Simplicity of spectra for certain multidimensional continued fraction algorithms

Charles Fougeron · Alexandra Skripchenko

Received: 7 June 2019 / Accepted: 15 January 2021 / Published online: 7 February 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, AT part of Springer Nature 2021

Abstract
We introduce a new strategy to prove simplicity of the spectrum of Lyapunov exponents that can be applied to a wide class of Markovian multidimensional continued fraction algorithms. As an application we use it for Selmer algorithm in dimension 2 and for the Triangle sequence algorithm and show that these algorithms are not optimal.

Keywords Multidimensional fraction algorithms · Diophantine approximation · Lyapunov exponents · Ergodic measure

Mathematics Subject Classification 11A55 · 11J70 · 37D25 · 28D05

There is a large diversity of multidimensional continued fraction algorithms to approximate a vector of real numbers with a rational vector whose denominators are uniformly bounded.

Whereas in dimension one, Gauss algorithm has a best approximation property which makes it the more natural algorithm to consider, none of the known multidimensional algorithm have such property. Nevertheless, according to the work of Lagarias [21], for Markovian multidimensional continued fraction algorithms there exists a uniform approximation exponent which measure their efficiency and can be estimated with Lyapunov exponents.

Hence to understand better approximation properties of these algorithms, it would be useful to have estimates on Lyapunov exponents. Most results on Lyapunov exponents for multidimensional continued fraction algorithms are showing that the second
Lyapunov exponent is negative (see e.g. [8,25,26]) implying strong convergence of the algorithm. In the present work, we prove simplicity of the Lyapunov spectrum which implies a negative result on the algorithms: they cannot be optimal.

We will consider two specific examples in dimension two: the Triangle Sequence, and Selmer algorithm in dimension 2. We expect our methods to work for any given linear simplex-splitting multidimensional continued fraction algorithms. Its application to those two examples will hopefully convince the reader.

1 Introduction

Let $\Delta := \{ x \in \mathbb{R}^{d+1}_+ \mid \|x\| = 1 \}$ be a $d$-dimensional simplex, where $\|x\| = \sum_{i=1}^{d+1} x_i$ and $\mathbb{R}_+$ is the set of positive real numbers.

A linear simplex-splitting multidimensional continued fraction algorithm (MCF), as in Lagarias [21], is given by a splitting of the simplex into a finite or countable number of subsimplices $\Delta_\alpha$ and a map

$$A : \Delta \rightarrow GL(d + 1, \mathbb{Z}).$$

And for each point $x = (x_0, x_1, \ldots, x_d)$ in $\Delta_\alpha$, we define the map

$$f(x) = \frac{A^{-1}x}{\|A^{-1}x\|}.$$ 

The $n$-th partial quotient matrix of $\theta \in [0, 1]^d$ is denoted by $A^{(n)}(\theta) = A(f^{n-1}(\theta))$ and we define the cocycle $C^n(\theta) = A^{(n)} \cdots A^{1}(\theta)$. For every $n > 0$, we denote the rows of the convergent matrix $C^n$ by $C^n_i = (c^n_{i,1}, \ldots, c^n_{i,d+1})$. They give Diophantine approximations of $\theta$:

$$\tilde{\omega}_i^n := \begin{pmatrix} c^n_{i,1} \\ \vdots \\ c^n_{i,d+1} \\ c^n_{i,d+1} \end{pmatrix}.$$ 

A single approximation of $\theta$,

$$\tilde{\omega} = \begin{pmatrix} p_1 \\ \vdots \\ p_d \\ q \end{pmatrix},$$

is given by the data

$$w = (p_1, \ldots, p_d, q) \in \mathbb{Z}^{d+1},$$

where $q = q(w)$ is its denominator.
For a matrix
\[ W = \begin{pmatrix} w_1 \\ \vdots \\ w_{d+1} \end{pmatrix} \in GL(d + 1, \mathbb{Z}), \]

the denominator is defined
\[ q(W) := \max_{1 \leq i \leq d+1} q(w_i). \]

The uniform approximation exponent of \( W \) is
\[ \eta^*(W, \theta) := \min_{1 \leq i \leq d+1} \left( -\frac{\log \|\theta - \tilde{w}_i\|}{\log q(W)} \right). \]

The best such approximation matrix having a denominator below \( Q \) gives
\[ \eta^*(\theta, Q) := \inf_{q(W) \leq Q} \left( \eta^*(W, \theta) \frac{\log q(W)}{\log Q} \right) \]

And the uniform approximation exponent for \( \theta \) is
\[ \eta^*(\theta) := \liminf_{Q \to \infty} \eta^*(\theta, Q). \]

**The triangle sequence**

Defined by Garrity in 2001 as an iteration of a map on a triangle which yields a sequence of nested triangles [17], the homogeneous Triangle sequence is an algorithm that is almost surely defined by
\[ F : (x_1, x_2, x_3) \in \mathbb{R}_+^3 \mapsto x' = (x'_1, x'_2, x'_3), \]

where if \( \{i, j, k\} = \{1, 2, 3\} \) and \( x_i \geq x_j \geq x_k \),
\[ x'_i = x_i - x_j - bx_k, \quad x'_j = x_j, \quad x'_k = x_k, \]

with \( b = \left\lfloor \frac{x_i - x_j}{x_k} \right\rfloor \).

The non-homogeneous Triangle sequence (a.k.a. Triangle sequence) is a renormalized version of the map \( F \):
\[ f(x) = \frac{F(x)}{\|F(x)\|}. \]

Topological ergodicity of the iterations of the algorithm was proved in [4]. Ergodicity of the algorithm, as well as weak convergence almost surely, were established.
Several new dynamical results have been shown recently on this algorithm in \([2,9,18]\).

**Cassaigne and 2-dimensional Selmer algorithm**

Introduced by Selmer \([28]\), the *homogeneous Selmer algorithm* is almost surely defined by

\[
F : \mathbb{R}_+^3 \rightarrow \mathbb{R}_+^3,
\]

where for \(\{i, j, k\} = \{1, 2, 3\}\) and \(x_i \geq x_j \geq x_k\)

\[
x'_i = x_i - x_k, \quad x'_j = x_j, \quad x'_k = x_k.
\]

Let \(D\) be subcone of \(\mathbb{R}_+^3\) defined by \(x_i < x_j + x_k\) for all \(\{i, j, k\} = \{1, 2, 3\}\). It is an invariant attractive subset of this algorithm.

Let us define the quotient \(S := \mathbb{R}_+^3/\sim\) where we identify \((x_1, x_2, x_3) \sim (x_2, x_3, x_1)\) and \(\pi : \mathbb{R}_+^3 \rightarrow S\) the associated projection. The projection of \(D\) is denoted by \(T := \pi(D)\). As the map \(F\) is equivariant for the identification we can define an induced map \(\tilde{F}\) on \(S\) which can be restricted to \(T\). Moreover the matrix

\[
M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}
\]

sends bijectively \(\mathbb{R}_+^3\) to \(T\) and we denote by \(f\) its inverse map.

Let us denote the conjugated map \(F_C = f \circ \tilde{F}_S \circ f^{-1}\). Using a fundamental domain of \(T\) defined by

\[
\{(x_1, x_2, x_3) \in \mathbb{R}_+^3 \mid \max(x_2, x_3) < x_1 < x_2 + x_3\}
\]

we compute

\[
F_C : (x_1, x_2, x_3) \in \mathbb{R}_+^3 \mapsto \begin{cases} (x_3, x_1, x_2 - x_3) & \text{if } x_2 > x_3 \\ (x_2, x_3 - x_2, x_1) & \text{if } x_3 > x_2 \end{cases}
\]

This map is also called Cassaigne algorithm and has been studied for its interesting combinatorial properties in \([13]\). A geometric explanation of this conjugacy maps together with general dynamical properties of the algorithm can be found in \([15]\).

We have the following commutative diagram.
The ergodicity of the *sorted* Selmer algorithm i.e. its quotient by all 3-permutations, restricted to its attractive invariant subsimplex $F_{|D}$ as well as its weak convergence almost surely is established in [27]. A proof of ergodicity for the *unsorted* algorithm and thus the two others is done in Section 5.2 of [15]. In the following we only need ergodicity of $F_C$, in other words the following result.

**Lemma 1** Cassaigne algorithm is ergodic.

We provide a self-contained proof of this fact in Appendix A.

In the current paper we introduce Lyapunov exponents for these algorithms and study its properties. Our main result is the following

**Theorem 2** The Lyapunov spectra of the cocycles associated to the Triangle sequence and Selmer algorithm in dimension 2 are simple.

Notice that according to the following proposition we can talk indistinctly of the Lyapunov spectrum of the sorted, semi-sorted or unsorted version of an algorithm.

**Proposition 3** For a $d$-dimensional linear simplex-splitting MCF algorithm $f$ and all subgroup $S$ of the permutation group $S(d)$ if $f$ is $S$-equivariant then the induced quotient map $\tilde{f}$ is also a linear simplex-splitting MCF and has the same Lyapunov spectrum.

**Proof** Just notice that for all $\theta \in [0, 1]^d$ and $n > 0$ there exists two permutation matrices $P_1, P_2$ such that $\tilde{C}^n$ the associated cocycle to $\tilde{f}$ satisfies

$$\tilde{C}^n(\theta) = P_1 \cdot C^n(\theta) \cdot P_2.$$  

As $P_1, P_2$ preserve the norm of vectors the Lyapunov spectrum is unchanged. □

Simplicity of spectrum of a dynamical cocycle (in particular, in case of multidimensional fraction algorithms) was established in different contexts. In particular, in case of products of random matrices several results were obtained in [16,19,20]. Simplicity of spectrum for Jacobi–Perron algorithm (in any dimension) was proved in [11].

In a breakthrough result, Avila and Viana introduced the first criterion for simplicity of spectrum for dynamical cocycle (see [7]); and used their criterion in the proof of Kontsevich–Zorich conjecture about the spectrum of Teichmüller flow on the moduli space of Abelian differentials. Later their technique was used by A. Herrera Torres who showed simplicity of spectrum for Selmer MCF (see [29]).
Matheus, M. Möller and J.-C. Yoccoz in [23] developed ideas of Avila and Viana and introduced a Galois-version of the criterion (see Sect. 5 for details). This version was used in [6] for a fully subtractive algorithm and the Rauzy gasket.

Our proof uses a slightly modified version of the same criterion: we check that the ergodic measure has bounded distortion in the sense of [7] and then show that there are two matrices of the cocycle such that both are Galois-pinching and one is twisting with respect to the other.

**Remark** Results proved in [11,29] are more general than our statement since in both cases the simplicity of spectrum of the algorithms of any dimension is established, while our algorithms are only defined in dimension two. However, if the dimension is fixed, the strategy we suggest can be applied for a very wide class of algorithms (basically, for any ergodic Markovian algorithm with integrable cocycle) and allows to get an elementary and straight-forward proof.

### 2 Special acceleration of Ergodic MCF and its properties

#### 2.1 Symbolic dynamics

One can associate a coding of any vector according to a linear simplex-splitting MCF: we consider word formed by the letters corresponding to the sub-simplex $\Delta_\alpha$ in which the vector is at each step of the induction. The cocyle matrix $C^n(\theta)$ only depends on $\gamma$, the coding of the $n$ first steps of the algorithm on $\theta$, and will also be denoted by $A_\gamma$.

**Definition 1** A path $\gamma$ is called positive if the corresponding matrix of the cocycle $A_\gamma$ has only positive entries.

One can check that the normalized image $A_\gamma(\Delta)$ is the subsimplex of points whose coding start by $\gamma$. Moreover, when a matrix $M$ is strictly positive, the normalization of $M(\Delta)$ is compactly included into $\Delta$. This enables us to introduce a special acceleration.

**Remark 1** Positiveness and ergodicity imply that one can consider the following acceleration of the given algorithm that was first defined in [5] for the Veech flow and interval exchange transformations (see Section 4.1.3): a special acceleration is a first return map to some subsimplex $\Delta_\gamma$ compactly contained in the parameter space $\Delta$ (here $\Delta_\gamma = \Delta_{\gamma_*}$, where $\gamma_*$ is some strictly positive path). Naturally, not all of the orbits of the original algorithm will return to $\Delta_\gamma$; nevertheless, due to ergodicity, it will happen to the orbits of almost every points.

### 3 Projective expanding maps

For any $C^1$ map $T : X \to Y$, we denote by $d_xT$ its differential at a point $x \in X$. A Jacobian (matrix) is a matrix that represents this linear map.

We follow a definition of Avila–Viana [7]:
Definition 2 A projective expanding map is a map $T: \bigcup \Delta^{(l)} = \Delta_* \to \Delta_*$, where $\Delta_*$ is a simplex compactly contained in the standard simplex, the $\Delta^{(l)}$'s form a finite or countable family of pairwise disjoint simplexes contained in $\Delta_*$ and covering almost all of $\Delta_*$, and $T^{(l)} = T: \Delta^{(l)} \to \Delta_*$ is a bijection such that $(T^{(l)})^{-1}$ is the restriction of a projective contraction i.e. it is represented by a positive matrix.

Remark 2 It is clear from the definition that a special acceleration of any ergodic simplex-splitting MCF is a projective expanding map.

To give the reader an idea why this definition is fundamental, we will prove the following implication.

Proposition 4 If $T$ is a special acceleration of a simplex-splitting MCF, there exists $k > 0$ such that the Jacobian of $T$ satisfies for all $x \in \Delta_*$, $\|d_x T\|_\infty \geq k$.

To prove this proposition, we need to introduce Hilbert pseudo-metric on $\mathbb{R}^A_+$:

$$
\text{dist}(x, y) = \max_{\alpha, \beta \in A} \log \frac{x_\alpha y_\beta}{y_\alpha x_\beta}.
$$

This pseudo-metric induces the metric on the space of rays $tx : t \in \mathbb{R}_+$. This metric is a complete Finsler metric. We will also need the operator norm $|A| = \sup_{\|v\|=1} A v$ and the norm on continuous operators on compact sets: $\|A\|_{C^0(\Delta)} = \sup_{\lambda \in \Delta} |A(\lambda)|$.

Proof Notice that the simplex $\Delta_*$ is partitioned by countably many subsimplices indexed by paths $\gamma \gamma_*$, where $\gamma_*$ is not a factor of $\gamma$. The inverse of the map $T^{(l)}$ is given by the projective map

$$(T^{(l)})^{-1}: \lambda \mapsto \frac{A_{\gamma \gamma_*} \lambda}{\|A_{\gamma \gamma_*} \lambda\|}.$$

The Hilbert metric is defined projectively so we can assume that the map above is the linear associated to the matrix $A_{\gamma_*} A_{\gamma}$. It is a well known fact about Hilbert metrics (see e.g. [30]) that a positive matrix is $\theta$-Lipschitz with respect to it, for some $\theta < 1$, and a non-negative matrix is 1-Lipschitz. Hence the inverse of $T$ is $\theta$-Lipschitz for some $\theta < 1$ depending only on $\gamma_*$. This implies a uniform lower bound on the Jacobian of $T$ for normalized vectors. \qed

Ergodic properties of expanding maps (whose Jacobian is uniformly bounded from below) are well-known (Theorem 1.3 in [22] and Section 4 in [1]):

Corollary 5 A special acceleration of a simplex-splitting MCF admits a unique absolutely continuous invariant measure. Moreover their density is bounded from above and from zero.

This measure corresponds to the given ergodic measure in the considered cases and has bounded distortion in the sense of Avila–Viana:
Definition 3 An invariant measure $\mu$ on a topological Markov shift $\Sigma_M$ has **bounded distortion property** if there exists a positive constant $C(\mu)$ such that for any (non-empty) cylinder set $[a_{i_1} \cdots a_{i_n}]$ and any $1 \leq j \leq n$ we have

$$\frac{1}{C(\mu)} \leq \frac{\mu([a_{i_1} \cdots a_{i_n}])}{\mu([a_{i_1} \cdots a_{i_j}])\mu([a_{i_{j+1}} \cdots a_{i_n}])} \leq C(\mu).$$

Corollary (Avila–Viana) The absolutely continuous ergodic measure of a special acceleration of a simplex-splitting MCF has bounded distortion.

Lemma 7.2 in [7] states that if the map is projectively expanding it has **approximate product structure** with respect to Lebesgue measure. Moreover, in Appendix A of [7] is proved the equivalence of the bounded distortion property of the measure defined above and the property of product structure. We will not give the definition of this latter property since it only serves to reformulate Avila–Viana’s lemma.

This implies bounded distortion of Lebesgue measure and thus of all equivalent measures.

This bounded distortion property is a key hypothesis to use Avila–Viana theorem and prove Theorem 10.

4 Lagarias conditions

Following [21], we give a list of properties on a MCF map $T : \Delta_* \to \Delta_*$ (which will be a special acceleration in our cases). These properties imply, by the work of Lagarias, existence of Lyapunov exponents and a formula, in terms of these exponents, for the uniform approximation exponent of the algorithm.

Notice that in our case, the invariant measure has full support hence we will not discuss Lagarias conditions in case of a smaller support.

(H1) **Ergodicity** $T$ has an ergodic absolutely continuous measure $d\mu$.

(H2) **Covering Property** $T$ is piecewise continuous with non-vanishing Jacobian almost everywhere.

(H3) **Semi-weak convergence** There is a set $Z$ of full Lebesgue measure and a constant $c_0 > 1$ such that for all $\theta \in Z$ and all $n \geq n_0(\theta)$,

$$\max_{1 \leq j \leq d+1} \| \theta - w_j^{(n)} \| \leq c_0^{-n(1+o(1))}.$$

(H4) **Boundedness** One has

$$\int_{\Delta_*} \log \| A(x) \|^+ \, d\mu(x) < \infty$$

where $\| A(x) \|^+ = \max(1, \| A(x) \|)$.
(H5) **Partial Quotient Mixing** All allowable partial quotient matrices of the algorithm are nonnegative matrices. For any $\theta \in \Delta_s$, we set

$$
\nu(\theta) = \min\{k : C^k(\theta) \text{ is strictly positive}\}
$$

then

$$
\int_{\Delta_s} \nu(\theta) d\theta < \infty.
$$

We now discuss these properties for a special acceleration of Triangle sequence and 2-dimensional Selmer algorithms and when they hold for special accelerations of other algorithms.

(H1) **Ergodicity**

As mentioned previously, ergodicity of Triangle sequence is prove in [4,24] and ergodicity of Selmer in [15,27]. It follows from Proposition 5 that the invariant measures for the accelerated Triangle sequence and Selmer algorithm are absolutely continuous with respect to Lebesgue measure and their density are bounded from above and from zero.

(H2) **Covering property**

The map $f$ is piecewise continuous with non-vanishing Jacobian almost everywhere. This property holds by Proposition 4.

(H3) **Semi-weak convergence**

We follow a result of Section 6 of [21].

**Lemma 6** A special acceleration of any ergodic simplex-splitting MCF has semi-weak convergence property.

**Proof** The cylinders of the Markov partition are simplices whose vertices are given by the rows of a convergent matrix. The diameters of these cylinders decrease geometrically for a set of full measure because of the lower bound on the Jacobian proved in Proposition 4 (or, equivalently, the statement about the size of the cylinders follows from the bounded distortion property of the measure).

(H4) **Boundedness**

This property is the standard log-integrability of the cocycle that is used in Oseledets theorem. Our proof is similar to the results in [21] for Selmer and Jacobi-Perron algorithms, and concern the slow version of the algorithm:
Lemma 7  The cocycle of Triangle sequence is log-integrable:

$$\int_{\Delta} \log \|A\|^+ d\mu < \infty.$$  

Proof  The statement follows from a direct calculation: for each step of the algorithm the matrix norm grows with linearly with a parameter $b$:

$$\|A\| = b + 4$$

and the measure (area) of the corresponding subsimplex decreases quadratically with the same parameter $b$. Indeed, using renormalization condition $x_i + x_j + x_k = 1$ we exclude one coordinate (say, $x_2$) and so it is easy to see that subsimplex $\Delta^{(l)}$ is a triangle with the following vertices: $(\frac{1}{2}, 0); \left(\frac{b+1}{b+3}, \frac{1}{b+3}\right); \left(\frac{b+2}{b+4}, \frac{1}{b+4}\right)$. Therefore

$$\mu(\Delta^{(l)}) = \frac{1}{4(b + 3)(b + 4)}.$$  

The statement about log-integrability follows from the convergence of the series $\sum_n \frac{\log n}{n^2}$.  

Lemma 8  The cocycle of Selmer algorithm is log-integrable.  

Proof  It is clear since the cocycle take only two finite values on two parts of the simplex.  

As the cocycle is induced on the acceleration of the algorithm and the log of its norm is sub-additive, the integral is bounded from above by the integral on the cocycle before acceleration. Moreover, as noticed in [7], page 46, the Lyapunov exponents get multiplied by the inverse of the measure of the subsimplex. Thus the simplicity property is preserved by acceleration.  

(H5) Partial Quotient Mixing  

For a special acceleration $\nu(\theta) = 1$ for all $\theta \in \Delta_s$. Thus $\int_{\Delta_s} \nu(\theta) d\lambda = \lambda(\Delta_s) < \infty$.  

Remark  Properties H1–H3 and H5 are satisfied for every special acceleration of any ergodic algorithm. Only property H4 needs to be checked on the algorithm before acceleration, but has already been proved for a large number of MCF.  

Lyapunov exponents and Convergence  

It was shown by Lagarias in [21] (see Theorem 4.1) that if conditions H1–H5 are satisfied, then the uniform approximation exponent of the algorithm for almost all the
parameters can be estimated in the following way:

$$\eta^*(\theta) = 1 - \frac{\lambda_2}{\lambda_1}.$$  

Notice that these exponents are zero if and only if $\lambda_1 = \lambda_2$. This implies the following application to our main result.

**Corollary 9** The Triangle sequence and Selmer algorithms have strictly positive best uniform approximation exponents and cannot be an optimal algorithm.

**Proof** The first statement is straightforward by the previous remark. For the second statement, one has to notice that the sum of Lyapunov exponents is equal to zero, since the matrices of the cocycle have determinant one. As proved by Lagarias, the uniform approximation exponent is always bounded by $1 + 1/d$. Thus an algorithm is optimal i.e. is an equality case for this bound, if and only if $\lambda_2 = \cdots = \lambda_d$. Which contradicts simplicity. □

## 5 Simplicity of spectrum

In this section we prove our main result (Theorem 2). First, we show simplicity for the spectrum of accelerated version of the Triangle sequence; as was mentioned above, Avila and Viana showed [7] that simplicity of spectrum of the induced cocycle implies the same property for the original one.

### 5.1 Avila–Viana’s criterion for simplicity

Let $\Lambda$ be a finite or countable alphabet and $\Sigma$ be the set of words: $\Sigma = \Lambda^\mathbb{N}$. The biinfinite version of $\Sigma$ is denoted by $\hat{\Sigma} = \Lambda^\mathbb{Z}$. Let $T$ and $\hat{T}$ be two shifts acting on $\Sigma$ and $\hat{\Sigma}$ respectively. With each word $l$ we associate a cylinder

$$\Sigma(l) = \{ x \in \Sigma : x \text{ starts in } l \}.$$  

Let $A : \Sigma \to G$, where $G$ is some matrix group (in our case $G = SL(d, \mathbb{Z})$). Then one can define a dynamical cocycle $(T, A)$ acting on $\Sigma \times \mathbb{R}^d$:

$$(T, A)(x, v) = (T(x), A(x)v).$$  

Then for each $l = (l_0, \cdots, l_{n-1}) \in \bigcup_{n \geq 0} \Lambda^n$ we define $A^L_l = A_{l_{n-1}} \cdots A_{l_0}$, so that we have $(T, A)^n(x, v) = (T^n(x), A^L_l(x)v)$ for each $x \in \Sigma(l)$.

**Definition 4** The cocycle is pinching if there exists a word $l^* \in \bigcup_{n \geq 0} \Lambda^n$ such that the spectrum of $A^L_l$ is simple. The matrix $A^L_l$ is said to be pinching.

Let $Gr(k)$ be a Grassmanian of $k$-planes of $\mathbb{R}^d$. 

 Springer
Definition 5 The cocycle is twisting if for any \( m \geq 1 \), any \( 1 \leq k_1, \ldots, k_m < n \), and any subspaces \( F_i \in Gr(k_i) \) and \( F'_i \in Gr(n - k_i) \) there exists a word \( l \) such that \( A^{l(F_i)} \cap F'_i = \{0\} \).

Then the following theorem was proved by Avila–Viana in [7]:

Theorem 10 Let \( \mu \) be a \( T \)-invariant probability measure on \( \Sigma \) with the bounded distortion property. Let \( A \) be a locally constant \( G \)-valued integrable cocycle. Assume that \( A \) is pinching and twisting. Then, the Lyapunov spectrum of \((T, A)\) with respect to \( \mu \) is simple.

5.2 Matheus–Möller–Yoccoz’s criterion

The criterion of Theorem 10 was improved by Matheus, Möller and Yoccoz in [23]. We explain briefly their results and adapt them to the case of \( G = SL(d, \mathbb{Z}) \).

Definition 6 Let \( A \in G \) be a pinching matrix. The matrix \( B \in G \) is \( k \)-twisting with respect to \( A \) if for every pair of \( A \)-invariant \( F \in Gr(k) \) and \( F' \in Gr(n - k) \), \( B(F) \cap F' = \{0\} \).

Let us quote Proposition 2.16 in [23].

Proposition 11 A cocycle \( A \) is pinching and twisting if and only if there exists a word \( l^* \) such that \( A^{l^*} \) is a pinching matrix and for some word \( l(k) \) \( A^{l(k)} \) is \( k \)-twisting with respect to \( A \).

As a corollary of Proposition 11 and Theorem 10, simplicity of spectrum for a given cocycle holds if there is a pinching matrix \( A \) and a matrix \( B \) that is \( k \)-twisting with respect to \( A \).

The next step is a so-called Galois-version of the previous simplicity criterion. We use a slightly modified definition from [23]:

Definition The element \( A \in SL(d, \mathbb{Z}) \) is called Galois-pinching if its characteristic polynomial is irreducible over \( \mathbb{Q} \), all its roots are real and the Galois group \( \mathbb{G} \) is the largest possible \( i.e. S_d \).

Similarly to Proposition 4.2 in [23],

Proposition A Galois-pinching matrix is pinching.

Proof For completeness we reproduce the proof of [23]. By the first two conditions, all eigenvalues are simple and real. The only possibility preventing the matrix to be pinching would be to have both \( \lambda \) and \( -\lambda \) as eigenvalues. But an element of the Galois group fixing \( \lambda \) will also fix \( -\lambda \) thus it is not \( S_d \).

Therefore, once we find a Galois-pinching matrix, we only need to explain how to find a \( k \)-twisting matrix. We prove a very close analogue of the Theorem 4.6 in [23]:

Theorem 12 Let \( A \) and \( B \) be two elements of \( SL(d, \mathbb{Z}) \). Assume that \( A \) and \( B \) do not share any common proper invariant subspace and that \( A \) is Galois-pinching. Then, there exist \( m \geq 1 \) and, for any \( l^* \), integers \( l_1, \ldots, l_m - 1 > l^* \) such that the product \( BA^{l_1} \cdots BA^{l_m - 1}B \) is \( k \)-twisting with respect to \( A \).
Following Lemma 4.8 in [23], we prove a transversality condition. Let $\lambda_i$ be the eigenvalues of the Galois-pinching matrix $A \in SL(d, \mathbb{Z})$ and let $v_{\lambda_i}$ be the corresponding eigenvectors. The coordinates of the eigenvectors belong to $\mathbb{Q}(\lambda_i)$ then, for any element $g \in \mathcal{G}$, one can assume that $g(v_{\lambda_i}) = v_{g\lambda_i}$. We denote $\lambda = \{\lambda_1 < \cdots < \lambda_k\}$ and the exterior product $v_\lambda = v_{\lambda_1} \wedge v_{\lambda_2} \cdots \wedge v_{\lambda_k}$.

Let us consider some matrix $C \in SL(d, \mathbb{Z})$.

**Lemma 13** For some $\lambda$ and $\lambda'$ we denote by $C_k^{\lambda\lambda'}$ the coefficients of the matrix $\wedge^k C$ in the basis $v_\lambda$ defined above. Then $C$ is $k$-twisting with respect to $A$ if and only if all coefficients $C_k^{\lambda\lambda'}$ are non-zero.

In other words, this statement means that we can express $k$-twisting properties in terms of entries of the $k$-th exterior power of the matrix.

**Proof** Indeed, the $A$-invariant spaces are generated by eigenvectors, therefore if $E$ and $F$ are two $A$-invariant subspaces of dimensions $k$ and $n - k$ respectively, then $C(E)$ is transversal to $F$ if and only if the exterior product $\wedge^k C(v_{\lambda_E}) \wedge \lambda_F$ is non-zero, which happens precisely when the coefficient $C_k^{\lambda_E\lambda_F}$ is non-zero. $\square$

Now, we introduce a family of oriented graphs associated to the matrix $C$. Namely, each set $\lambda$ defines a vertex of the graph $\Gamma_k$ and two vertices $\lambda$ and $\lambda'$ are connected with an arrow if and only the coefficient $C_k^{\lambda\lambda'}$ written in the basis of eigenvectors is non-zero. Note that $\Gamma_k$ is invariant under the natural action of the Galois group $\mathcal{G}$. Lemma 13 means that $C$ is $k$-twisting with respect to $A$ if and only if $\Gamma_k$ is complete.

**Proof of Theorem 12** Using the previous remarks, one only needs to check that the graph $\Gamma_k$ is complete. Since the matrix $C$ is invertible, each graph $\Gamma_k(C)$ contains at least one arrow. But $A$ is Galois-pinching therefore its Galois group is the group of all permutation. As the graph $\Gamma_k(C)$ is $\mathcal{G}$-invariant, the existence of one arrow implies existence of all arrows. $\square$

**Remark 3** The proof of the $k$-twisting criterion in case of the symplectic group $Sp$ provided in [23] is remarkably more complicated. It comes from the fact that the Galois group has a more sophisticated nature and completeness of the graph does not hold for every $k$.

We are now ready to apply the criterion to our setting. We have already checked that the measure has bounded distortion. Therefore, it is enough to find two explicit paths for which the cocycles are Galois-pinching and one is $k$-twisting with respect to the other. We do so in the next two paragraphs.

**Triangle sequence**

Consider the following paths in the Triangle sequence algorithm, coded by the order of the coordinates, $(i, j, k)$ if $x_i < x_j < x_k$, and the number $b$:

$\gamma_1 = ((1, 2, 3), 0) \rightarrow ((3, 1, 2), 1) \rightarrow ((2, 3, 1), 2)$,
\( \gamma_2 = ((1, 3, 2), 0) \rightarrow ((2, 1, 3), 0) \rightarrow ((3, 2, 1), 3) \)

Along these paths the cocycles are equal to,

\[
A^{\gamma_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 2 \\ 1 & 3 & 3 \end{pmatrix}
\]

\[
A^{\gamma_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 4 \\ 1 & 2 & 5 \\ 1 & 1 & 5 \end{pmatrix}
\]

their characteristic polynomials are,

\[
P^{\gamma_1} = x^3 - 7x^2 + 6x - 1,
\]

\[
P^{\gamma_2} = x^3 - 8x^2 + 7x - 1,
\]

which discriminants are equal to,

\[
\Delta(P^{\gamma_1}) = 697 = 17 \cdot 41,
\]

\[
\Delta(P^{\gamma_2}) = 257.
\]

**Selmer algorithm**

We code the path, in Cassaigne’s presentation of the restricted algorithm, by words in numbers 1, 2, depending if they satisfy the first or second case in the definition of the function \( F \) at each step. Consider the paths

\[
\gamma_1 = 2 \rightarrow 1 \rightarrow 2 \rightarrow 2 \rightarrow 1
\]

\[
\gamma_2 = 1 \rightarrow 2 \rightarrow 2 \rightarrow 2 \rightarrow 1 \rightarrow 2 \rightarrow 1.
\]

Along these paths the cocycles are equal to

\[
A^{\gamma_1} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 2 \end{pmatrix}
\]

\[
A^{\gamma_2} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 3 \\ 1 & 1 & 1 \end{pmatrix},
\]

their characteristic polynomials are

\[
P^{\gamma_1} = x^3 - 4x^2 + 1
\]

\[
P^{\gamma_2} = x^3 - 6x^2 + 1,
\]
Simplicity of spectra for certain MCF… 781

which discriminants are equal to

\[ \Delta(P^{\gamma_1}) = 229, \]
\[ \Delta(P^{\gamma_2}) = 3^3 \cdot 31. \]

**Pinching and twisting**

For both algorithms, we establish a proof of the two following propositions.

**Proposition 14** The matrices \( A^{\gamma_1} \) and \( A^{\gamma_2} \) are Galois-pinching.

**Proof** Observe that their characteristic polynomials,

- their first and the last coefficients are 1 and \(-1\), which implies according to the rational root theorem that they do not have a rational root, and thus are irreducible over \( \mathbb{Q} \);
- have positive discriminants, thus all of their roots are real;
- moreover the discriminants are not a square of a rational number, this implies (see [14]) that their Galois groups are isomorphic to \( S_3 \).

\[ \square \]

**Proposition 15** The matrices \( A^{\gamma_1} \) and \( A^{\gamma_2} \) do not share a common proper invariant subspace.

**Proof** The statement follows, for example, from [10] (see Theorems 2 and 3): if two Galois-pinching matrices share a common proper invariant space, then they necessarily commute. A simple direct check shows that \( A^{\gamma_1} \) and \( A^{\gamma_2} \) do not commute.

\[ \square \]

**Proof of Theorem 2** Propositions 14 and 15 imply that the cocycle is pinching and twisting in the sense of [7]. Now Theorem 2 follows using bounded distortion property for the measure and Theorem 10.

\[ \square \]

**Appendix A. Proof of Lemma 1**

**A.1 Strategy**

In this appendix we give a proof of the ergodicity of Cassaigne algorithm with classical techniques.

First, we recall the definition of *ergodicity* for multidimensional fraction algorithm given by Schweiger [27] (Definition 21 at p 21).

**Definition 7** \( T \) is called ergodic if \( T^{-1}(E) = E \mod 0 \) implies \( \lambda(E) = 0 \) or \( \lambda(B/E) = 0 \) for the Lebesgue measure \( \lambda \).

**Remark 4** One can easily see that this definition is equivalent to the standard definition of ergodicity when the invariant measure is equivalent to the Lebesgue measure.
We follow the strategy suggested by Schweiger in [27] for Selmer algorithm (Theorem 23, p 60):

1. describe the associated Markov partition and symbolic coding for the cylinders;
2. define the absorbing and so called bad sets (in a sense of distortion properties) and check that the bad set is of measure zero;
3. define Schweiger jump transformation as a natural acceleration of the algorithm that takes each point of the simplex except those in the bad set to the absorbing set;
4. prove that Schweiger jump transformation has weak bounded distortion property and therefore is ergodic.
5. conclude that the original algorithm was ergodic as well.

A.2 Markov partition

We define a Markov partition for the Cassaigne algorithm in the following way. The alphabet for the coding contains two letters (let denote them by 1 and 2). After one step of the algorithm we have the following partition of the parameter space $\Delta = (x_1, x_2, x_3 : \sum_i x_i = 1)$,

$$B_1 = \{(x_1, x_2, x_3) \in \Delta \mid x_1 > x_3\}$$

and

$$B_2 = \{(x_1, x_2, x_3) \in \Delta \mid x_1 < x_3\}.$$

A direct calculation reveals that $f = \frac{F_C}{\|F_C\|}$ maps $B_i$ onto the whole parameter space and that vertices of each simplex are mapped into the vertices. For example, for $(x_1, x_3) \in B_1$, $f(x_1, x_3) = (\frac{x_1 - x_3}{1 - x_3}, \frac{1 - x_1 - x_3}{1 - x_3})$.

Iterating the algorithm, we get the Markov partition $\alpha$. Symbolic dynamics is given by a Markov shift defined on the cylinders described by 2 letters.

A.3 Absorbing set versus bad set

Absorbing set is formed by the cylinders that contain words where both letters are included:

$$D = \bigcup_k (1^k 2 \cdots) \cup \bigcup_k (2^k 1 \cdots).$$

Here $1^k$ means a sequence of $k$ times 1.

The bad set is formed by infinite cylinders without change of letters: $H = (1 \cdots 1 \cdots) \cup (2 \cdots 2 \cdots)$. Naturally, $H \cup D$ gives the whole parameter space. And the following holds:

Lemma 16 The measure of the set $H$ is 0, the set $D$ is of full measure.
A.4 A Schweiger jump transformation

For any cylinder $B$ (except those that are in a bad set $H$) we can define the arrival time to the absorbing set $D$ and denote it by $N_B$. For Cassaigne algorithm, we arrive to the absorbing set once we change letter. For example, for any cylinder $B_n = (1 \cdots 12)$ or $B'_n = (2 \cdots 21)$ this time is equal to the number of 1s (respectively, 2s). It allows us to define an accelerated version of the algorithm (an analogue of Zorich acceleration for the Rauzy induction) that contains all the consequent stages of the original algorithm coded by the same letter and the first stage after the change. We denote this accelerated algorithm by $\hat{F}_C : \hat{F}_C = F_C^{N_B}$.

A.5 Distortion properties

We recall some definitions from [1] (see Chapter 4).

**Definition 8** For a given Markov map $T$ acting on a Polish space $X$ with Markov partition $\alpha$ and measure $m$ the distortion at $(x, y) \in Dv_a \times DV_a$ is the ratio

$$G(T, x, y) = \frac{v'_a(x)}{v'_a(y)}$$

where $v'_a$ is the Radon–Nikodym derivative of the non-singular inverse branches $v_a$ of $T^n$ with respect to Lebesgue measure.

**Strong bounded distortion property** means that there exists a constant $C > 1$ such that, for all Markov cells $a$ of positive measure, $G(T, x, y) \leq C$ for almost every $(x, y) \in Dv_a \times DV_a$.

**Schweiger collection** is a special subset of the partition $\tau \subset \alpha$ such that the strong bounded distortion property with some $C > 1$ holds for all cylinders in $\tau$ and for any two cylinders of positive measure $[b] \in \tau$ and $[a] \in \alpha$ $[a, b] \in \tau$ and $\bigcup_{B \in \tau} B = X$.

**Weak bounded distortion property** means that for a given Markov map there exists a Schweiger collection.

**Lemma 17** Schweiger jump transformation of $F_C$ has weak bounded distortion property.

**Proof** It is enough to check that the statement holds for almost all points in the absorbing set. The proof is based on a direct calculation and is organized as follows. First we check that the condition holds for one step of our accelerated algorithm; then we use it to prove that a stronger condition holds and this condition implies the statement for all the iterations of the accelerated algorithm.

Lets consider the cylinder $B_n = (1 \cdots 12)$ (the opposite case of $C_n = (2 \cdots 21)$ is similar). The index $n$ refers to the number of 1 in the cylinder coding. The projective version of $F_C$ is one of the two maps: either if $n = 2k$,

$$\hat{F}_C(x_1, x_3) = \left( \frac{1 - x_1 - x_3}{1 - x_1}, \frac{k + x_3 - (k + 1)x_1}{1 - x_1} \right).$$
either if \( n = 2k + 1 \),

\[
\hat{F}_C(x_1, x_3) = \left( \frac{x_3}{1 - x_1}, \frac{1 + k - (k + 2)x_1}{1 - x_1} \right).
\]

The Jacobian of the map in both cases is \( J = \frac{1}{(1 - x_1)^3} \).

We treat here only the situation \( n = 2k \) (the other one is completely similar). In this case the cylinder identifies the triangle in the parameter space with the following vertices:

\[
\left( \frac{k}{k + 1}, \frac{1}{k + 1} \right); \quad \left( \frac{k}{k + 1}, 0 \right); \quad \left( \frac{k + 1}{k + 2}, \frac{1}{k + 2} \right).
\]

Hence we know that \( x_1 \in [k/k + 1, (k + 1)/(k + 2)] \).

Therefore, the ratio of two Jacobian’s taken at points \( a \) and \( b \) in the same cylinder satisfies \( \frac{(k + 1)^3}{(k + 2)^3} < \frac{J(a)}{J(b)} < \frac{(k + 2)^3}{(k + 1)^3} \). It follows that this ratio is uniformly bounded (e.g. by \( \frac{1}{8} \) and 8).

Now we need to check that the ratio of two Jacobians remains uniformly bounded after any number of steps of the accelerated algorithm. In order to do it, we introduce another bounded distortion property.

**Definition 9** We say that the map \( T : X \to X \) satisfies Renyi condition if two following properties hold:

- for two points \( a, b \) from the same element of the partition we have
  \[
  \left| \frac{J_T(a)}{J_T(b)} - 1 \right| \leq C \| T(a) - T(b) \| \tag{1}
  \]
  with some uniform constant \( C \) (\( J_T \) is the Jacobian map of \( T \));
- \( |J_{T^n}| \geq K\lambda^n \) for some uniform constants \( K > 0 \) and \( \lambda > 1 \).

This condition is a classical version of the bounded distortion property widely used in the study of ergodic properties of different interval maps (see, for example, [12]).

It is easy to check that if \( T \) satisfies Renyi condition, then any iteration \( T^k \) of the map \( T \) also satisfies the above conditions with some uniform constant \( C' \) (that is the same for all \( k \)) (see Theorem 1.2 in Chapter 3 in [22]). This fact implies, in particular, that the ratio of Jacobians of the iterated map at any two points from the same elements in the corresponding partition is uniformly bounded. Therefore, if one checks that our jump transformation (a.k.a. acceleration of Cassaigne algorithm) \( F_C \) satisfies Renyi condition for the partition given by \( B_n \) and \( C_n \) for the first step of the algorithm, it is enough to conclude that the union of \( B_n \) and \( C_n \) gives us a Schweiger collection. And thus the map \( F_C \) satisfies weak distortion property.

Now it remains to check that Renyi condition is satisfied for \( F_C \). We consider the case of \( B_n \) (the other case is very similar).
As we explained above, for any \( a = (a_1, a_3) \) and \( b = (b_1, b_3) \) from the same cell of the partition \( \left| \frac{1-a_1}{1-b_1} \right| \) is bounded by some uniform constants from above and from below (we took 1/8 and 8). So

\[
\left| \frac{J(b)}{J(a)} - 1 \right| = \left| \left( \frac{1-a_1}{1-b_1} \right)^2 + \frac{1-a_1}{1-b_1} + 1 \right| \cdot \left| \frac{1-a_1}{1-b_1} - 1 \right|
\leq (2^2 + 2 + 1) \left| \frac{1-a_1}{1-b_1} - 1 \right|.
\]

Now we note that

\[
\hat{F}_C(a) - \hat{F}_C(b) = (A, B),
\]

where \( A = \frac{b_3}{1-b_1} - \frac{a_3}{1-a_1} \) and \( B = \frac{1-b_3}{1-b_1} - \frac{1-a_3}{1-a_1} \). Since \( A^2 + B^2 \geq \frac{(A+B)^2}{2} \), we conclude that

\[
\| \hat{F}_C(a) - \hat{F}_C(b) \| \geq \frac{1}{\sqrt{2}} \cdot \frac{|b_1 - a_1|}{(1-a_1)(1-b_1)} \geq \frac{1}{\sqrt{2}} \cdot \frac{|b_1 - a_1|}{1-b_1} = \frac{1}{\sqrt{2}} \cdot \left| \frac{1-a_1}{1-b_1} - 1 \right|.
\]

Therefore the condition (1) is satisfied with constant \( C = 7 \sqrt{2} \).

Also \( J_T = \frac{1}{(1-a_1)^2} \geq 8 \) because as we know \( a_1 \geq \left( \frac{k}{k+1} \right) \). It implies the second part of Renyi condition. \( \square \)

To conclude we use the following famous result on ergodicity of Markov maps whose proof can be found in [1] (Theorem 4.6.3).

**Theorem 18** (Aaronson–Denker–Urbanski) A topologically transitive Markov map with weak bounded distortion property is either conservative and ergodic or totally dissipative.

Since the Cassaigne algorithm (and therefore its acceleration) can not be totally dissipative (the density of its invariant measure was calculated in [3]), Schweiger jump transformation of \( F_C \) is ergodic.

**Remark 5** The same conclusion can be obtained from the so called Adler Folklore theorem (see discussion after Theorem 1.2 in [22] where the argument can be done for \( \gamma = 1 \)) using the Renyi property we established above.

### A.6 Ergodicity of the original algorithm

The following theorem proved by Schweiger (see [27], Theorem 11, p 19) now implies ergodicity of the normalized Cassaigne algorithm:

**Theorem 19** If Schweiger jump transformation of the Markovian algorithm is ergodic, then the original algorithm is also ergodic.
Acknowledgements  We heartily thank Carlos Matheus who explained to us the proof of theorem 2.17 in [23] in the case of $SL(n, \mathbb{Z})$ and other important remarks. We also thank Pascal Hubert and Vincent Delecroix for several useful discussions. We also thank our referee for many improvements for the first version of this paper. The work was done while the first author was visiting the Max-Planck-Institut für Mathematik in Bonn, he is grateful to MPI for the excellent working conditions, The second author was partially supported by RFBR-CNRS Grant No. 18-51–15010 and The study has been funded within the framework of the HSE University Basic Research Program.

References

1. Aaronson, J.: An introduction to infinite ergodic theory. Mathematical Surveys and Monographs, vol. 50. American Mathematical Society, Providence (1997)
2. Amburg, I., Garrity, T.: Functional analysis behind a family of multidimensional continued fractions: Triangle partition maps (2017)
3. Arnoux, Pierre, Labbé, Sébastien: On some symmetric multidimensional continued fraction algorithms. Ergod. Theory Dyn. Syst. 38(5), 1601–1626 (2018)
4. Assaf, S., Chen, L.-C., Cheslack-Postava, T., Cooper, B., Diesl, A., Garrity, T., Lepinski, M., Schuyler, A.: A dual approach to triangle sequences: a multidimensional continued fraction algorithm. Integers 5(1), 39 (2005)
5. Avila, Artur, Gouëzel, Sébastien, Yoccoz, Jean-Christophe: Exponential mixing for the Teichmüller flow. Publ. Math. Inst. Inst. Hautes Études Sci. 104, 143–211 (2006)
6. Avila, Artur, Hubert, Pascal, Skripchenko, Alexandra: Diffusion for chaotic plane sections of 3-periodic surfaces. Invent. Math. 206(1), 109–146 (2016)
7. Avila, Artur, Viana, Marcelo: Simplicity of Lyapunov spectra: proof of the Zorich–Kontsevich conjecture. Acta Math. 198(1), 1–56 (2007)
8. Berthé, V., Steiner, W., Thuswaldner, J.: A slow triangle map with a segment of indifferent fixed points and a complete tree of rational pairs (2019)
9. Bourgeois, G.: Common invariant subspace and commuting matrices. Linear Algebra Appl. 438(7), 3030–3038 (2013)
10. Broise-Alamichel, A., Guivarc’h, Y.: Exposants caractéristiques de l’algorithme de Jacobi-Perron et de la transformation associée. Ann. Inst. Fourier (Grenoble) 51(3), 565–686 (2001)
11. Bruin, Henk, Clack, Gregory: Inducing and unique ergodicity of double rotations. Discrete Contin. Dyn. Syst. 32(12), 4133–4147 (2012)
12. Cassaigne, J., Labbé, S., Leroy, J.: A set of sequences of complexity $2n + 1$. Combinatorics on Words. Lecture Notes in Computer Science, vol. 10432, pp. 144–156. Springer, Cham (2017)
13. Conrad, K.: Galois groups of cubics and quadrics
14. Fougeron, C.: Dynamical properties of simplicial systems and continued fraction algorithms. arXiv:2001.01367v3 (2020)
15. Furstenberg, Harry: Noncommuting random products. Trans. Am. Math. Soc. 108, 377–428 (1963)
16. Garrity, Thomas: On periodic sequences for algebraic numbers. J. Number Theory 88(1), 86–103 (2001)
17. Garrity, T.: On Gauss–Kuzmin statistics and the transfer operator for a multidimensional continued fraction algorithm: the triangle map (2015)
18. Goldsheid, I.Y., Margulis, G.A.: Lyapunov exponents of a product of random matrices. Uspekhi Mat. Nauk 445(269), 13–60 (1989)
19. ) Guivarc’h, Y., Raugi, A.: Products of random matrices: convergence theorems. In Random matrices and their applications (Brunswick, Maine, 1984), Contemp. Math., Amer. Math. Soc., vol. 50, pp. 31–54, Providence, RI (1986)
20. Lagarias, J.C.: The quality of the Diophantine approximations found by the Jacobi–Perron algorithm and related algorithms. Monatsh. Math. 115(4), 299–328 (1993)
21. Mañé, R.: Ergodic theory and differentiable dynamics. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 8. Springer, Berlin (1987). Translated from the Portuguese by Silvio Levy
23. Matheus, Carlos, Möller, Martin, Yoccoz, Jean-Christophe: A criterion for the simplicity of the Lyapunov spectrum of square-tiled surfaces. Invent. Math. 202(1), 333–425 (2015)
24. Messaoudi, Ali, Nogueira, Arnaldo, Schweiger, Fritz: Ergodic properties of triangle partitions. Monatsh. Math. 157(3), 283–299 (2009)
25. Paley, R.E.A.C., Ursell, H.D.: Continued fractions in several dimensions. Proc. Camb. Philos. Soc. 26, 127–144 (1930)
26. Schratzberger, B.R.: The exponent of convergence for Brun’s algorithm in two dimensions. Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II, 207:229–238 (1999), 1998
27. Schweiger, Fritz: Multidimensional Continued Fractions. Oxford Science Publications. Oxford University Press, Oxford (2000)
28. Ernst, S.S.: Continued fractions in several dimensions. Nordisk Nat. Tidskr. 9, 37–43 (1961)
29. Torres, A.H.: Simplicity of Lyapunov spectrum for multidimensional continued fraction algorithms. Ph.D. Thesis, IMPA, (2009), unpublished thesis
30. Viana, M.: Stochastic dynamics of deterministic systems. Brazilian Math. Colloquium. IMPA (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.