proved that each of the cases (i), (ii) and (iii) takes place at least on one infinite family of disjoint matching intervals clustering at the origin, thus proving the non-monotonicity of the entropy function. Moreover, they conjectured that the union of all matching intervals is a dense, open subset of [0, 1] with full Lebesgue measure.

In the following we will analyse more closely the mechanism which leads to the existence of such matchings. As a consequence, we shall see that it looks more natural to drop condition (III) from the previous definition and replace (II) with

$$(\mathrm{II}') \quad M_{\alpha,\alpha,k_1-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} M_{\alpha,\alpha-1,k_2-1} \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}$$

(which implies $\frac{1}{T^{k_1-1}(\alpha)} + \frac{1}{T^{k_2-1}(\alpha-1)} = -1$, see remark at the end of section 3.1).

We can now define the *matching set* as

$$\mathcal{M} = \{\alpha \in (0, 1] \text{ s.t. (I) and (II}') \text{ hold}\}.$$

Note $\mathcal{M}$ is open, since the symbolic codings of α up to step $k_1 - 1$ and of $\alpha - 1$ up to step $k_2 - 1$ are locally constant.

Moreover, we will see that under this condition it is possible to predict the symbolic orbit of $\alpha - 1$ given the symbolic orbit of α, and vice versa. As an application, we will construct a countable family of intervals which accumulates at a point different from 0.

Let us point out that our matching set $\mathcal{M}$ is a set slightly bigger than the union of all matching intervals satisfying condition (I, II, III): in fact the difference is just a countable set of points.

3.1. Encoding of matchings

Let us consider $PGL(2, \mathbb{Z}) := GL(2, \mathbb{Z})/\{\pm I\}$ as acting on $\mathbb{R} \cup \{\infty\}$ via Möbius transformations, and denote by S, T, V the elements represented by the matrices

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad V = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Using the fact that $PSL(2,\mathbb{Z})$ is the free product of two cyclic groups of order 2 and 3 one can write the presentation

$$PGL(2,\mathbb{Z}) = \{S, T, V \mid S^2 = I, (ST)^3 = I, V^2 = I, VSV^{-1} = S, VTV^{-1} = T^{-1}\}.$$

Now, every step of the algorithm generating α-continued fractions consists of an operation of the type

$$z \mapsto \frac{\epsilon}{z} - c, \qquad \epsilon \in \{\pm 1\}, \quad c \in \mathbb{N},$$

which corresponds to the matrix $T^{-c}SV^{e(\epsilon)}$ with

$$e(\epsilon) = \begin{cases} 0 & \text{if } \epsilon = -1, \\ 1 & \text{if } \epsilon = 1, \end{cases}$$

so if x belongs to the cylinder $((c_1, \epsilon_1), \ldots, (c_k, \epsilon_k))$ we can express

$$T_\alpha^k(x) = T^{-c_k}SV^{e(\epsilon_k)} \cdots T^{-c_1}SV^{e(\epsilon_1)}(x).$$

Now, suppose we have a matching $T_\alpha^{k_1}(\alpha) = T_\alpha^{k_2}(\alpha - 1)$ and let α belong to the cylinder $((a_1, \epsilon_1), \ldots, (a_{k_1}, \epsilon_{k_1}))$ and $\alpha - 1$ belong to the cylinder $((b_1, \eta_1), \ldots, (b_{k_2}, \eta_{k_2}))$. One can rewrite the matching condition as

$$T^{-a_{k_1}}SV^{e(\epsilon_{k_1})} \cdots T^{-a_1}SV^{e(\epsilon_1)}(\alpha) = T^{-b_{k_2}}SV^{e(\eta_{k_2})} \cdots T^{-b_1}SV^{e(\eta_1)}T^{-1}(\alpha),$$

hence it is sufficient to have an equality of the two Möbius transformations

$$T^{-a_{k_1}}SV^{e(\epsilon_{k_1})} \cdots T^{-a_1}SV^{e(\epsilon_1)} = T^{-b_{k_2}}SV^{e(\eta_{k_2})} \cdots T^{-b_1}SV^{e(\eta_1)}T^{-1}.$$

We call such a matching an *algebraic matching*. Now, numerical evidence shows that, if a matching occurs, then

$$\epsilon_1 = +1, \qquad \epsilon_i = -1 \quad \text{for } 2 \leqslant i \leqslant k_1 - 1, \qquad \eta_i = -1 \quad \text{for } 1 \leqslant i \leqslant k_2 - 1.$$

If we make this assumption we can rewrite the matching condition as

$$V^{e(\epsilon_{k_1})+1}T^{a_{k_1}(-1)^{e(\epsilon_{k_1})}}ST^{a_{k_1-1}}S \cdots T^{a_1}S = V^{e(\eta_{k_2})}T^{b_{k_2}(-1)^{[e(\eta_{k_2})+1]}}ST^{-b_{k_2-1}}S \cdots T^{-b_1}ST^{-1},$$

which implies $e(\epsilon_{k_1}) = e(\eta_{k_2}) + 1 \bmod 2$, i.e. $\epsilon_{k_1}\eta_{k_2} = -1$. If for instance $e(\epsilon_{k_1}) = 1$ and $e(\eta_{k_2}) = 0$, by letting

$$U = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

so that $T = SU$, one has

$$\begin{aligned}(U^2S)^{a_{k_1}}U(SU)^{a_{k_1-1}-2}SU^2 \cdots SU^2(SU)^{a_1-2}SUS \\ = (U^2S)^{b_{k_2}-1}US(U^2S)^{b_{k_2-1}-2}US \cdots US(U^2S)^{b_1-2}US. \qquad (3)\end{aligned}$$

Since every element of $PSL(2,\mathbb{Z})$ can be written as a product of S and U in a unique way, one can get a relation between a_r and b_r. Note that, since we are interested in $\alpha \leqslant \sqrt{2} - 1$, $a_i \geqslant 2$ and $b_i \geqslant 2$ for every i, hence there is no cancellation in equation (3). By counting the number of (U^2S) blocks at the beginning of the word, one has $a_{k_1} = b_{k_2} - 1$, and by simplifying,

$$(SU)^{a_{k_1-1}-2}SU^2 \cdots SU^2(SU)^{a_1-2}SUS = S(U^2S)^{b_{k_2-1}-2}US \cdots US(U^2S)^{b_1-2}US. \qquad (4)$$

If we have $e(\epsilon_{k_1}) = 0$ and $e(\eta_{k_2}) = 1$ instead, the matching condition is

$$\begin{aligned}(SU)^{a_{k_1}-1}SU^2(SU)^{a_{k_1-1}-2}SU^2 \cdots SU^2(SU)^{a_1-2}SUS \\ = (SU)^{b_{k_2}}SU^2S(U^2S)^{b_{k_2-1}-2}US \cdots US(U^2S)^{b_1-2}US,\end{aligned}$$

which implies $b_{k_2} = a_{k_1} - 1$, and if you simplify you still get equation (4).

From (4) one has that to every a_r bigger than 2 it corresponds exactly to a sequence of $b_i = 2$ of length precisely $a_r - 2$, and vice versa. More formally, one can give the following algorithm to produce the coding of the orbit of $\alpha - 1$ up to step $k_2 - 1$ given the coding of the orbit of α up to step $k_1 - 1$ (under the hypothesis that an algebraic matching occurs, and at least k_1 is known).

1. Write down the coding of α from step 1 to $k_1 - 1$, separated by a symbol $\star$:
$$a_1 \star a_2 \star \cdots \star a_{k_1-1}.$$
2. Subtract 2 from every a_r; if $a_r = 2$, then leave the space empty instead of writing 0:
$$a_1 - 2 \star a_2 - 2 \star \cdots \star a_{k_1-1} - 2.$$
3. Replace stars with numbers and vice versa (replace the number n with n consecutive stars, and write the number n in place of n stars in a row).
4. Add 2 to every number you find and remove the stars: you will get the sequence $(b_1, \ldots, b_{k_2-1})$.

Example. Let us suppose there is a matching with $k_1 = n + 3$ and α has initial coding $((3, +), (4, -)^n, (2, -))$. The steps of the algorithm are as follows:

Step 1
$$3 \star \underbrace{4 \star 4 \star \cdots \star 4\star}_{n \text{ times}} 2$$
Step 2
$$1 \star \underbrace{2 \star 2 \star \cdots \star 2\star}_{n \text{ times}}$$
Step 3
$$\star 1 \underbrace{\star\star 1 \star\star 1 \cdots 1 \star\star 1}_{n \text{ times}}$$
Step 4
$$2\;3\; \underbrace{2\;3\; \cdots 2\;3}_{n \text{ times}}$$
so the coding of $\alpha - 1$ is $((2, -)(3, -))^{n+1}$, and $k_2 = 2n + 3$.

Remark. Let us remark that (4) is equivalent to
$$T^{-1}ST^{-a_{k_1-1}}S \cdots T^{-a_1}SV = VST^{-b_{k_2-1}}S \cdots T^{-b_1}ST^{-1},$$
which is precisely condition (II′): by evaluating both sides on α you get
$$\frac{1}{T^{k_1-1}(\alpha)} + \frac{1}{T^{k_2-1}(\alpha - 1)} = -1.$$

3.2. *Construction of matchings*

Let us now use this knowledge to construct explicitly an infinite family of matching intervals which accumulates on a non-zero value of α. For every n, let us consider the values of α such that α belongs to the cylinder $((3,+),(4,-)^n,(2,-))$ with the respect to T_α. Let us compute the endpoints of such a cylinder.

- The right endpoint is defined by
$$\begin{pmatrix}-4 & -1\\ 1 & 0\end{pmatrix}^n \begin{pmatrix}-3 & 1\\ 1 & 0\end{pmatrix}(\alpha) = \alpha - 1,$$
i.e.
$$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}-4 & -1\\ 1 & 0\end{pmatrix}^n \begin{pmatrix}-3 & 1\\ 1 & 0\end{pmatrix}(\alpha) = \alpha.$$
- The left endpoint is defined by
$$\begin{pmatrix}-4 & -1\\ 1 & 0\end{pmatrix}^n \begin{pmatrix}-3 & 1\\ 1 & 0\end{pmatrix}(\alpha) = -\frac{1}{\alpha+2},$$
i.e.
$$\begin{pmatrix}-2 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-4 & -1\\ 1 & 0\end{pmatrix}^n \begin{pmatrix}-3 & 1\\ 1 & 0\end{pmatrix}(\alpha) = \alpha.$$

By diagonalizing the matrices and computing the powers one can compute these values explicitly. In particular,
$$\alpha^1_{\min} = \frac{\sqrt{3}-1}{2} + \frac{40\sqrt{3}-69}{13}(2+\sqrt{3})^{-2n} + O((2+\sqrt{3})^{-4n}),$$
$$\alpha^1_{\max} = \frac{\sqrt{3}-1}{2} + \frac{10\sqrt{3}-12}{13}(2+\sqrt{3})^{-2n} + O((2+\sqrt{3})^{-4n}).$$

The αs such that $\alpha-1$ belongs to the cylinder $((2,-),(3,-))^{n+1}$ are defined by the equations
$$\left[\begin{pmatrix}-3 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-2 & -1\\ 1 & 0\end{pmatrix}\right]^{n+1}(\alpha-1) = \alpha - 1$$
for the left endpoint and
$$\left[\begin{pmatrix}-3 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-2 & -1\\ 1 & 0\end{pmatrix}\right]^{n+1}(\alpha-1) = \alpha$$
for the right endpoint, so the left endpoint corresponds to the periodic point such that
$$\left[\begin{pmatrix}-3 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}-2 & -1\\ 1 & 0\end{pmatrix}\right](\alpha-1) = \alpha - 1,$$
i.e.
$$\alpha^2_{\min} = \frac{\sqrt{3}-1}{2}$$
and
$$\alpha^2_{\max} = \frac{\sqrt{3}-1}{2} + \frac{33-19\sqrt{3}}{2}(2+\sqrt{3})^{-2n} + O((2+\sqrt{3})^{-4n}).$$

By comparing the first order terms one gets asymptotically
$$\alpha^2_{\min} < \alpha^1_{\min} < \alpha^2_{\max} < \alpha^1_{\max},$$

hence the two intervals intersect for infinitely many n, producing infinitely many matching intervals which accumulate at the point $\alpha_0 = \frac{\sqrt{3}-1}{2}$. The length of such intervals is

$$\alpha^2_{\max} - \alpha^1_{\min} = \frac{567 - 327\sqrt{3}}{26}(2+\sqrt{3})^{-2n} + O((2+\sqrt{3})^{-4n}).$$

4. Numerical production of matchings

In this section we will describe an algorithm to produce a lot of matching intervals (i.e. find out their endpoints exactly), as well as the results we obtained through its implementation. Our first attempt to find matching intervals used the following scheme:

1. We generate a random seed of values α_i belonging to $[0, 1]$ (or some other interval of interest). When a high precision is needed (we manage to detect intervals of size 10^{-60}) the random seed is composed of algebraic numbers, in order to allow symbolic (i.e. non-floating point) computation.
2. We find numerically candidates for the values of k_1 and k_2 (if any) simply by computing the orbits of α and of $\alpha - 1$ up to some finite number of steps, and numerically checking if $T_\alpha^{k_1}(\alpha) = T_\alpha^{k_2}(\alpha - 1)$ holds approximately for some k_1 and k_2 smaller than some bound.
3. Given any triplet $(\bar{\alpha}, k_1, k_2)$ determined as above, we compute the symbolic orbit of $\bar{\alpha}$ up to step $k_1 - 1$ and the orbit of $\bar{\alpha} - 1$ up to step $k_2 - 1$.
4. We check that the two Möbius transformations associated with these symbolic orbits satisfy condition (II′):
$$M_{\alpha,\alpha,k_1-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} M_{\alpha,\alpha-1,k_2-1} \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix}.$$
5. We solve the system of quadratic equations which correspond to imposing that α and $\alpha - 1$ have the same symbolic orbit as $\bar{\alpha}$ and $\bar{\alpha} - 1$, respectively.
Let us remark that this is the heaviest step of the whole procedure since we must solve $k_1 + k_2 - 2$ quadratic inequalities; for this reason the value $k = k_1 + k_2$ may be thought of as a measure of the computational cost of the matching interval and will be referred to as *order of matching*.

Following this scheme, we detected more than 10^7 matching intervals, whose endpoints are quadratic surds; their union still leaves many gaps, each of which is smaller than 6.6×10^{-6}. A table with a sample of such data is given in the appendix[6].

In order to detect some patterns in the data, let us plot the size of these intervals (figure 4, left). For each matching interval $]\alpha_-, \alpha_+[$, we drew the point of coordinates $(\alpha_-, \alpha_+ - \alpha_-)$. It seems there is some self-similar pattern: in order to understand its structure better it is useful to identify some 'borderline' families of points. The most evident family is the one that appears as the higher line of points in figure 4 (left) (which we have highlighted by circles): these points correspond to matching intervals which contain the values $1/n$, and their endpoints are $\alpha_-(n) = \frac{1}{2}[\sqrt{n^2+4} - n]$, $\alpha_+(n) = \frac{1}{2n-2}[\sqrt{n^2+2n-3} - n + 1]$; this is the family I_n already exhibited in [10]. Since $\alpha_-(n) = 1/n - 1/n^3 + o(1/n^3)$ and $\alpha_+(n) = 1/n + 1/n^3 + o(1/n^3)$, for $n \gg 1$ the points $(\alpha_-(n), \alpha_+(n) - \alpha_-(n))$ are very close to $(\frac{1}{n}, \frac{2}{n^3})$. This suggests that this family will 'straighten' if we replot our data in the log–log scale. This is indeed the case, and in fact it seems that there are also other families which, in log–log scale, get perfectly aligned along parallel lines of slope 3 (see figure 4, right).

[6] A more efficient algorithm, which avoids random sampling, will be discussed in section 4.1.

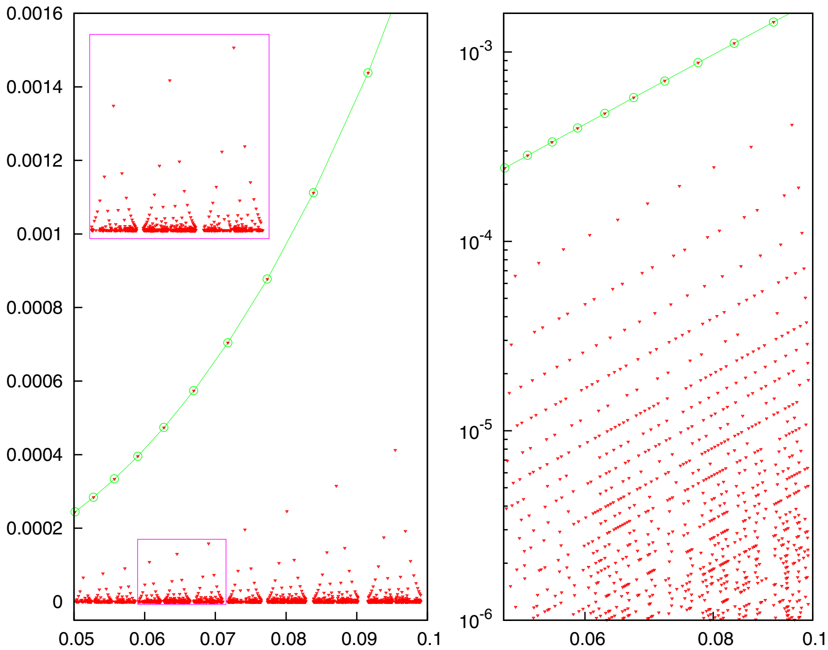

Figure 4. Size of matchings both in normal (right) and log–log scale (left).

If we consider the ordinary continued fraction expansion of the elements of these families we realize that they obey some very simple empirical rules[7]:

(i) the endpoints of any matching interval have a purely periodic continued fraction expansion of the type $[0; \overline{a_1, a_2, ..., a_m, 1}]$ and $[0; \overline{a_1, a_2, ..., a_m + 1}]$; this implies that the rational number corresponding to $[0; a_1, a_2, ..., a_m + 1]$ is a common convergent of both endpoints and is the rational with smallest denominator which falls inside the matching interval;

(ii) any endpoint $[0; \overline{a_1, a_2, ..., a_m}]$ of a matching interval belongs to a family $\{[0; \overline{a, a_2, ..., a_m}] : a \geqslant \max_{2\leqslant i\leqslant m} a_i\}$; in particular this family has a member in each cylinder $B_n := \{\alpha : 1/(n+1) < \alpha < 1/n\}$ for $n \geqslant a$, so that each family will cluster at the origin.

(ii′) other families can be detected in terms of the continued fraction expansion: for instance on each cylinder B_n $(n \geqslant 3)$ the largest matching interval on which h is decreasing has endpoints with expansion $[0; \overline{n, 2, 1, n-1, 1}]$ and $[0; \overline{n, 2, 1, n}]$;

(iii) matching intervals seem to be organized in a binary tree structure, which is related to the Stern–Brocot tree[8] (figure 5): one can thus design a bisection algorithm to fill in the gaps between intervals, and what is left over is a closed, nowhere dense set; this and the following points will be analysed extensively in section 4.1;

(iv) if $\alpha \in B_n$ is the endpoint of some matching interval then $\alpha = [0; \overline{a_1, a_2, ..., a_m}]$ with $a_i \leqslant n\ \forall i \in \{1, \ldots, m\}$; this implies that the values $\alpha \in B_n$ which do not belong to any matching interval must be bounded-type numbers with partial quotients bounded above by n;

[7] These rules have been proved to be correct, see [2].

[8] Sometimes also known as Farey tree, see [4].

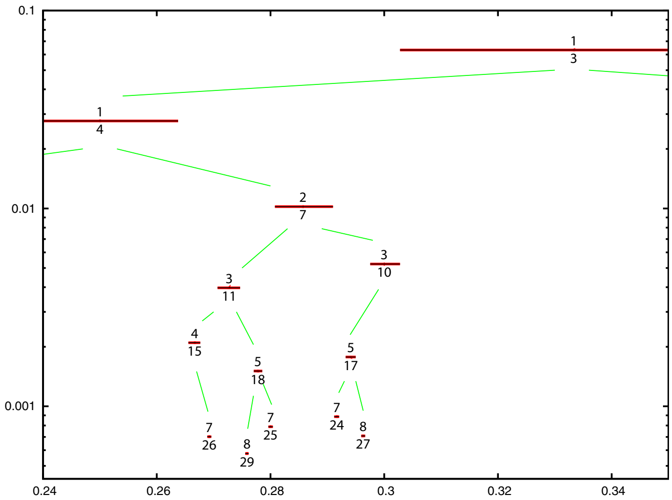

Figure 5. A few of the largest matching intervals in the window $[1/4, 1/3]$, and the corresponding nodes of Stern–Brocot tree. The value on the y-axis is the logarithm of the size of each interval.

(v) it is possible to compute the exponent (k_1, k_2) of a matching from the continued fraction expansion of any one of its endpoints.

From our data it is also evident that the size of these intervals decreases as k_1+k_2 increases, and low order matchings tend to disappear as α approaches zero (see figure 6).

Moreover, as α tends to 0 the space covered by members of families of type (ii) previously encountered decreases, hence new families have to appear. One can quantify this phenomenon from figure 4: since the size of matching intervals in any family decreases as $2/n^3$ on the interval cylinder B_n (whose size decreases as $1/n^2$): this means that, as n increases, the mass of B_n gets more and more split among a huge number of tiny intervals.

This fact compromises our numerical algorithm: it is clear that choosing floating point values at random becomes a hopeless strategy when approaching zero. Indeed, even if there still are intervals bigger than the double-precision threshold, in most cases the random seed will fall in a really tiny interval corresponding to a very high matching order: this amounts to having very little gain as the result of a really heavy computation.

We still can try to test numerically the conjecture that the matching set has full measure on $[0, 1]$; but we must expect that the percentage of space covered by matching intervals (found numerically) will decrease dramatically near the origin, since we only detect intervals with $k_1 + k_2$ bounded by some threshold. The matching intervals we have found so far cover a portion of 0.884 of the interval $[0, 1]$; this ratio increases to 0.989 if we restrict to the interval $[0.1, 1]$ and it reaches 0.9989 restricting to the interval $[0.2, 1]$.

Figure 7 represents the percentage of the interval $[x, 1]$ which is covered by matching intervals of order $k = k_1 + k_2$ for different values of k[9]. It gives an idea of the gain, in terms of the total size covered by matching intervals, one gets when refining the gaps (i.e. considering matching intervals of higher order).

Finally, by looking at the scattered plot in figure 8, one can have a more precise picture of the relationship between the order of matching (on the x-axis) and the size of the matching

[9] Let us point out that for big values of k the graph does not take into account all matching intervals of order k but only those we have found so far.

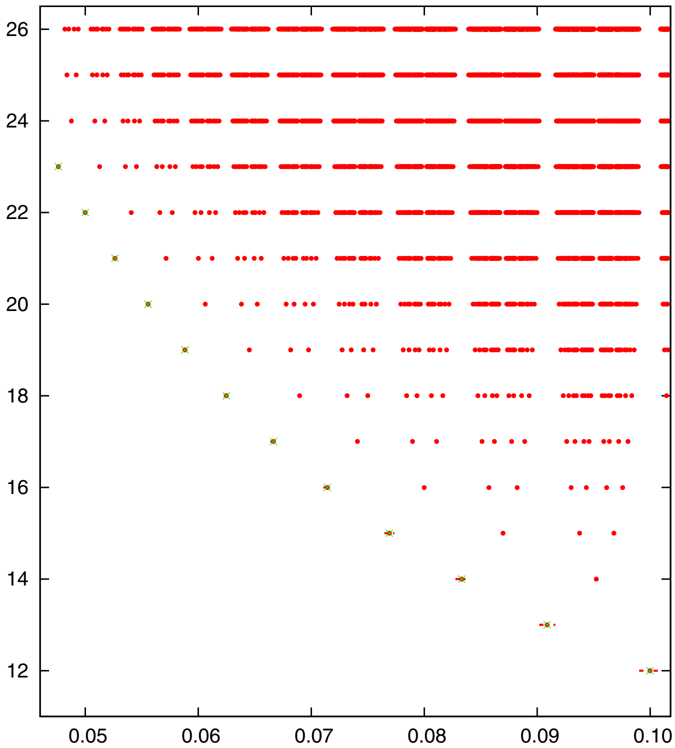

Figure 6. Plot of the order $k = k_1 + k_2$ of matching intervals versus their position.

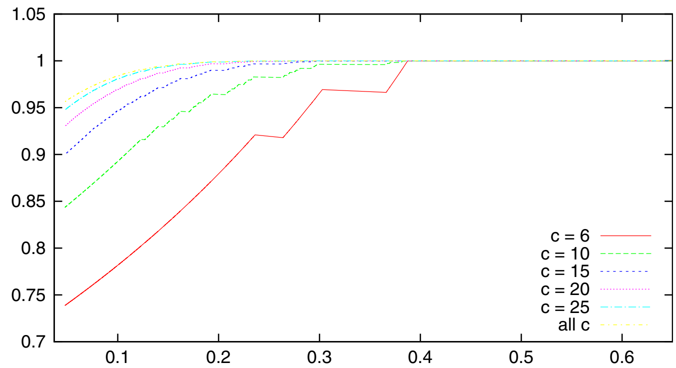

Figure 7. Percentage of covering by matching intervals with $k_1 + k_2 \leqslant c$ for different values of c.

interval (on the y-axis). The two data sets correspond, respectively, to the matching intervals constructed via the bisection scheme and those found using a random seed. The two lines bounding the cloud correspond to matching intervals with very definite patterns: the upper line corresponds to the family I_n (with endpoints of type $[0; \overline{n}]$ and $[0; \overline{n-1, 1}]$), the lower line corresponds to matching intervals with endpoints of type $[0; \overline{2, 1, 1, \ldots, 1, 1, 1}]$ and $[0; \overline{2, 1, 1, \ldots, 1, 2}]$. The latter ones converge to $\frac{3-\sqrt{5}}{2}$, which is the supremum of all values where the entropy is increasing.

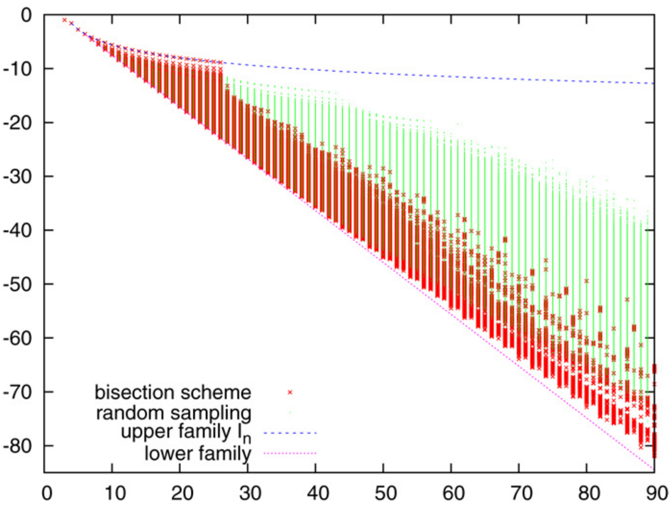

Figure 8. The order $k_1 + k_2$ versus the logarithm of the size of matching intervals.

Thus numerical evidence shows that if J is an interval with matching order $k = k_1 + k_2$, then the size of J is bounded below by $|J| \geqslant c_0 \mathrm{e}^{-c_1 k}$ where $c_0 = 8.4423...$ and $c_1 = 0.9624....$ On the other hand we know for sure that, on the right of 0.0475 (which corresponds to the leftmost matching interval of our list), the biggest gap left by the matching intervals found so far is of order 6.6×10^{-6}. So, if J is a matching interval which still does not belong to our list, either $J \subset [0, 0.0475]$ and $k \geqslant 20$, or its size must be smaller than 6.6×10^{-6} and by the aforementioned empirical rule, its order must be $k > 14.6$. Hence, our list should include all matching intervals with $k_1 + k_2 \leqslant 14$.

4.1. The matching tree

As mentioned before, it seems that matching intervals are organized in a binary tree structure. To describe such a structure, we will provide an algorithm which allows one to construct all matching intervals by recursively 'filling the gaps' between matching intervals previously obtained, similarly to the way the usual Cantor middle third set is constructed.

In order to do so, let us first note that every rational value $r \in \mathbb{Q} \cap (0, 1]$ has two (standard) continued fraction expansions:

$$r = [0; a_1, a_2, \ldots, a_m, 1] = [0; a_1, a_2, \ldots, a_m + 1].$$

One can associate with r the interval whose endpoints are the two quadratic surds with continued fraction obtained by endless repetition of the two expansions of r:

Definition 4.1. *Given $r \in \mathbb{Q} \cap (0, 1]$ with continued fraction expansion as above, we define I_r to be the interval with endpoints*

$$[0; \overline{a_1, a_2, \ldots, a_m, 1}] \text{ and } [0; \overline{a_1, a_2, \ldots, a_m + 1}]$$

(in any order). The strings $S_1 := \{a_1, \ldots, a_m, 1\}$ and $S_2 := \{a_1, \ldots, a_m + 1\}$ will be said to be conjugate and we will write $S_2 = (S_1)'$.

Note that $r \in I_r$.

Definition 4.2. *Given an open interval $I \subseteq [0, 1]$ one can define the pseudocentre of I as the rational number $r \in I \cap \mathbb{Q}$ which has the minimum denominator among all rational numbers contained in I.*

It is straightforward to prove that the pseudocentre of an interval is unique, and the pseudocentre of I_r is r itself.

We are now ready to describe the *bisection algorithm*:

1. The rightmost matching interval is $[\frac{\sqrt{5}-1}{2}, 1]$; its complement is the *gap* $J = [0, \frac{\sqrt{5}-1}{2}]$.
2. Suppose we are given a finite set of intervals, called *gaps of level n*, so that their complement is a union of matching intervals. Given each gap $J = [\alpha^-, \alpha^+]$, we determine its *pseudocentre r*. Let $\alpha^\pm = [0; S, a^\pm, S^\pm]$ be the continued fraction expansion of $\alpha^\pm$, where S is the finite string containing the first common partial quotients, $a^+ \neq a^-$ the first partial quotient on which the two values differ and $S^\pm$ the rest of the expansion of $\alpha^\pm$, respectively. The pseudocentre of $[\alpha^-, \alpha^+]$ will be the rational number r with expansions $[0; S, a, 1] = p/q = [0; S, a+1]$ where $a := \min(a^+, a^-)$.
3. We remove from the gap J the matching interval I_r corresponding to the pseudocentre r: in this way the complement of I_r in J will consist of two intervals J_1 and J_2, which we will add to the list of gaps of level $n+1$. It might occur that one of these new intervals consists of only one point, i.e. two matching intervals are adjacent.

By iterating this procedure, after n steps we will get a finite set $\mathcal{G}_n$ of gaps, and clearly $\bigcup_{J \in \mathcal{G}_{n+1}} J \subseteq \bigcup_{J \in \mathcal{G}_n} J$. We expect that all intervals obtained by taking pseudocentres of gaps are matching intervals, and that the set on which matching fails is the intersection

$$\mathcal{G}_\infty := \bigcap_{n \in \mathbb{N}} \bigcup_{J \in \mathcal{G}_n} J.$$

Table 1 contains the list of the elements of the family $\mathcal{G}_n$ of gaps of level n for $n = 0, \ldots, 4$: when a gap is reduced to a point we mark the corresponding line with the symbol $\star$. A few steps of the algorithm are also displayed in figure 9.

The numerical evidence supporting the correctness of this scheme is quite robust: all 1 169 731 intervals obtained by running the first 23 steps turn out to be real matching intervals.

We can also prove the following:

Lemma 4.1. *$\mathcal{G}_\infty$ consists of numbers of bounded type; more precisely, the elements of $\mathcal{G}_\infty \cap (\frac{1}{n+1}, \frac{1}{n}]$ have regular continued fraction bounded by n.*

Proof. The scheme described before forces all endpoints of matching intervals containted in the cylinder $B_n =]1/(n+1), 1/n[$ to have quotients bounded by n. We now claim that, if $\gamma = [0; c_1, c_2, \ldots, c_n, \ldots] \notin \mathcal{M}$, then, $c_k \leqslant c_1$ for all $k \in \mathbb{N}$.

If $\gamma \notin \mathcal{M}$ then $\gamma \in \bigcup_{J \in \mathcal{G}_n} J$ for all $n \in \mathbb{N}$; let us call J_n the member of the family $\mathcal{G}_n$ containing γ. It may happen that there exists n_0 such that $J_n = \{\gamma\}\ \forall n \geqslant n_0$. γ is an endpoint of two adjacent matching intervals, hence it has bounded type. Otherwise, $J_n = [\alpha_n, \beta_n]$ with $\beta_n - \alpha_n > 0\ \forall n > c_1$, where α_n, β_n are the endpoints of two matching intervals. Now, if p_n/q_n is the pseudocentre of J_n from the minimality of q_n it follows that $|\beta_n - \alpha_n| < 2/q_n$, but also that $q_{n+1} > q_n$ (since $p_{n+1}/q_{n+1} \in J_{n+1} \subset J_n$); these two properties together imply that $0 \leqslant \gamma - \alpha_n < 2/q_n \to 0$ as $n \to +\infty$. This implies γ cannot be rational, since $\gamma \in J_n\ \forall n$ and the minimum denominator of a rational sitting in J_n is $q_n \to +\infty$. Hence, since $\alpha_n \to \gamma$, for every fixed $k \in \mathbb{N}$, there is some $n(k)$ such that for all $n \geqslant n(k)$ all the partial quotients up to level k of γ coincide with those of α_n, which are bounded by c_1. $\square$

Table 1. Gaps of level n for $n = 0, \dots, 4$.

		α^-	α^+
$\mathcal{G}_0$		0	$[0; \overline{1}]$
$\mathcal{G}_1$		0	$[0; \overline{2}]$
	$\star$	$[0; \overline{1, 1}]$	$[0; \overline{1}]$
$\mathcal{G}_2$		0	$[0; \overline{3}]$
		$[0; \overline{2, 1}]$	$[0; \overline{2}]$
	$\star$	$[0; \overline{1, 1}]$	$[0; \overline{1}]$
$\mathcal{G}_3$		0	$[0; \overline{4}]$
		$[0; \overline{3, 1}]$	$[0; \overline{3}]$
		$[0; \overline{2, 1}]$	$[0; \overline{2, 1, 1}]$
	$\star$	$[0; \overline{2, 2}]$	$[0; \overline{2}]$
	$\star$	$[0; \overline{1, 1}]$	$[0; \overline{1}]$
$\mathcal{G}_4$		0	$[0; \overline{5}]$
		$[0; \overline{4, 1}]$	$[0; \overline{4}]$
		$[0; \overline{3, 1}]$	$[0; \overline{3, 1, 1}]$
		$[0; \overline{3, 2}]$	$[0; \overline{3}]$
		$[0; \overline{2, 1}]$	$[0; \overline{2, 1, 2}]$
		$[0; \overline{2, 1, 1, 1}]$	$[0; \overline{2, 1, 1}]$
	$\star$	$[0; \overline{2, 2}]$	$[0; \overline{2}]$
	$\star$	$[0; \overline{1, 1}]$	$[0; \overline{1}]$
...		...	...

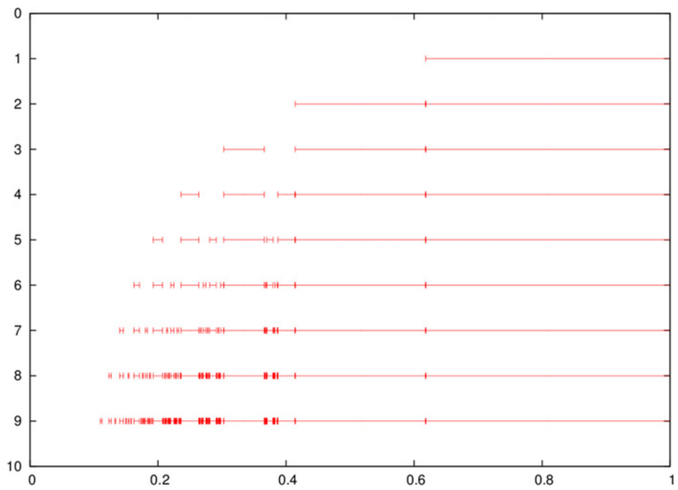

Figure 9. Recursive construction of the matching set.

Lemma 4.1 shows that the correctness of our bisection scheme (i.e. the fact that $\mathcal{G}_\infty = [0, 1] \setminus \mathcal{M}$) implies claim 1.1.

Note that $\mathcal{G}_\infty \cap (1/(n + 1), 1/n]$ has Hausdorff dimension strictly smaller than one for each n. Moreover, the Hausdorff dimension of n-bounded numbers tends to 1 as $n \to \infty$. We think that, similarly, $\mathrm{H.dim}\{(\frac{1}{n+1}, \frac{1}{n}] \setminus \mathcal{M}\} \to 1$; this could explain why finding matching intervals near the origin becomes a tough task.

Remark. Since we have associated a rational number with each matching interval, one can think of the bisection algorithm as acting on $\mathbb{Q}$, and get a binary tree whose nodes are rationals: this object has already been widely studied in the number-theoretical literature, and it is known as the *Stern–Brocot* tree (see [4]).

Given that all matching intervals correspond to some rational number, one can ask which subset of $\mathbb{Q}$ actually arises in that way.

Definition 4.3. *An interval I_r, $r \in \mathbb{Q} \cap (0, 1]$ is maximal if $I_r \not\subseteq I_{r'}$ $\forall r' \in \mathbb{Q} \cap (0, 1]$, $r' \neq r$.*

We expect[10] that the matching intervals are *precisely* the maximal intervals, so that the matching set is

$$\mathcal{M} = \bigcup_{r \in (0,1] \cap \mathbb{Q}} I_r = \bigcup_{\substack{r \in (0,1] \cap \mathbb{Q} \\ I_r \text{maximal}}} I_r.$$

We have also found an empirical rule to reconstruct the periods (k_1, k_2) of a matching interval from the labels of its endpoints. Let $S = [a_1, \ldots, a_\ell]$ be a label of the endpoint s of some matching interval:

1. If s is a left endpoint then
$$k_1 = 2 + \sum_{j \text{ even}} a_j, \qquad k_2 = \sum_{j \text{ odd}} a_j.$$
2. If s is a right endpoint then
$$k_1 = 1 + \sum_{j \text{ even}} a_j, \qquad k_2 = 1 + \sum_{j \text{ odd}} a_j.$$

Using this rule, we are able to prove that every neighbourhood of the point $[0; 3, \bar{1}]$ contains intervals of matching of all types: with $k_1 < k_2$, $k_1 = k_2$ and $k_1 > k_2$. Indeed, it is not difficult to realize that $[0; 3, \bar{1}]$ is contained in the family of gaps J_P of endpoints $[0; \overline{3, P}]$ and $[0; \overline{3, P, 1}]$ where P is a string of the type $1, 1, \ldots, 1, 1$ of even length; by our rule the left endpoint of J_P is the right endpoint of an interval of matching where $k_1 < k_2$. Nevertheless, by performing a few steps of the algorithm, it is not difficult to check that the gap J_P contains the interval C_P of endpoints $[0; \overline{3, P, 2, 1, 1}]$ and $[0; \overline{3, P, 2, 1, 1}]$ (on which $k_1 = k_2$) but also D_P of endpoints $[0; \overline{3, P, 2, 1, 2, 1}]$and $[0; \overline{3, P, 2, 1, 3}]$ (on which $k_1 > k_2$).

4.2. Adjacent intervals and period doubling

Let us now focus on pairs of adjacent intervals (corresponding to isolated points in $[0, 1] \setminus \mathcal{M}$): our data show they all come in infinite chains, and can be obtained from some starting matching interval via a 'period doubling' construction.

Let us start with a matching interval $]\alpha, \beta[$; $\alpha = [0; \overline{S}]$ where S is a sequence of positive integers of odd length; define the sequence of strings

$$\begin{cases} S_0 = S, \\ S_{n+1} = (S_n S_n)', \end{cases} \tag{5}$$

where S' denotes the conjugate of S as in definition 4.1. Let $a_n := [0; \overline{S_n}]$ and $b_n := [0; \overline{S_n'}]$; then the sequence $I_n :=]a_n, b_n[$ is formed by a chain of adjacent intervals: clearly $b_{n+1} = a_n$; moreover $a_n < b_n$ because $|S_n|$ is odd for all n.

[10] This fact has now been proved in [2].

Assuming this scheme, we can construct many cluster points of matching intervals. For instance, let us look at the first (i.e. rightmost) one: we start with the interval $](\sqrt{5}-1)/2,\ 1[$ so that the first terms of the sequence S_n are

$$S_0 = (1),$$
$$S_1 = (2),$$
$$S_3 = (2, 1, 1),$$
$$S_4 = (2, 1, 1, 2, 2),$$
$$S_5 = (2, 1, 1, 2, 2, 2, 1, 1, 2, 1, 1),$$

The corresponding sequence a_n converges to the first (i.e. rightmost) point $\hat{\alpha}$ where intervals of matching cluster. We can also determine the continued fraction expansion of the value $\hat{\alpha}$, since it can be obtained by just merging[11] the strings $(S_n)_n \in \mathbb{N}$:

$$\hat{\alpha} = [0; 2, 1, 1, 2, 2, 2, 1, 1, 2, 1, 1, 2, 1, 1, 2, 2, 2, 1, 1, 2, 2, 2, 1, 1, 2, 2, 2, 1, 1, 2, 1, 1, \ldots].$$

Numerically[12], $\hat{\alpha} \cong 0.386749970714300706171524803485580939661\ldots$.

It is evident from formula (5) that any such cluster point will be a bounded-type number; one can indeed prove that no cluster point of this type is a quadratic surd.

5. Behaviour of entropy inside the matching set

In [10], the following formula is used to relate the change of entropy between two sufficiently close values of α to the invariant measure corresponding to one of these values: more precisely

Proposition 5.1. *Let us suppose the hypotheses of proposition 3.1 hold for α: then for $\eta > 0$ small enough*

$$h(T_{\alpha-\eta}) = \frac{h(T_\alpha)}{1 + (k_2 - k_1)\mu_\alpha([\alpha - \eta, \alpha])} \tag{6}$$

and similarly

$$h(T_\alpha) = \frac{h(T_{\alpha+\eta})}{1 + (k_2 - k_1)\mu_{\alpha+\eta}([\alpha, \alpha + \eta])}. \tag{7}$$

By exploiting these formulae, we will get some results on the behaviour of $h(T_\alpha)$.

5.1. One-sided differentiability of $h(T_\alpha)$

Equation (6) has interesting consequences on the differentiability of h: we can rewrite it as

$$h(T_\alpha) - h(T_{\alpha-\eta}) = h(T_{\alpha-\eta})(k_2 - k_1)\mu_\alpha([\alpha - \eta, \alpha])$$

and dividing by η

$$\frac{h(T_\alpha) - h(T_{\alpha-\eta})}{\eta} = h(T_{\alpha-\eta})(k_2 - k_1)\frac{\mu_\alpha([\alpha - \eta, \alpha])}{\eta}.$$

[11] This can be done since, by (5), S_n is a substring of S_{n+1}.
[12] This pattern has been checked up to level 10, which corresponds to a matching interval of size smaller than 10^{-200}; see also table A.1 in the appendix.

Since ρ_α has bounded variation, then there exists $R(\alpha) = \lim_{x\to\alpha^-} \rho_\alpha(x)$, therefore

$$\lim_{\eta\to 0} \frac{\mu_\alpha([\alpha-\eta,\alpha])}{\eta} = R(\alpha)$$

and by the continuity of h (which is obvious in this case by equation (6))

$$\lim_{\eta\to 0} \frac{h(T_\alpha) - h(T_{\alpha-\eta})}{\eta} = h(T_\alpha)(k_2 - k_1) \lim_{x\to\alpha^-} \rho_\alpha(x),$$

hence the function $\alpha \mapsto h(T_\alpha)$ is left differentiable in α. On the other hand, one can slightly modify the proof of (7) and realize it is equivalent to

$$h(T_{\alpha+\eta}) = \frac{h(T_\alpha)}{1 + (k_1 - k_2)\mu_\alpha([\alpha-1, \alpha-1+\eta])},$$

which reduces to

$$\frac{h(T_{\alpha+\eta}) - h(T_\alpha)}{\eta} = \frac{\mu_\alpha([\alpha-1,\alpha-1+\eta])}{\eta} \frac{h(T_\alpha)(k_2-k_1)}{1 + (k_1 - k_2)\mu_\alpha([\alpha-1,\alpha-1+\eta])}.$$

Since the limit

$$\lim_{\eta\to 0} \frac{\mu_\alpha([\alpha-1,\alpha-1+\eta])}{\eta} = \lim_{x\to(\alpha-1)^+} \rho_\alpha(x)$$

also exists, then $h(T_\alpha)$ is also right differentiable in α; more precisely

$$\lim_{\eta\to 0} \frac{h(T_{\alpha+\eta}) - h(T_\alpha)}{\eta} = h(T_\alpha)(k_2 - k_1) \lim_{x\to(\alpha-1)^+} \rho_\alpha(x).$$

We conjecture that in such points the left and right derivatives are equal. This is trivial for $k_1 = k_2$; for $k_1 \neq k_2$ it is equivalent to say $\lim_{x\to\alpha^-} \rho_\alpha(x) = \lim_{x\to(\alpha-1)^+} \rho_\alpha(x)$.

5.2. The entropy for $\alpha \geqslant \frac{2}{5}$

Corollary 5.2. *For $\frac{2}{5} \leqslant \alpha \leqslant \sqrt{2}-1$, the entropy is*

$$h(T_\alpha) = \frac{\pi^2}{6\log\left(\dfrac{\sqrt{5}+1}{2}\right)}.$$

Proof. Every α in the interval $(0.4, \sqrt{2}-1)$ satisfies the hypotheses of the theorem with $k_1 = k_2 = 3$, hence $h(T_\alpha)$ is locally constant, and by continuity $h(T_\alpha) = h(T_{\sqrt{2}-1})$, whose value was already known. $\square$

Remark. By using our computer-generated matching intervals, we can analogously prove $h(T_\alpha) = h(T_{\sqrt{2}-1})$ for

$$\sqrt{2} - 1 \geqslant \alpha \geqslant 0.386749970714300706171524\ldots.$$

5.3. Invariant densities

In the case $\alpha \geqslant \sqrt{2}-1$ it is known that invariant densities are of the form

$$\rho_\alpha(x) = \sum_{i=1}^{r} \chi_{I_i}(x) \frac{A_i}{x + B_i},$$

where I_i are subintervals of $[\alpha-1, \alpha]$.

For these values of α, a matching condition is present and the endpoints of I_i (i.e. the values where the density may 'jump') correspond exactly to the first few iterates of α and $\alpha - 1$ under the action of T_α. We present some numerical evidence in order to support the following:

Conjecture 5.3. *Let $\alpha \in [0, 1]$ be a value such that one has a matching of type (k_1, k_2) (i.e. with $T_\alpha^{k_1}(\alpha) = T_\alpha^{k_2}(\alpha - 1)$). Then the invariant density has the form*

$$\rho_\alpha(x) = \sum_{i=1}^{r} \chi_{I_i}(x) \frac{A_i}{x + B_i} \tag{8}$$

where each I_i is an interval with endpoints contained in the set

$$S := \{T_\alpha^m(\alpha) : 0 \leqslant m < k_1\} \cup \{T_\alpha^n(\alpha - 1) : 0 \leqslant n < k_2\}.$$

Therefore, the number of branches is bounded above by $k_1 + k_2 - 1$.

In all known cases, moreover, there exists exactly one I_i which contains α and exactly one which contains $\alpha - 1$; thus, on neighbourhoods of α and $\alpha - 1$, the invariant density has the simple form $\rho_\alpha|_{I_i}(x) = \frac{A_i}{x+B_i}$.

As an example of such numerical evidence we report a numerical simulation of the invariant density for some values of α in the interval $[\frac{\sqrt{13}-3}{2}, \frac{\sqrt{3}-1}{2}]$ where a matching of type (2, 3) occurs. We fit the invariant density with the function $A_+/(x + B_+)$ on the interval $[\max\{S\}, \alpha]$ and with the function $A_-/(x + B_-)$ on $[\alpha - 1, \min\{S\}]$.

	$\alpha = 0.310$	$\alpha = 0.320$	$\alpha = \frac{1}{3}$	$\alpha = 0.338$	$\alpha = 0.350$	$\alpha = 0.360$
A_+	1.76114	1.76525	1.77603	1.78963	1.81981	1.84658
B_+	1.64768	1.63487	1.62374	1.62987	1.64092	1.65138
A_-	1.77289	1.78874	1.81488	1.82411	1.84562	1.85959
B_-	2.66097	2.66081	2.66583	2.66751	2.66915	2.6658

Moreover, from these numerical data it is apparent that the leftmost branch of the hyperbola is nothing else than a translation by 1 of the rightmost one (i.e. $A_+ = A_-$, $B_- = B_+ + 1$); see also figure 10.

5.4. Comparison with the entropy

If $I \subset [0, 1]$ is a matching interval, the knowledge of the invariant density for one single value of $\alpha \in I$ plus equation (6) allows us to recover the entropy in the whole interval. Let α belong to an interval where a matching of type (k_1, k_2) occurs and suppose, according to the previous conjecture, that on $[\overline{x}, \alpha]$ the invariant density has the form

$$\rho_\alpha(x) = \frac{A}{x + B}$$

for some $A, B \in \mathbb{R}$ and $\overline{x} = \max\{T_\alpha^n(\alpha), 1 \leqslant n < k_1\} \cup \{T_\alpha^m(\alpha - 1), 1 \leqslant m < k_2\}$. Then by (6), for $x < \alpha$ sufficiently close to α

$$h(x) = \frac{h(\alpha)}{1 + (k_2 - k_1) A \log\left(\dfrac{B + \alpha}{B + x}\right)}. \tag{9}$$

We think that the entropy has in general such form for values of α where a matching occurs.

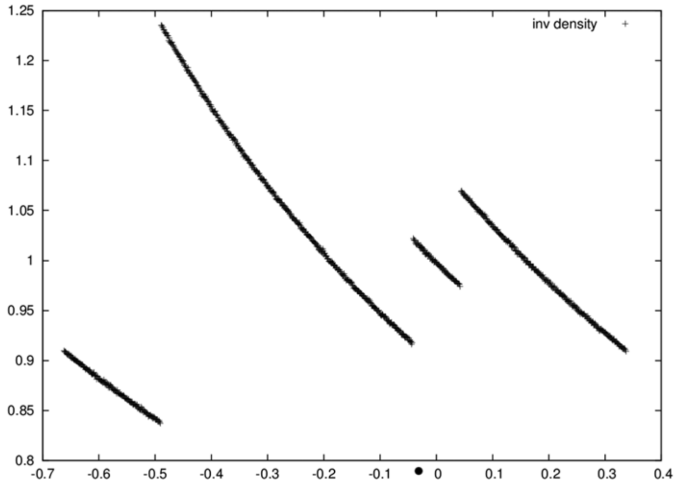

Figure 10. Invariant density for $\alpha = 0.338$.

Let us consider the particular case of the interval [0.295, 0.3042]. In the region to the right of the big central plateau (i.e. for $\alpha > \frac{-3+\sqrt{13}}{2}$) the behaviour of entropy looks approximately linearly increasing, as conjectured in [7, section 3]. We will provide numerical evidence that it actually has the logarithmic form given by equation (9) on the interval $[\frac{\sqrt{13}-3}{2}, \frac{\sqrt{3}-1}{2}]$. To test this hypothesis, we proceed as follows:

1. We fit the data of the invariant density for $\alpha = 0.338$, obtaining the constants A_+ and B_+ which refer to the rightmost branch of the hyperbola (the data are already given in the previous table).
2. We fit the data of the entropy already calculated (relative to the window [0.30277, 0.3042]) with function (9). We assume A_+ and B_+ as given constants and we look for the best possible value of $h(\alpha)$ (which we did not have from previous computations). The result given is $h(\alpha) \cong 3.28302$. In the figure 11 we plot the obtained function in the known window, as well as a linear fit. In this interval, the difference between the two functions is negligible (figure 11, left).
3. In order to really distinguish between linear and logarithmic behaviour of the entropy, we computed some more numerical data for the entropy far away to the right but in the same matching interval. In this region the linear and logarithmic plots are clearly distinguishable, and the new points seem to perfectly agree with the logarithmic formula[13] (figure 11, right).

Note that these data agree with equation (9) also for $x > \alpha$, which is equivalent to say $\rho_\alpha(x) = \frac{A}{B+1+x}$ for x in a right neighbourhood of $\alpha - 1$.

Acknowledgments

This research was partially supported by the project 'Dynamical Systems and Applications' of the Italian Ministry of University and Research (PRIN 2007B3RBEY) and the Centro di

[13] Let us remark that the new values computed are just a few, but are more accurate than those in the interval [0.30277, 0.3042] since we used the package CLN, a C++ library to perform computations in arbitrary precision.

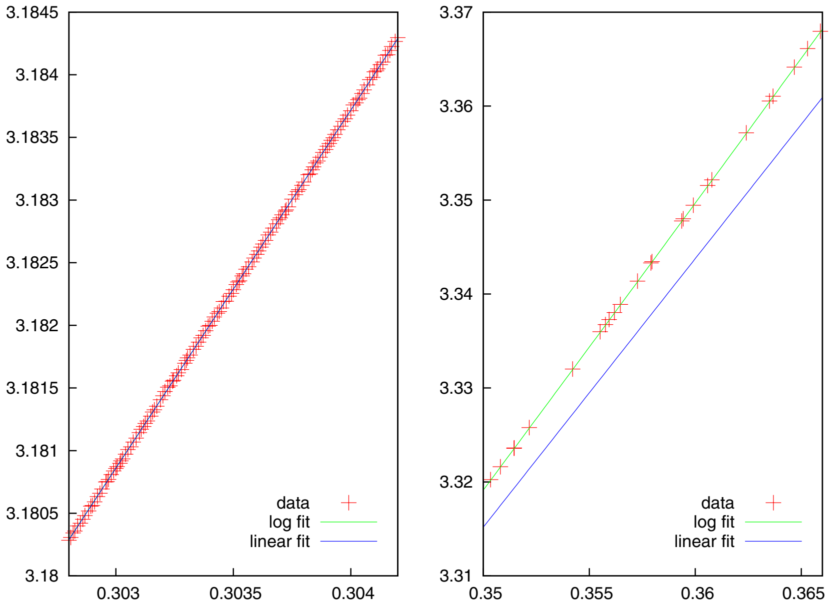

Figure 11. Invariant density for $\alpha = 0.338$.

Ricerca Matematica 'Ennio De Giorgi'. We also wish to thank the Max Planck Institute in Bonn for their hospitality in July 2009.

Appendix

In this appendix we give the proof of two simple results which are of some relevance for the issues discussed in this paper.

Proposition A.1. *If x_0 is a quadratic surd then x_0 is a pre-periodic point for T_α, $\alpha \in [0, 1]$.*

For $\alpha = 1$ this is the well-known Lagrange theorem, and this statement is known to be true for $\alpha = 0$ and $\alpha \in [1/2, 1]$ [6]. Since we did not find a reference containing a simple proof of this fact for all $\alpha \in [0, 1]$ we sketch it here, in few lines: this proof follows closely the classical proof of Lagrange theorem for regular continued fractions given by [3] which relies on the approximation properties of convergents, therefore it works for $\alpha > 0$.

If x_0 is a quadratic surd then $F_0(x_0) = 0$ for some $F_0(x) := A_0x^2 + B_0x + C_0$ quadratic polynomial with integer coefficients. On the other hand, since[14] $x_0 = \frac{p_{n-1}x_n+p_n}{q_{n-1}x_n+q_n}$, setting $F_n(x) := F_0(\frac{p_{n-1}x+p_n}{q_{n-1}x+q_n})(q_{n-1}x + q_n)^2$, we get that $F_n(x_n) = F_0(x_0) = 0$.

Moreover $F_n(x) = A_nx^2 + B_nx + C_n$ with

$$\begin{cases} A_n = F_0(p_{n-1}/q_{n-1})q_{n-1}^2, \\ C_n = F_0(p_n/q_n)q_n^2, \\ B_n^2 - 4A_nC_n = B_0^2 - 4A_0C_0. \end{cases} \tag{10}$$

Both A_n, B_n are bounded since $|F_0(p_n/q_n)| = |F_0(p_n/q_n) - F_0(x_0)| = |F_0'(\xi)| \cdot |\frac{p_n}{q_n} - x_0| \leqslant \frac{C}{\alpha q_n^2}$; moreover from the last equation in (10) it follows that B_n are bounded as well.

[14] To simplify notation we shall write p_n, q_n instead of $p_{n,\alpha}$, $q_{n,\alpha}$.

Proposition A.2. *The variance $\sigma^2(\alpha)$ is constant for $\alpha \in [\sqrt{2}-1, (\sqrt{5}-1)/2]$.*

This result relies on the fact that for all $\alpha \in [\sqrt{2}-1, (\sqrt{5}-1)/2]$ the maps T_α have natural extensions $\tilde{T}_\alpha$ which are all isomorphic to $\tilde{T}_{1/2}$. In the following we shall prove the claim for $\alpha \in [\sqrt{2}-1, 1/2]$ and we shall write T_1 instead of T_α and T_2 instead of $T_{1/2}$. So $T_j : I_j \to I_j$, $(j = 1, 2)$ are one-dimensional maps with invariant measure μ_j; $\tilde{T}_j : \tilde{I}_j \to \tilde{I}_j$, $(j = 1, 2)$ are the corresponding two-dimensional representations of the natural extension with invariant measure $\tilde{\mu}_j$, and $\Phi : \tilde{I}_1 \to \tilde{I}_2$ is the (measurable) isomorphism

$$\Phi \circ \tilde{T}_1 = \tilde{T}_2 \circ \Phi, \qquad \Phi_* \tilde{\mu}_1 = \tilde{\mu}_2.$$

First let us point out (see [10, pp 1222–3]) that Φ is almost everywhere differentiable and has a diagonal differential; moreover $\tilde{T}_j$ are almost everywhere differentiable as well and have triangular differential. Therefore

$$\mathrm{d}\Phi|_{T_1(x,y)}\, \mathrm{d}\tilde{T}_1|_{(x,y)} = \mathrm{d}\tilde{T}_2|_{\Phi(x,y)}\, \mathrm{d}\Phi_{(x,y)} \tag{11}$$

and it is easy to check that, setting $\tilde{T}_j^x$ the first component of $\tilde{T}_j$, a scalar analogue holds as well:

$$\left.\frac{\partial \Phi^x}{\partial x}\right|_{T_1(x,y)} \left.\frac{\partial \tilde{T}_1^x}{\partial x}\right|_{(x,y)} = \left.\frac{\partial \tilde{T}_2^x}{\partial x}\right|_{\Phi(x,y)} \left.\frac{\partial \Phi^x}{\partial x}\right|_{(x,y)}. \tag{12}$$

So we get that, for all k,

$$\log\left|\frac{\partial \tilde{T}_1^x}{\partial x}\right| = \log\left|\frac{\partial \tilde{T}_2^x}{\partial x} \circ \Phi\right| + \log\left|\frac{\partial \Phi^x}{\partial x}\right| - \log\left|\frac{\partial \Phi^x}{\partial x} \circ \tilde{T}_1\right|.$$

Since $\tilde{T}_1^x$ is $\tilde{\mu}_1$-measure preserving $\int_{\tilde{I}_1} \log|\frac{\partial \Phi^x}{\partial x}| - \log|\frac{\partial \Phi^x}{\partial x} \circ \tilde{T}_1|\, \mathrm{d}\tilde{\mu}_1 = 0$; so, taking into account that $\Phi_{\tilde{\mu}_1} = \tilde{\mu}_2$ we get

$$\int_{\tilde{I}_1} \log\left|\frac{\partial \tilde{T}_1^x}{\partial x}\right| \mathrm{d}\tilde{\mu}_1 = \int_{\tilde{I}_1} \log\left|\frac{\partial \tilde{T}_2^x}{\partial x} \circ \Phi\right| \mathrm{d}\tilde{\mu}_1 = \int_{\tilde{I}_2} \log\left|\frac{\partial \tilde{T}_2^x}{\partial x}\right| \mathrm{d}\tilde{\mu}_2 := m.$$

Let us define $g_1 := \log|\frac{\partial \tilde{T}_1^x}{\partial x}|$ and $g_2 := \log|\frac{\partial \tilde{T}_2^x}{\partial x}|$ (so that $\int_{\tilde{I}_1} g_1\, \mathrm{d}\tilde{\mu}_1 = \int_{\tilde{I}_2} g_2\, \mathrm{d}\tilde{\mu}_2 = 0$) and $S_N^T g := \sum_{k=0}^{N-1} g \circ T^k$; we easily see that

$$S_N^{\tilde{T}_1} g_1 = S_N^{\tilde{T}_2} g_1 \circ \Phi \log\left|\frac{\partial \Phi^x}{\partial x} \circ \tilde{T}_1^k\right| - \log\left|\frac{\partial \Phi^x}{\partial x} \circ \tilde{T}_1^{k+1}\right|,$$

which means that $S_N^{\tilde{T}_1} g_1$ and $S_N^{\tilde{T}_2} g_2 \circ \Phi$ differ by a coboundary.

Lemma A.3. *Let u, v be two observables such that*

1. $\lim_{N\to+\infty} \int (\frac{S_N v}{\sqrt{N}})^2\, \mathrm{d}\mu = l \in \mathbb{R}$;
2. $u = v + (f - f \circ T)$ *for some* $f \in L^2$.

Then

$$\lim_{N\to+\infty} \int \left(\frac{S_N v}{\sqrt{N}}\right)^2 \mathrm{d}\mu = \lim_{N\to+\infty} \int \left(\frac{S_N u}{\sqrt{N}}\right)^2 \mathrm{d}\mu.$$

The lemma implies

$$\lim_{N\to+\infty} \int_{\tilde{I}_1} \left(\frac{S_N^{\tilde{T}_1} g_1}{\sqrt{N}}\right)^2 \mathrm{d}\tilde{\mu}_1 = \lim_{N\to+\infty} \int_{\tilde{I}_2} \left(\frac{S_N^{\tilde{T}_2} g_2}{\sqrt{N}}\right)^2 \mathrm{d}\tilde{\mu}_2. \tag{13}$$

This information can be translated back to the original systems: since $\frac{\partial \tilde{T}_1^x}{\partial x}|_{(x,y)} = T_1'(x)$, $\frac{\partial \tilde{T}_2^x}{\partial x}|_{(x,y)} = T_2'(x)$ if we define

$$G_1 := \log|T_1'(x)| - \int_{I_1} \log|T_1'(x)|\,\mathrm{d}\mu_1,$$

$$G_2 = \log|T_2'(x)| - \int_{I_2} \log|T_2'(x)|\,\mathrm{d}\mu_2,$$

we get $g_1(x, y) = G_1(x)$ and $g_2(x, y) = G_2(x)$; therefore $S_N^{\tilde{T}_1} g_1 = S_N^{T_1} G_1$ and $S_N^{\tilde{T}_2} g_2 = S_N^{T_2} G_2$. Finally, by equation (13), we get

$$\lim_{N\to+\infty} \int_{I_1} \left(\frac{S_N^{T_1} G_1}{\sqrt{N}}\right)^2 \mathrm{d}\mu_1 = \lim_{N\to+\infty} \int_{I_2} \left(\frac{S_N^{T_2} G_2}{\sqrt{N}}\right)^2 \mathrm{d}\mu_2.$$

A.1. Tables

Table A.1. A chain of adjacent matching intervals (see section 4.2).

$(k_1\ k_2)$	Size	$(\alpha^-,$	$\alpha^+)$
(257 257)	5.43e-201	(...........,	...........)
(129 129)	7.27e-101	(...........,	...........)
(65 65)	7.98e-51	(...........,	...........,)
(33 33)	8.81e-26	$\Big(\{-1051803916417 + 5\sqrt{110424870216034832616745}\}/,\ 1576491320449$	$-1+\dfrac{\sqrt{31529826409}}{128045}\Big)$
(17 17)	2.78e-13	$\left(-1+\dfrac{\sqrt{31529826409}}{128045},\right.$	$\left.\dfrac{-433+\sqrt{467857}}{649}\right)$
(9 9)	5.2e-7	$\left(\dfrac{-433+\sqrt{467857}}{649},\right.$	$\left.\dfrac{-13+5\sqrt{13}}{13}\right)$
(5 5)	6.75e-4	$\left(\dfrac{-13+5\sqrt{13}}{13},\right.$	$\left.\dfrac{-2+\sqrt{10}}{3}\right)$
(3 3)	2.68e-2	$\left(\dfrac{-2+\sqrt{10}}{3},\right.$	$\left.-1+\sqrt{2}\right)$
(2 2)	2.04e-1	$\left(-1+\sqrt{2},\right.$	$\left.\dfrac{-1+\sqrt{5}}{2}\right)$

Table A.2. A sample of matching intervals found as in section 4.

$(k_1\ k_2)$	Size	$(\alpha^-,$	$\alpha^+)$	$(k_1\ k_2)$	Size	$(\alpha^-,$	$\alpha^+)$
(3 9)	7.69e-4	$\left(\dfrac{-8+\sqrt{82}}{9},\right.$	$\left.\dfrac{-2+\sqrt{5}}{2}\right)$	(8 6)	6.42e-5	$\left(\dfrac{-33+\sqrt{2305}}{64},\right.$	$\left.\dfrac{-77+\sqrt{7221}}{34}\right)$
(2 8)	3.68e-3	$\left(-4+\sqrt{17},\right.$	$\left.\dfrac{-7+\sqrt{77}}{14}\right)$	(5 5)	1.46e-3	$\left(\dfrac{-7+\sqrt{101}}{13},\right.$	$\left.-2+\sqrt{5}\right)$
(3 8)	1.11e-3	$\left(\dfrac{-7+\sqrt{65}}{8},\right.$	$\left.\dfrac{-7+3\sqrt{7}}{7}\right)$	(2 4)	2.77e-2	$\left(-2+\sqrt{5},\right.$	$\left.\dfrac{-3+\sqrt{21}}{6}\right)$

(2 7)	5.44e-3	$\left(\frac{-7+\sqrt{53}}{2}, \frac{-3+\sqrt{15}}{6}\right)$	(3 6)	2.1e-3	$\left(\frac{-7+\sqrt{65}}{4}, \frac{-6+4\sqrt{5}}{11}\right)$
(3 8)	6.98e-4	$\left(\frac{-19+\sqrt{445}}{14}, \frac{-9+2\sqrt{30}}{13}\right)$	(4 6)	7.02e-4	$\left(\frac{-11+\sqrt{226}}{15}, \frac{-23+3\sqrt{93}}{22}\right)$
(3 7)	1.69e-3	$\left(\frac{-6+5\sqrt{2}}{7}, \frac{-3+2\sqrt{3}}{3}\right)$	(3 5)	3.97e-3	$\left(\frac{-5+\sqrt{37}}{4}, \frac{-9+\sqrt{165}}{14}\right)$
(4 7)	8.12e-4	$\left(\frac{-17+\sqrt{445}}{26}, \frac{-3+\sqrt{11}}{2}\right)$	(4 6)	5.77e-4	$\left(\frac{-13+\sqrt{257}}{11}, \frac{-2+2\sqrt{2}}{3}\right)$
(2 6)	8.54e-3	$\left(-3+\sqrt{10}, \frac{-5+3\sqrt{5}}{10}\right)$	(4 5)	1.51e-3	$\left(\frac{-15+\sqrt{445}}{22}, \frac{-8+3\sqrt{11}}{7}\right)$
(3 8)	6.06e-4	$\left(\frac{-11+\sqrt{145}}{6}, \frac{-10+2\sqrt{42}}{17}\right)$	(5 5)	7.88e-4	$\left(\frac{-10+\sqrt{226}}{18}, \frac{-23+5\sqrt{29}}{14}\right)$
(3 7)	1.12e-3	$\left(\frac{-8+\sqrt{82}}{6}, \frac{-15+\sqrt{357}}{22}\right)$	(3 4)	1.02e-2	$\left(\frac{-3+\sqrt{17}}{4}, \frac{-3+\sqrt{15}}{3}\right)$
(3 6)	2.76e-3	$\left(\frac{-5+\sqrt{37}}{6}, \frac{-5+\sqrt{35}}{5}\right)$	(4 6)	8.86e-4	$\left(\frac{-11+\sqrt{170}}{7}, \frac{-19+3\sqrt{93}}{34}\right)$
(4 6)	1.34e-3	$\left(\frac{-7+\sqrt{82}}{11}, \frac{-15+\sqrt{285}}{10}\right)$	(4 5)	1.78e-3	$\left(\frac{-15+\sqrt{365}}{14}, \frac{-7+3\sqrt{11}}{10}\right)$
(9 7)	2.38e-5	$\left(\frac{-51+13\sqrt{29}}{100}, \frac{-117+\sqrt{15621}}{42}\right)$	(5 5)	7.09e-4	$\left(\frac{-11+\sqrt{257}}{17}, \frac{-6+2\sqrt{14}}{5}\right)$
(5 6)	7.91e-4	$\left(\frac{-9+\sqrt{145}}{16}, \frac{-10+2\sqrt{30}}{5}\right)$	(8 6)	2.73e-5	$\left(\frac{-54+\sqrt{7057}}{101}, \frac{-127+7\sqrt{453}}{74}\right)$
(9 7)	2.25e-5	$\left(\frac{-53+\sqrt{5185}}{99}, \frac{-30+4\sqrt{66}}{13}\right)$	(4 4)	5.24e-3	$\left(\frac{-4+\sqrt{37}}{7}, \frac{-3+\sqrt{13}}{2}\right)$
(10 7)	1.54e-5	$\left(\frac{-127+\sqrt{30629}}{250}, \frac{-73+\sqrt{6083}}{26}\right)$	(8 6)	2.73e-5	$\left(\frac{-54+\sqrt{7057}}{101}, \frac{-127+7\sqrt{453}}{74}\right)$
(2 5)	1.45e-2	$\left(\frac{-5+\sqrt{29}}{2}, \frac{-1+\sqrt{2}}{2}\right)$	(2 3)	6.32e-2	$\left(\frac{-3+\sqrt{13}}{2}, \frac{-1+\sqrt{3}}{2}\right)$
(3 8)	6.57e-4	$\left(\frac{-23+\sqrt{629}}{10}, \frac{-10+\sqrt{195}}{19}\right)$	(4 6)	6.9e-4	$\left(\frac{-13+\sqrt{290}}{11}, \frac{-23+\sqrt{1365}}{38}\right)$
(3 7)	1.06e-3	$\left(\frac{-9+\sqrt{101}}{5}, \frac{-4+\sqrt{30}}{7}\right)$	(4 5)	1.72e-3	$\left(\frac{-15+\sqrt{533}}{22}, \frac{-4+\sqrt{30}}{4}\right)$
(3 6)	1.98e-3	$\left(\frac{-13+\sqrt{229}}{10}, \frac{-2+\sqrt{7}}{3}\right)$	(3 4)	9.87e-3	$\left(\frac{-7+\sqrt{85}}{6}, \frac{-3+2\sqrt{6}}{5}\right)$
(4 6)	7.42e-4	$\left(\frac{-10+\sqrt{170}}{14}, \frac{-7+\sqrt{69}}{6}\right)$	(4 5)	1.45e-3	$\left(\frac{-9+\sqrt{145}}{8}, \frac{-8+2\sqrt{42}}{13}\right)$
(9 7)	1.03e-5	$\left(\frac{-81+\sqrt{13226}}{155}, \frac{-187+3\sqrt{4669}}{82}\right)$	(4 4)	3.82e-3	$\left(\frac{-5+\sqrt{65}}{8}, \frac{-11+\sqrt{221}}{10}\right)$
(3 5)	4.94e-3	$\left(\frac{-4+\sqrt{26}}{5}, \frac{-2+\sqrt{6}}{2}\right)$	(5 5)	6.75e-4	$\left(\frac{-13+5\sqrt{13}}{13}, \frac{-2+\sqrt{10}}{3}\right)$
(4 6)	8.44e-4	$\left(\frac{-10+\sqrt{145}}{9}, \frac{-19+3\sqrt{69}}{26}\right)$	(3 3)	2.68e-2	$\left(\frac{-2+\sqrt{10}}{3}, -1+\sqrt{2}\right)$
(7 6)	1.11e-4	$\left(\frac{-25+\sqrt{1297}}{48}, \frac{-29+\sqrt{1023}}{13}\right)$	(2 2)	2.04e-1	$\left(-1+\sqrt{2}, \frac{-1+\sqrt{5}}{2}\right)$
(4 5)	2.45e-3	$\left(\frac{-11+\sqrt{229}}{18}, \frac{-3+2\sqrt{3}}{2}\right)$	(2 1)	3.82e-1	$\left(\frac{-1+\sqrt{5}}{2}, 1\right]$

References

[1] Broise A 1996 Transformations dilatantes de l'intervalle et théorèmes limites *Astérisque* **238** 1–109

[2] Carminati C and Tiozzo G 2010 A canonical thickening of Q and the dynamics of continued fractions arXiv:1004.3790

[3] Charves M 1877 Démonstration de la périodicité des fractions continues, engendrées par les racines d'une èquation du deuxième degré *Bull. Sci. Math. (2)* **1** 41–3

[4] Graham R L, Knuth D E and Patashnik O 1994 *Concrete Mathematics* (Reading, MA: Addison-Wesley)

[5] Kraaikamp C 1991 A new class of continued fraction expansions *Acta Arith.* **57** 1–39

[6] Hartono Y and Kraaikamp C 2002 A note on Hurwitzian numbers *Tokyo J. Math.* **25** 353–62

[7] Luzzi L and Marmi S 2008 On the entropy of Japanese continued fractions *Discrete Contin. Dyn. Syst.* **20** 673–711 (arXiv:math.DS/0601576v2)

[8] Moussa P, Cassa A and Marmi S 1999 Continued fractions and Brjuno functions *J. Comput. Appl. Math.* **105** 403–15

[9] Nakada H 1981 Metrical theory for a class of continued fraction transformations and their natural extensions *Tokyo J. Math.* **4** 399–426

[10] Nakada H and Natsui R 2008 The non-monotonicity of the entropy of α-continued fraction transformations *Nonlinearity* **21** 1207–25

[11] Rohlin V A 1961 Exact endomorphisms of a Lebesgue space *Izv. Akad. Nauk SSSR Ser. Mat.* **25** 499–530
Rohlin V A 1964 *Am. Math. Soc. Transl. (2)* **39** 1–36 (Engl. Transl.)

[12] Rychlik M 1983 Bounded variation and invariant measures *Studia Math.* **76** 69–80

[13] Schweiger F 1995 *Ergodic Theory of Fibered Systems and Metric Number Theory* (Oxford: Oxford Science Publications/Clarendon Press)

[14] Tiozzo G 2009 The entropy of α-continued fractions: analytical results arXiv:0912.2379

[15] Viana M 1997 *Stochastic Dynamics of Deterministic Systems* (*Lecture Notes 21st Brazilian Mathematics Colloquium (IMPA), Rio de Janeiro*)

[16] Zweimüller R 1998 Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points *Nonlinearity* **11** 1263–76