Ergodic and Diophantine properties
of algorithms of Selmer type

by

FRITZ SCHWEIGER (Salzburg)

1. Introduction. The important work Lagarias (1993) has renewed the interest in Diophantine approximation by multidimensional continued fractions and their ergodic properties. The connections between multidimensional continued fractions and ergodic theory are given a careful study in Broise-Alamichel & Guivarc'h (2001). A further survey of recent results is given in Schweiger (2000, 2002).

In this paper a two-dimensional continued fraction will be defined as follows.

DEFINITION 1. Let B be a subset of the Euclidean plane $\mathbb{R}^2$. A map $T : B \to B$ is said to generate a two-dimensional continued fraction if there is a partition $\{B(k) : k \in I\}$ of B and a set of invertible matrices $\{\alpha(k) \in \mathrm{GL}(3, \mathbb{Z}) : k \in I\}$, where I is finite or countable and $\alpha(k) = ((A_{ij}(k)))$, such that

$$(Tx)_i = y_i = \frac{A_{i0} + A_{i1}x_1 + A_{i2}x_2}{A_{00} + A_{01}x_1 + A_{02}x_2}, \quad 1 \le i \le 2,$$

maps every cell $B(k)$ onto a subset $TB(k)$ of B. Then the fibred system (B, T) is called a *two-dimensional continued fraction*. The partition is called the *time-1-partition*.

The most famous example is given by the Jacobi algorithm which is a straightforward generalization of regular continued fractions. It is defined piecewise as

$$T(x_1, x_2) = \left(\frac{x_2}{x_1} - \left[\frac{x_2}{x_1} \right], \frac{1}{x_1} - \left[\frac{1}{x_1} \right] \right).$$

DEFINITION 2. Let $\beta(k) = ((B_{ij}(k)))$ denote the inverse matrices of $\alpha(k)$, $k \in I$. Then we define

$$\beta(k_1, \ldots, k_s) := \beta(k_1, \ldots, k_{s-1})\beta(k_s) = ((B_{ij}^{(s)})).$$

Then $y = T^s x$ is equivalent to

$$x_i = \frac{B_{i0}^{(s)} + B_{i1}^{(s)} y_1 + B_{i2}^{(s)} y_2}{B_{00}^{(s)} + B_{01}^{(s)} y_1 + B_{02}^{(s)} y_2}, \quad 1 \le i \le 2,$$

for some matrix $\beta(k_1, \ldots, k_s)$.

The matrices $\beta(k_1, \ldots, k_s) = ((B_{ij}^{(s)}))$ produce rational approximations

$$\left(\frac{B_{1g}^{(s)}}{B_{0g}^{(s)}}, \frac{B_{2g}^{(s)}}{B_{0g}^{(s)}} \right), \quad 0 \le g \le 2.$$

Therefore the problem arises to estimate the differences

$$\left| x_i - \frac{B_{ig}^{(s)}}{B_{0g}^{(s)}} \right|, \quad 1 \le i \le 2, \, 0 \le g \le 2.$$

In this paper we will show that Selmer's algorithm (Section 2) produces "good" approximations for almost all points (x_1, x_2) in the following sense. There is a constant $\delta > 0$ such that for almost all (x_1, x_2) the inequalities

$$\left| x_i - \frac{B_{i1}^{(s)}}{B_{01}^{(s)}} \right| \ll \frac{1}{(B_{01})^{1+\delta}}, \quad 1 \le i \le 2,$$

are valid for $s \ge s(x_1, x_2)$. The result for Selmer's algorithm is an improvement of Lagarias (1993).

In Section 3 the fully subtractive algorithm and a new multiplicative version obtained by parallel division are considered. These algorithms are not convergent but exhibit interesting ergodic properties, namely absorbing regions and invariant measures. Since the transient regions of Selmer's algorithm and the fully subtractive algorithm are complementary sets, in Section 4 the interesting question is taken up what will happen when one mixes both. The new algorithm has a Cantor-like exceptional set with positive Lebesgue measure. The existence of such sets was first detected as a strange feature of the Parry–Daniels map (Schweiger (1981), Nogueira (1995)).

We will use the notations $K^3 := \{b = (b_0, b_1, b_2) : b_0 \ge b_1 \ge b_2 \ge 0\}$, $B^2 := \{x = (x_1, x_2) : 1 \ge x_1 \ge x_2 \ge 0\}$ and the projection map $p : K^3 \to B^2$, defined as $p(b_0, b_1, b_2) := \left(\frac{b_1}{b_0}, \frac{b_2}{b_0} \right)$.

2. Selmer's algorithm

DEFINITION 3. Let $K^3 = \{b = (b_0, b_1, b_2) : b_0 \ge b_1 \ge b_2 \ge 0\}$ as defined before. Then define $\sigma b := (b_0 - b_2, b_1, b_2)$. There are matrices $\alpha(i)$, $0 \le i \le 2$, such that the permuted triple $\pi \sigma b = \alpha(i)$ is in K^3. The map S which makes

the diagram

$$
\begin{array}{ccc}
K^3 & \xrightarrow{\ \pi\sigma\ } & K^3 \\
\downarrow{\scriptstyle p} & & \downarrow{\scriptstyle p} \\
B^2 & \xrightarrow[\ S\ ]{} & B^2
\end{array}
$$

commutative is said to generate *Selmer's algorithm*.

We list the three matrices $\alpha(i)$ and their inverses $\beta(i)$, $0 \le i \le 2$:

$$
\alpha(0) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad
\alpha(1) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad
\alpha(2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix},
$$

$$
\beta(0) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad
\beta(1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad
\beta(2) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.
$$

The 3 cells of the time-1-partition are denoted as $\Delta(i)$, $0 \le i \le 2$. It is well known that $D := \{x : x_1 + x_2 \ge 1\}$ is an absorbing set for this algorithm. Note that $D = \Delta(1) \cup \Delta(2)$ and $SD = D$. Since $\alpha^k(0)b = (b_0 - kb_2, b_1, b_2)$ we see that if $k := \left[\frac{b_0 - b_1}{b_2}\right]$ then $S^k x \in D$. Furthermore S restricted to D admits the invariant density

$$
g(x_1, x_2) = \frac{1}{x_1 x_2}.
$$

The properties of Selmer's algorithm can be well understood if one considers the jump transformation $R : D \to D$ with respect to the set $\Delta(2) = \{(x_1, x_2) : 1 \le 2x_2\}$ (the technique of jump transformation is explained in Schweiger (1995) or (2000)). To ensure that the map R is defined almost everywhere on D we must show that the orbit of almost every point $(x_1, x_2) \in D$ eventually enters $\Delta(2)$. The residual set D_n is the cylinder $\Delta(\underbrace{1, \ldots, 1}_{n \text{ times}})$ which has the following vertices.

$$
n = 2k+1: \quad (1,0), \ \left(\frac{k+1}{k+2}, \frac{1}{k+2}\right), \ \left(1, \frac{1}{k+2}\right),
$$

$$
n = 2k+2: \quad (1,0), \ \left(1, \frac{1}{k+2}\right), \ \left(\frac{k+2}{k+3}, \frac{1}{k+3}\right).
$$

Clearly,

$$
\operatorname{diam} \Delta(\underbrace{1, \ldots, 1}_{n \text{ times}}) \ll \frac{1}{n} \quad (n \to \infty)
$$

and $\lambda(D_n) \downarrow 0$. Therefore the jump transformation is defined almost everywhere. In this way a new two-dimensional algorithm appears which is called *Baldwin's algorithm*. It is easy to show that R is ergodic and has the invariant density

$$f(x_1, x_2) = \frac{1}{x_1(x_1 + x_2)}.$$

In Schweiger (2001) it is shown that the approximations $B_{01}^{(s)}, B_{11}^{(s)}, B_{21}^{(s)}$ generated by Baldwin's algorithm satisfy

$$\left| x_i - \frac{B_{i1}^{(s)}}{B_{01}^{(s)}} \right| \ll \frac{1}{B_{01}^{(s)}}, \quad i = 1, 2,$$

for all points $x = (x_1, x_2)$ and furthermore there is a constant $\delta > 0$ such that

$$\left| x_i - \frac{B_{i1}^{(s)}}{B_{01}^{(s)}} \right| \ll \frac{1}{(B_{01}^{(s)})^{1+\delta}}, \quad i = 1, 2,$$

is true for almost all points $x = (x_1, x_2)$.

We will show that this result can be extended to Selmer's algorithm.

To distinguish both algorithms we denote the entries of the matrices which belong to Selmer's algorithm by $G_{ij}^{(s)}$, i.e.

$$((G_{ij}^{(s)})) = \beta(\varepsilon_1, \ldots, \varepsilon_s), \quad \varepsilon_j \in \{1, 2\}, \ 1 \le j \le s.$$

THEOREM 1. *For Selmer's algorithm there is a constant $\delta > 0$ such that for almost all $(x_1, x_2) \in D$ the inequalities*

$$\left| x_1 - \frac{G_{11}^{(N)}}{G_{01}^{(N)}} \right| \ll \frac{1}{(G_{01}^{(N)})^{1+\delta}}, \quad \left| x_2 - \frac{G_{21}^{(N)}}{G_{01}^{(N)}} \right| \ll \frac{1}{(G_{01}^{(N)})^{1+\delta}}$$

are valid for $N \ge N(x_1, x_2)$.

Proof. We recall that $\varepsilon_{k+1}(x_1, x_2) = j$ if $S^k(x_1, x_2) \in \Delta(j)$. If $e(x_1, x_2) := \min\{k \ge 0 : \varepsilon_{k+1}(x_1, x_2) = 2\}$ then Baldwin's algorithm is given by the equation

$$R(x_1, x_2) := S^{e(x_1, x_2)+1}(x_1, x_2).$$

If we denote as before the corresponding matrix elements for Baldwin's algorithm by $B_{ij}^{(s)}$ then

$$B_{ij}^{(s)} = G_{ij}^{(N)}$$

if $\varepsilon_N = 2$ and

$$s = \sum_{k=0}^{N-1} c_{\Delta(2)}(S^k(x_1, x_2)).$$

Here c_E denotes the indicator function of a set E.

Note that

$$G_{i0}^{(g+1)} = G_{i1}^{(g)}, \quad G_{i1}^{(g+1)} = G_{i0}^{(g)}, \quad G_{i2}^{(g+1)} = G_{i0}^{(g)} + G_{i2}^{(g)} \quad \text{if } \varepsilon_{g+1} = 1,$$
$$G_{i0}^{(g+1)} = G_{i1}^{(g)}, \quad G_{i1}^{(g+1)} = G_{i0}^{(g)} + G_{i2}^{(g)}, \quad G_{i2}^{(g+1)} = G_{i0}^{(g)} \quad \text{if } \varepsilon_{g+1} = 2.$$

Therefore, if $\varepsilon_N = 2$ is followed by a block $\varepsilon_{N+1} = \cdots = \varepsilon_{N+j} = 1$, $0 \le j$, then

$$G_{i0}^{(N)} = G_{i1}^{(N+1)} = G_{i1}^{(N+3)} = \cdots = G_{i1}^{(N+t)}, \quad t \equiv 1 \bmod 2, \ t \le j,$$
$$G_{i1}^{(N)} = G_{i1}^{(N+2)} = G_{i1}^{(N+4)} = \cdots = G_{i1}^{(N+t)}, \quad t \equiv 0 \bmod 2, \ t \le j.$$

This means that $\dfrac{G_{i1}^{(N+t)}}{G_{01}^{(N+t)}}$ equals $\dfrac{G_{i1}^{(N)}}{G_{01}^{(N)}} = \dfrac{B_{i1}^{(s)}}{B_{01}^{(s)}}$ or $\dfrac{G_{i0}^{(N)}}{G_{00}^{(N)}} = \dfrac{B_{i0}^{(s)}}{B_{00}^{(s)}}$ for $1 \le t \le j$. The conditions for Baldwin's algorithm imply that $B_{i0}^{(s)} = B_{i0}^{(s-1)}$ or $B_{i0}^{(s)} = B_{i1}^{(s-1)}$ (see Schweiger (2001), Lemma 2). Therefore it is easy to see that the result given in Schweiger (2001) remains valid if one replaces $B_{i1}^{(s)}$ by $B_{i0}^{(s)}$ and the result follows from Baldwin's algorithm.

REMARKS. (1) A block of consecutive digits $\varepsilon = 1$ does not increase the growth of $G_{00}^{(N)}$ and $G_{01}^{(N)}$. In Schweiger (2001) the result

$$\limsup_{s \to \infty} \frac{\log B_{01}^{(s)}}{s} \le \Theta < \infty$$

is used. It is easy to show that even

$$\lim_{s \to \infty} \frac{\log B_{01}^{(s)}}{s} = \frac{h(R)}{3}$$

exists a.e., where $h(R)$ is the entropy of the map R.

If

$$c_0 := \int_D \frac{dx_1 \, dx_2}{x_1 x_2}$$

then we define

$$\kappa := \frac{1}{c_0} \int_{\Delta(2)} \frac{dx_1 \, dx_2}{x_1 x_2}$$

and we obtain

$$\lim_{N \to \infty} \frac{\log G_{01}^{(N)}}{N} = \kappa \lim_{s \to \infty} \frac{\log B_{01}^{(s)}}{s}.$$

For the entropies of S and R this means

$$h(S) = \kappa h(R).$$

(2) The following multiplicative acceleration of Selmer's algorithm seems to be straightforward:

$$\delta b := (b_0 - mb_2, b_1, b_2), \quad m = \left[\frac{b_0}{b_2}\right].$$

Permutation and projection gives the algorithm

$$M(x_1, x_2) = \left(\frac{x_2}{x_1}, \frac{1 - mx_2}{x_1}\right)$$

with cells

$$Z(m) = \left\{x \in B^2 : \frac{1}{m+1} < x_2 \leq \frac{1}{m}\right\}, \quad k = 1, 2, \ldots.$$

It is easy to see that $MZ(m)$ is the convex set with vertices $\left(\frac{1}{m}, 0\right)$, $\left(\frac{1}{m+1}, \frac{1}{m+1}\right)$, $(1, 1)$, and $(1, 0)$. Therefore no cylinder is full and $MZ(m)$ is not a union of cylinders of rank 1.

This algorithm has bad Diophantine approximation properties. This can be illustrated with periodic algorithms of period length 1. Let $\varrho_0 > 1$ be the largest root of $\varrho^3 - m\varrho - 1 = 0$. Then the pair $(\xi, \eta) = (1/\varrho_0, 1/\varrho_0^2)$ has the periodic expansion $m_k = m$ for all $k \geq 1$. But for $m \geq 3$ the second root ϱ_1 satisfies $\varrho_1 < -1$, i.e. $|\varrho_1| > 1$. Since $\xi B_{00}^{(s)} - B_{10}^{(s)} \sim \varrho_1^s$ we see that

$$\liminf_{s \to \infty} |\xi B_{00}^{(s)} - B_{10}^{(s)}| = \infty.$$

3. The Fully Subtractive Algorithm

DEFINITION 4. We define $\tau b := (b_0 - b_2, b_1 - b_2, b_2)$. Again there are matrices $\gamma(i)$, $0 \leq i \leq 2$, such that the permuted triple $\pi\tau b = \gamma(i)b$ is in K^3. Projection on B^2 generates a map $T : B^2 \to B^2$ which is said to yield the *Fully Subtractive Algorithm*.

The complete list of the matrices $\gamma(i)$ and their inverses $\delta(i)$, $0 \leq i \leq 2$, is the following:

$$\gamma(0) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \gamma(1) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \gamma(2) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix},$$

$$\delta(0) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \delta(1) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \delta(2) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

The 3 cells of the time-1-partition are denoted as $E(i)$, $0 \leq i \leq 2$. The

following maps are the inverse branches of T:

$$W(0) = \left(\frac{x_1 + x_2}{1 + x_2}, \frac{x_2}{1 + x_2} \right),$$

$$W(1) = \left(\frac{x_1 + x_2}{1 + x_1}, \frac{x_1}{1 + x_1} \right),$$

$$W(2) = \left(\frac{1 + x_2}{1 + x_1}, \frac{1}{1 + x_1} \right).$$

Again, the line $x_1 + x_2 = 1$ plays a crucial role for this algorithm. It is known that in this case the set $E := \{x \in B^2 : x_1 + x_2 \leq 1\}$ is absorbing, i.e. $TE = E$ and for almost every $x \in B^2$ there is an $N = N(x)$ such that $T^N x \in E$. Furthermore T restricted to E admits the invariant density

$$h(x_1, x_2) = \frac{1}{x_1 x_2 (1 - x_1 - x_2)}.$$

However, this result is much more difficult to obtain (Meester & Nowicki (1989)). The segment $x_1 + x_2 = 1$, $0 \leq x_2 \leq x_1 \leq 1$, is fixed under T. The restriction of T to this segment is isomorphic to the map $\tau : [0, 1/2] \to [0, 1/2]$ defined by

$$\tau(t) = \begin{cases} t/(1-t), & 0 \leq t \leq 1/3, \\ 1 - t/(1-t), & 1/3 \leq t \leq 1/2. \end{cases}$$

Furthermore

$$E = \{x : T^j x \in E(0) \cup E(1) \text{ for all } j \geq 0\}.$$

REMARK. The dynamics of the system (E, T) can be completely reduced to the dynamics of continued fractions as follows. Let $\psi(x_1, x_2) := (x_2/x_1, (1 - x_1 - x_2)/x_1)$; then we see that $\psi E = \{(u, v) : 0 \leq u \leq 1, 0 \leq v\}$. The map T is therefore isomorphic to the map

$$\Theta(u, v) = \begin{cases} \left(\dfrac{u}{1-u}, \dfrac{v}{1-u} \right), & 0 \leq u \leq 1/2, \\ \left(\dfrac{1-u}{u}, \dfrac{v}{u} \right), & 1/2 \leq u \leq 1 \end{cases}$$

with the invariant density

$$\chi(u, v) = \frac{1}{uv}.$$

If we form the jump transformation with respect to the set $\{(u, v) : 1/2 \leq u \leq 1\}$ we obtain the map

$$P(u, v) = \left(\frac{1}{u} - a, \frac{v}{u} \right), \qquad \frac{1}{a+1} < u \leq \frac{1}{a},$$

with the invariant density

$$\pi(u, v) = \frac{1}{(1+u)v}.$$

Note that the Jacobians of the inverse branches of P with respect to the measure

$$\mu(X) = \iint_X \frac{du\,dv}{(1+u+v)^3}$$

are given as

$$\kappa(a_1, \ldots, a_n; u, v) = \frac{1}{(q_n + q_{n-1}u + v)^3},$$

where $q_k = a_k q_{k-1} + q_{k-2}$ as in the continued fraction algorithm.

From this picture it is clear that the algorithm is not convergent. The fixed points of T restricted to E are $(0,0)$ and the periodic points on the segment $x_1 + x_2 = 1$ (or equivalently the points $(u,0)$ with u having a periodic continued fraction). Therefore this algorithm cannot produce "good" Diophantine approximations.

One can also consider the multiplicative acceleration defined by the parallel division process $\delta b := (b_0 - mb_2, b_1 - mb_2, b_2)$, where $m = [b_1/b_2]$. It is easy to verify that the associated map is given piecewise as

$$Rx = \begin{cases} \left(\dfrac{x_2}{1 - mx_2}, \dfrac{x_1 - mx_2}{1 - mx_2} \right) & \text{on } \Delta(\alpha; m), \\[2mm] \left(\dfrac{1 - mx_2}{x_2}, \dfrac{x_1 - mx_2}{x_2} \right) & \text{on } \Delta(\beta; m), \end{cases}$$

where $\Delta(\alpha; m)$ is the triangle with vertices $(0,0)$, $\left(\frac{m}{1+m}, \frac{1}{1+m} \right)$, $\left(1, \frac{1}{1+m} \right)$ and $\Delta(\beta; m)$ is the triangle with vertices $\left(1, \frac{1}{m} \right)$, $\left(\frac{m}{1+m}, \frac{1}{1+m} \right)$, $\left(1, \frac{1}{1+m} \right)$.

THEOREM 2. *R is the jump transformation of T with respect to the set $E(1) \cup E(2)$. Therefore the set E is absorbing for R and*

$$k(x_1, x_2) = \frac{1}{x_1(x_1 + x_2)(1 - x_1 - x_2)}$$

is an invariant density for R restricted to E.

Proof. Iteration of T on $E(3)$ gives

$$T^m x = \left(\frac{x_1 - mx_2}{1 - mx_2}, \frac{x_2}{1 - mx_2} \right).$$

If $T^m x \in E(1)$ we obtain

$$Rx = T^{m+1} x = \left(\frac{x_2}{1 - (m+1)x_2}, \frac{x_1 - (m+1)x_2}{1 - (m+1)x_2} \right).$$

If $T^m x \in E(2)$ we obtain

$$Rx = T^{m+1} x = \left(\frac{1 - (m+1)x_2}{x_2}, \frac{x_1 - (m+1)x_2}{x_2} \right).$$

The set E is absorbing for T and therefore absorbing for R. But the restriction of R to E is the jump transformation of the restriction of T to E with respect to the set $E \cap E(2)$. Therefore

$$k(x_1, x_2) = h\left(\frac{x_1 + x_2}{1 + x_1}, \frac{x_1}{1 + x_1} \right) \frac{1}{(1 + x_1)^3} = \frac{1}{x_1(x_1 + x_2)(1 - x_1 - x_2)}.$$

However, Kuzmin's equation can also be verified directly:

$$\sum_{m=1}^{\infty} \frac{1}{(mx_1 + x_2)((m+1)x_1 + x_2)(1 - x_1 - x_2)}$$

$$= \sum_{m=1}^{\infty} \frac{1}{x_1} \left(\frac{1}{mx_1 + x_2} - \frac{1}{(m+1)x_1 + x_2} \right) \frac{1}{1 - x_1 - x_2}$$

$$= \frac{1}{x_1(x_1 + x_2)(1 - x_1 - x_2)}.$$

LEMMA 1. *If E is an absorbing set for T then E is an absorbing set for any jump transformation R stemming from T.*

Proof. For almost every $x \in B$ there is an $N = N(x)$ such that $T^n x \in E$ for $n \geq N$. Since $(R^n x)$ is a subsequence of $(T^n x)$, clearly E is absorbing for R.

REMARK. The converse statement is not true (even if R is ergodic). An easy counterexample is

$$Tx = \begin{cases} 2x, & 0 \leq x < 1/2, \\ x - 1/2, & 1/2 \leq x < 1. \end{cases}$$

Consider the jump transformation R with respect to the subset $B_1 = [1/2, 1[$. One calculates that

$$Rx = 2^k x - \frac{1}{2}, \quad \frac{1}{2^{k+1}} \leq x < \frac{1}{2^k}, \ k = 0, 1, 2, \ldots.$$

So R is ergodic on $[0, 1[$ and $E = [0, 1/2[$ is an absorbing set.

4. A strange exceptional set. Both Selmer's algorithm and the fully subtractive version have absorbing regions. However, for Selmer's algorithm S this region is $D = \{(x_1, x_2) \in B^2 : x_1 + x_2 \geq 1\}$, while for the fully subtractive version T it is its complement $E = \{(x_1, x_2) \in B^2 : x_1 + x_2 \leq 1\}$. This behaviour leads to the interesting question what will be the properties of the compositions $S \circ T$ and $T \circ S$.

We first consider $(S \circ T)x = S(Tx)$. The 9 cells of the time-1-partition can be calculated from the compositions $W(i)\,V(j)$, or equivalently from the matrices $\delta(i)\,\beta(j)$, $1 \le i, j \le 3$. The complete list is given as follows:

$$\delta(0)\beta(0) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \delta(0)\beta(1) = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\delta(0)\beta(2) = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \delta(1)\beta(0) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\delta(1)\beta(1) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \delta(1)\beta(2) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$\delta(2)\beta(0) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix}, \qquad \delta(2)\beta(1) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix},$$

$$\delta(2)\beta(2) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The associated cylinders of the time-1-partition will be denoted by $[i\ j]$.

Only the 3 cylinders $[0\,0], [1\,0], [2\,0]$ are full. The other 6 cylinders are mapped onto the triangle D which itself is the union of 7 cylinders, namely $[2\,0], [0\,1], [0\,2], [1\,1], [1\,2], [2\,1], [2\,2]$. The triangle E is the union of $[0\,0]$ and $[1\,0]$.

A similar calculation can be made for $(T \circ S)x = T(Sx)$. The associated cylinders will be denoted by $\{i\ j\}$. Here 5 cylinders are full and the 4 cylinders $\{1\,0\}, \{1\,1\}, \{2\,0\}, \{2\,2\}$ are mapped onto the upper triangle D which itself is a union of 6 cylinders, namely $\{1\,0\}, \{1\,1\}, \{1\,2\}, \{2\,1\}, \{2\,2\}, \{2\,3\}$. The lower triangle is the union of $\{0\,1\}, \{0\,2\}, \{0\,3\}$.

Now it turns out that the ergodic behaviour of $S \circ T$ is similar to that of the Parry–Daniels map (Nogueira (1995), Schweiger (1981)).

THEOREM 3. *Let* $\Gamma = \{(x_1, x_2) \in B^2 : (S \circ T)^j x \in E \text{ for all } j \ge 0\}$. *Then* $\lambda(\Gamma) > 0$.

REMARKS. (1) Note that $E = [0\,0] \cup [1\,0]$. Since $TE = \{0\,0\} \cup \{0\,1\}$ a similar statement is true for $T \circ S$.

(2) It remains to explain why this set is "exceptional". Let an algorithm produce digits from a digit set I. If J is a proper subset of I then we consider the set $E(J)$ of all points whose expansions contain digits from J only. For many algorithms, the set $E(J)$ is a set of measure zero. For example, for $Tx = 3x \bmod 1$ we find $I = \{0, 1, 2\}$. For $J = \{0, 2\}$ the set $E(J)$ is essentially the well known Cantor set. For Selmer's algorithm again we have $I = \{0, 1, 2\}$. Since $\Delta(0)$ is a transient set for $J = \{1, 2\}$ the set $E(J)$ is the triangle D. But since there is no escape from D possible this behaviour is not so exciting. Here the situation is different. The proof will show that Γ consists of infinitely many segments of variable length which are glued together in the point $(0, 0)$. Since the cylinders $[0\,0]$ and $[1\,0]$ are full cylinders one may think that Γ is more like the Cantor set but the proof shows that it has positive Lebesgue measure.

Proof. We again use the maps

$$\psi(x_1, x_2) = \left(\frac{x_2}{x_1}, \frac{1 - x_1 - x_2}{x_1} \right), \quad \psi^{-1}(u, v) = \left(\frac{1}{1 + u + v}, \frac{u}{1 + u + v} \right).$$

Note that $\psi B^2 = \{(u, v) : 0 \leq u \leq 1,\, u + v \geq 0\}$ and $\psi E = \{(u, v) : 0 \leq u \leq 1,\, v \geq 0\}$. A calculation shows that the restriction of $\psi \circ (S \circ T) \circ \psi^{-1}$ to the set ψE is given as

$$\Theta(u, v) = \begin{cases} \left(\dfrac{u}{1 - u}, \dfrac{-u + v}{1 - u} \right), & 0 \leq u \leq 1/2, \\[2mm] \left(\dfrac{1 - u}{u}, \dfrac{-1 + u + v}{u} \right), & 1/2 \leq u \leq 1. \end{cases}$$

We form the jump transformation with respect to the set $\{(u, v) : 1/2 \leq u \leq 1\}$ and obtain the map

$$P(u, v) = \left(\frac{1}{u} - a, \frac{-1 + u + v}{u} \right), \quad \frac{1}{a + 1} < u \leq \frac{1}{a}.$$

Lebesgue measure λ on B^2 induces on ψE the finite measure

$$\mu(X) = \iint_X \frac{du\, dv}{(1 + u + v)^3}.$$

The Jacobian of a cylinder $B^\star(a_1, \ldots, a_n) = \{(u, v) : u \in B(a_1, \ldots, a_n)\}$ (here $B(a_1, \ldots, a_n)$ refers to the associated continued fraction expansion) with respect to the map P is given as

$$\kappa(a_1, \ldots, a_n; u, v) = \frac{1}{(A_n + B_n u + v)^3}.$$

The numbers A_n, B_n satisfy the recursion relations

$$A_{n+1} = A_n a_{n+1} + B_n + a_{n+1} - 1, \quad B_{n+1} = A_n + 1.$$

Hence

$$A_{n+1} = A_n a_{n+1} + A_{n-1} + a_{n+1}.$$

The growth of A_n can be compared with the growth of the corresponding continued fraction

$$q_{n+1} = q_n a_{n+1} + q_{n-1}.$$

Therefore, clearly $q_n \leq A_n$. The theorem of Borel–Bernstein implies that for almost all $u \in [0,1]$ the inequality

$$a_{n+1} < \frac{q_{n-1}}{(n+1)^2}$$

is true for $n \geq N(u)$.

We claim that for some constant $K = K(u)$,

$$A_n \leq K \sum_{j=1}^{n} \frac{1}{j^2} q_n.$$

For $n < N(u)$ we choose K large enough to start the induction. But then

$$A_{n+1} = a_{n+1} A_n + A_{n-1} + a_{n+1}$$
$$\leq K \left(\sum_{j=1}^{n} \frac{1}{j^2} \right) q_n a_{n+1} + K \left(\sum_{j=1}^{n-1} \frac{1}{j^2} \right) q_{n-1} + \frac{q_{n-1}}{(n+1)^2}$$
$$\leq K \left(\sum_{j=1}^{n+1} \frac{1}{j^2} \right) q_{n+1}.$$

Therefore

$$A_n \leq K \frac{\pi^2}{6} q_n$$

for all $n \geq 1$ for some $K = K(u)$. This means

$$\frac{q_n}{A_n} \geq \gamma > 0, \quad \gamma = \gamma(u) \quad \text{a.e.}$$

Then

$$\sum_{(a_1,\ldots,a_n)} \frac{1}{A_n^2} = \sum_{(a_1,\ldots,a_n)} \frac{q_n^2}{A_n^2} \frac{1}{q_n^2} \geq \sum_{(a_1,\ldots,a_n)} \gamma^2(u) \frac{1}{q_n(q_n + q_{n-1})}.$$

For the sequence of functions f_n defined by

$$f_n(u) = \frac{q_n^2}{A_n^2}, \quad u \in B(a_1, \ldots, a_n),$$

we find $\inf f_n \geq \gamma^2(u) > 0$ a.e. Hence

$$\sum_{(a_1,\ldots,a_n)} \frac{1}{A_n^2} \geq \int\limits_0^1 f_n(u)\,du \geq \int\limits_0^1 \inf f_n(u)\,du > 0.$$

Since $\mu(B^\star(a_1,\ldots,a_n)) \gg 1/A_n^2$ we see that

$$\mu(\Gamma) = \mu\Big(\bigcap_{n=1}^\infty \bigcup_{(a_1,\ldots,a_n)} B^\star(a_1,\ldots,a_n)\Big) > 0$$

and hence $\lambda(\Gamma) > 0$.

CONJECTURE. *The set Γ is an absorbing set.*

References

A. Broise-Alamichel et Y. Guivarc'h (2001), *Exposants caractéristiques de l'algorithme de Jacobi–Perron et de la transformation associée*, Ann. Inst. Fourier (Grenoble) 51, 565–686.

J. C. Lagarias (1993), *The quality of the Diophantine approximations found by the Jacobi–Perron algorithm and related algorithms*, Monatsh. Math. 115, 299–328.

R. Meester and T. Nowicki (1989), *Infinite clusters and critical values in two-dimensional circle percolation*, Israel J. Math. 68, 63–81.

A. Nogueira (1995), *The three-dimensional Poincaré continued fraction algorithm*, ibid. 90, 373–401.

F. Schweiger (1981), *On the Parry–Daniels transformation*, Analysis 1, 171–175.

F. Schweiger (1995), *Ergodic Theory of Fibred Systems and Metric Number Theory*, Clarendon Press, Oxford.

F. Schweiger (2000), *Multidimensional Continued Fractions*, Oxford Univ. Press.

F. Schweiger (2001), *Invariant measure and exponent of convergence for Baldwin's algorithm GCFP*, Sitzungsber, Österr. Akad. Wiss. Math.-Natur. Kl. Abt. II 210, 11–23.

F. Schweiger (2002), *Diophantine properties of multidimensional continued fractions*, in: A. Dubickas *et al.* (eds.), Analytic and Probabilistic Methods in Number Theory (in Honour of J. Kubilius), 242–255.

Department of Mathematics
University of Salzburg
Hellbrunnerstrasse 34
A-5020 Salzburg, Austria
E-mail: fritz.schweiger@sbg.ac.at

(4261)