Rational approximations, multidimensional continued fractions, and lattice reduction

Berthé, V.; Dajani, K.; Kalle, C.C.C.J.; Krawczyk, E.; Kuru, H.; Thevis, A.; ... ; Smajlovic, L.

Rational Approximations, Multidimensional Continued Fractions, and Lattice Reduction

V. Berthé, K. Dajani, C. Kalle, E. Krawczyk, H. Kuru, and A. Thevis

1 Introduction

Continued fraction type expansions aim (among other properties) at providing increasingly good rational Diophantine approximations of real numbers. More precisely, a multidimensional continued fraction is expected to produce simultaneous better and better rational approximations with the same denominator $\mathbf{p}^{(n)}/q^{(n)} = (p_1^{(n)}/q^{(n)}, \dots p_d^{(n)}/q^{(n)})_{n\in\mathbb{N}}$ for d-tuples $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d)$ of real numbers, with the fractions $p_i^{(n)}/q^{(n)}$ converging to α_i for each $1 \le i \le d$.

The usual regular continued fractions are known to provide extremely good (and even the best) rational approximations for positive real numbers [41, 89]. The

V. Berthé
Université Paris Cité, CNRS, Paris, France
e-mail: berthe@irif.fr

K. Dajani
Department of Mathematics, Utrecht University, Utrecht, The Netherlands
e-mail: k.dajani1@uu.nl

C. Kalle
Mathematisch Instituut, Leiden University, Leiden, The Netherlands
e-mail: kallecccj@math.leidenuniv.nl

E. Krawczyk
Faculty of Mathematics and Computer Science, Institute of Mathematics, Jagiellonian University, Kraków, Poland

H. Kuru
Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Turkey
e-mail: hamidekuru@sabanciuniv.edu

A. Thevis (✉)
Institut für Mathematik, Goethe-Universität Frankfurt, Frankfurt am Main, Germany
e-mail: thevis@math.uni-frankfurt.de

R. Abdellatif et al. (eds.), *Women in Numbers Europe IV*, Association for Women in Mathematics Series 32, https://doi.org/10.1007/978-3-031-52163-8_5

situation is more complicated in higher dimension. Indeed, there is no canonical extension of regular continued fractions to higher dimensions (see Sect. 4.1), and the zoology of existing algorithms is particularly rich (see Sect. 4.5 as an illustration). The main advantage of most classical (unimodular) continued fractions is that they can be expressed as dynamical systems whose ergodic study has already been well understood (such as described in Sect. 4.4). Ergodic theory allows a precise description of the long-range statistical properties of the expansions that are produced (e.g. their mean behavior). Indeed, ergodic theory extends basic laws of large numbers in probability by dropping the assumption of intertemporal independence. Thus, it relates spatial averages $\int_X f\, d\mu$ to time averages $\frac{1}{n}\sum_{0\le i<n} f\circ T^i$ along trajectories; in other words, the system has the same behavior when averaged over time as averaged over the whole space.

However, the main disadvantages of these algorithms lie, firstly, in the fact that the behavior of the continued fraction expansion of a given d-tuple $\boldsymbol{\alpha}$ can be difficult to grasp (it might not behave in a generic way), and secondly, in the quality of the rational approximations that are produced. Indeed, the convergence of multidimensional continued fractions is governed by their (first and second) Lyapunov exponents (see [105]), which describe the asymptotic behavior of the singular values of large products of matrices, under the ergodic hypothesis. More precisely, their approximation exponent can be expressed as $1-\frac{\lambda_2}{\lambda_1}$ according to [105] (λ_1 and λ_2 being the two largest Lyapunov exponents of the associated dynamical system). It has to be compared with Dirichlet's exponent $1+1/d$ (see Theorem 1 in Sect. 3.1). However, there is numerical evidence [38] that the second Lyapunov exponent is not even negative in higher dimensions for most classical algorithms such as the Jacobi–Perron [20, 124, 134], Brun [34–36], or Selmer [142] algorithms, which prevents strong convergence of these algorithms. In a nutshell, strong (resp. weak) convergence refers to the convergence of quantities of the type $|||q^{(n)}\boldsymbol{\alpha}|||$, where $|||\cdot|||$ stands for the distance to the nearest integer (resp., $||\alpha-\mathbf{p}^{(n)}/q^{(n)}||$). In other words, these algorithms converge weakly usually, but they fail to have strong convergence (see Sect. 3.3 for the definitions of weak and strong convergence).

In terms of the quality of rational approximations that are produced, there is a second strategy which relies on lattice reduction, where rational approximations are obtained by exhibiting short vectors in a lattice attached to some given d-tuple $\boldsymbol{\alpha}$. Lattice reduction algorithms aim to find reduced bases of Euclidean lattices, formed by short and almost orthogonal vectors. The most celebrated one is the LLL algorithm, designed by Lenstra, Lenstra and Lovász in 1982 [108]. It relies heavily on the use of Gram–Schmidt orthogonalization. Its overall algorithmic structure is simple and yet, its general probabilistic behavior is far from being understood; this includes the gap between its practical performances and its proved worst-case estimates. Hence, although finding rational approximations works quite well in practice, the average behavior of such a strategy is not well understood. In particular, the lack of a description of reduction algorithms as dynamical systems prevents the use of tools from ergodic theory.

In the expository part of the paper, we focus on the two main classes of algorithms that produce rational approximations as discussed above. The first type of algorithms can be expressed via dynamical systems defined on a compact set (usually of the form $[0, 1]^d$). Such an algorithm associates with some given vector an infinite sequence of matrices, and one can consider the quality of convergence of this product of matrices. The most classical examples of such algorithms are the Jacobi–Perron [20, 68, 124, 134], the Brun [34–36], or the Selmer algorithms (the last of which is conjugate on the absorbing simplex to Mönkemeyer's algorithm [111, 123]). They are described, e.g., in [31, 105, 138]. The second type of algorithm is based on lattice reduction algorithms, such as the LLL algorithm (see Sect. 5). We focus on these two families since they share as a common feature the fact that they rely on a choice of a basis of the integer lattice $\mathbb{Z}^d$.

We then illustrate in Sect. 7 the dynamical approach with the ergodic study of a version of the Jacobi–Perron algorithm based on the use of the nearest integer part for the partial quotients. One motivation for studying the nearest integer Jacobi–Perron algorithm is to confirm the idea that working with the nearest integer part improves the quality of continued fraction algorithms, such as indicated numerically by the experimental results from Steiner [145] (see Sect. 7.2). The partial quotients produced by the usual Jacobi–Perron algorithm satisfy a simple Markovian rule. In the case of the nearest integer Jacobi–Perron algorithm the description of the admissible sequences of digits is much more involved. Hence, a simple modification—such as changing the choice of the integer part—leads to much more delicate conditions for the description of the algorithm. As a first step toward a theoretical confirmation of the above-mentioned estimates, we prove the existence of a Markov partition for the nearest integer Jacobi–Perron algorithm and suggest a possible procedure for proving the existence of a finite ergodic invariant measure absolutely continuous with respect to Lebesgue measure.

Let us sketch the contents of this paper. Section 2 recalls basic notions concerning classical continued fractions. We present their main properties that we will use as a guideline for possible generalizations to the higher-dimensional case. We then focus in Sect. 3 on the two main strategies that can be used for producing rational approximations in an effective way. Section 4 deals with the classical dynamical unimodular continued fraction algorithms. Algorithms based on the lattice reduction algorithms and homogeneous dynamics are considered in Sect. 5. We briefly discuss applications and possible ways to improve algorithms in Sect. 6. Lastly, in Sect. 7 we focus on the nearest integer Jacobi–Perron algorithm. We describe its associated Markov partition and provide a strategy for proving the existence of an absolutely continuous invariant measure.

2 Continued Fractions

In this section we briefly recall the main properties of the usual regular continued fractions. They will serve as a guideline for the discussion on the higher-dimensional case. For general references on continued fractions, see, e.g., [24, 51, 74, 75, 89]. For any positive real number $\alpha \in [0, 1]$, its continued fraction expansion is

$$\alpha = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}},$$

where the digits a_n are positive integers, called *partial quotients*. The rational numbers p_n/q_n, where p_n, q_n are coprime positive integers defined as

$$\frac{p_n}{q_n} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}},$$

are called *convergents*. The sequence of rational numbers p_n/q_n approximates α up to an error of order $1/q_n^2$: one has

$$|\alpha - p_n/q_n| \leq \frac{1}{q_n^2} \quad \text{for all } n.$$

Dynamically, continued fraction expansions can be obtained by applying the Gauss map $T_G : [0, 1] \to [0, 1]$ defined by

$$T_G(0) = 0 \quad \text{and} \quad T_G(\alpha) = \{1/\alpha\} \text{ if } \alpha \neq 0,$$

where $\{\cdot\}$ is the fractional part of a real number. If for an $\alpha \in (0, 1]$ we write $T_G(\alpha) = \{1/\alpha\} = \frac{1}{\alpha} - \lfloor \frac{1}{\alpha} \rfloor = \frac{1}{\alpha} - a_1$, then $\alpha = \frac{1}{a_1 + T_G(\alpha)}$. Now, by setting

$$a_n = \left\lfloor \frac{1}{T_G^{n-1}(\alpha)} \right\rfloor$$

for $n \geq 1$, one gets the digits in the continued fraction expansion of α.

For $\alpha \notin (0, 1]$, one sets $a_0 = \lfloor \alpha \rfloor$ and one considers the digits generated by T_G for $\alpha - a_0$. Alternatively we can use the following form for $\gamma_0 := 1/\alpha$ in $\mathbb{R}^+$, obtained by setting, for $n \geq 0$,

$$\begin{cases} a_n &= \lfloor \gamma_n \rfloor \\ \gamma_{n+1} &= \frac{1}{\gamma_n - a_n}. \end{cases}$$

One then has $\gamma_n = \frac{1}{T_G^{n-1}(\alpha)}$ for all $n \geq 1$.

Note that the Gauss map is closely related to Euclid's algorithm: starting with two (coprime) positive integers $\ell^{(0)}$ and $\ell^{(1)}$, Euclid's algorithm works by subtracting as often as possible the smallest of both numbers from the largest one (that is, one performs the Euclidean division of the largest one by the smallest); this yields $\ell^{(0)} = \ell^{(1)} \lfloor \frac{\ell^{(0)}}{\ell^{(1)}} \rfloor + \ell^{(2)}$, $\ell^{(1)} = \ell^{(2)} \lfloor \frac{\ell^{(1)}}{\ell^{(2)}} \rfloor + \ell^{(3)}$, etc., until we reach $\ell^{(m+1)} = 1 = \gcd(\ell^{(0)}, \ell^{(1)})$. By setting, for $n \in \mathbb{N}$, $\alpha^{(n)} = \frac{\ell^{(n)}}{\ell^{(n+1)}}$ and $a_n = \lfloor \alpha^{(n)} \rfloor$, one gets $\alpha^{(n-1)} = a_{n-1} + \frac{1}{\alpha^{(n)}}$, and

$$\alpha^{(0)} = \frac{\ell^{(0)}}{\ell^{(1)}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots + \frac{1}{a_{m-1} + \frac{1}{a_m}}}}}.$$

Let us now revisit the action of the Gauss map in matricial terms. Let $\alpha \in [0, 1]$. For all n, we have

$$\begin{aligned} \begin{bmatrix} \alpha \\ 1 \end{bmatrix} &= \alpha T_G(\alpha) \cdots T_G^{n-1}(\alpha) \begin{bmatrix} 0 & 1 \\ 1 & a_1 \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & a_n \end{bmatrix} \begin{bmatrix} T_G^n(\alpha) \\ 1 \end{bmatrix} \\ &= \alpha T_G(\alpha) \cdots T_G^{n-1}(\alpha) \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix} \begin{bmatrix} T_G^n(\alpha) \\ 1 \end{bmatrix}, \end{aligned}$$

by using the classical relations between convergents and partial quotients, namely $q_{-1} = 0$, $p_{-1} = 1$, $q_0 = 1$, $p_0 = 0$, and, for all n,

$$q_{n+1} = a_{n+1} q_n + q_{n-1}, \;\; p_{n+1} = a_{n+1} p_n + p_{n-1}.$$

The matrix $\begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix}$ is a square matrix with integer entries that has determinant of absolute value 1, which is to say that it is *unimodular*. We denote the set of unimodular matrices by $\mathrm{GL}(d, \mathbb{Z})$. (We also use the following standard notation: $\mathrm{GL}(d, \mathbb{R})$ stands for the set of $d \times d$ invertible matrices with real entries, $SL(d, \mathbb{N})$ stands for the set of $d \times d$ matrices of determinant 1 with non-negative integer coefficients.) Note that the entries of this matrix are even positive. To understand the convergence of such a sequence of matrices one can use generalizations of the Perron–Frobenius theorem, such as Theorem 2 below.

Dynamically, the Gauss map T_G goes with the map

$$A_G : [0,1] \to \mathrm{GL}(2,\mathbb{N}),\ \alpha \mapsto \begin{bmatrix} 0 & 1 \\ 1 & \lfloor 1/\alpha \rfloor \end{bmatrix}. \tag{1}$$

Such a map is called a *cocycle* in the terminology of (random) dynamical systems (see, e.g., [10, 11, 151]). Let us set $A^{(n)} := A(T_G^n)(\alpha)$ for all positive n. The line directed by the vector $(\alpha, 1)$ in $\mathbb{R}^2$ belongs to the sequence of nested cones $A^{(1)} \cdots A^{(n)}\mathbb{R}_+^2$, i.e.,

$$(\alpha, 1) \in \bigcap_n A^{(1)} \cdots A^{(n)}\mathbb{R}_+^2 = \bigcap_n A(T_G)(\alpha) \cdots A(T_G^n)(\alpha)\mathbb{R}_+^2.$$

One has even more: this sequence of nested cones converges to the line directed by $(\alpha, 1)$. The *convergence* is said to be *weak* if the angles between the (column) vectors of the product matrices $A^{(1)} \cdots A^{(n)}$ tend to 0 (as $n \to \infty$), and *strong* if the distances between the vectors tend to 0. Continued fractions can thus be seen as producing dynamically infinite convergent sequences of unimodular matrices. We revisit the notions of convergence of infinite products of matrices in Sect. 3.3 in more detail.

3 On Simultaneous Approximation

This section is devoted to simultaneous rational approximations. In Sect. 3.1 we first recall Dirichlet's theorem. We then describe in Sect. 3.2 the two main strategies for producing such approximations that we consider in this survey. We focus on the quality of approximations in Sect. 3.3.

3.1 Dirichlet's Bound

Let us recall Dirichlet's theorem; it can be obtained as a direct application of the pigeonhole principle (see, e.g., [74]) or of Minkowski's first theorem.

Theorem 1 (Dirichlet's Theorem) *For any $(\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ and any Q, there exists a positive integer q with $q \leq Q^d$ and integers p_i such that*

$$\max_{1 \leq i \leq d} |q\alpha_i - p_i| < \frac{1}{Q}.$$

Theorem 1 immediately implies that for any $(\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ the system of inequalities

$$\left|\frac{p_i}{q} - \alpha_i\right| < \frac{1}{q^{1+\frac{1}{d}}}, \quad \text{for } i = 1, 2, \ldots, d,$$

admits infinitely many integer solutions $(p_1, \ldots, p_d, q)$.

The exponent $1 + 1/d$ is optimal as shown in [125], see also [41, 136]. The Dirichlet's theorem provides the existence of "good" approximations. One can thus make an exhaustive search, though this is not an efficient algorithmic method. See [121, Chapter 6] for a discussion on effective methods. See also [104], where the computational complexity concerning simultaneous Diophantine approximations is investigated. When the dimension d is fixed, [104] gives algorithms which, for a given N, find a good rational approximation with denominator $1 \leq q \leq N$ with respect to a specified accuracy, or which find all best approximations (in the sense of Definition 1 below) located in $[1, \cdots, N]$ in polynomial time (using methods based on the LLL algorithm). Note that the following problem is proved to be NP-hard: for a given vector $\boldsymbol{\alpha} \in \mathbb{Q}^d$, positive integer N and accuracy s_1/s_2, is there an integer Q with $1 \leq Q \leq N$ such that $|||Q\boldsymbol{\alpha}||| \leq s_1/s_2$? (The distance to the nearest integer is expressed here with respect to the supremum norm.) In a similar flavor, see also [69] concerning the problem of finding integer relations, and [25].

Continued fractions are known to provide good (and even the best) rational approximations of a given real number α in [0, 1] (see, e.g., [41, 89]). One would desire to have similar algorithms yielding good rational approximations with the same denominator of d-tuples of positive real numbers. That is, for a given $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in [0, 1]^d$, one looks for sequences of positive integers $(q_n)_n$ and positive integer d-tuples $\mathbf{p}^{(n)} = (p_1^{(n)}, \cdots, p_d^{(n)})_n$ such that

$$\lim p_i^{(n)}/q_n = \alpha_i, \quad i = 1, \cdots, d$$

with a good quality of rational approximation of $\boldsymbol{\alpha}$. Geometrically, this corresponds to looking for approximations of a line in $\mathbb{R}^{d+1}$ by points in $\mathbb{Z}^{d+1}$. Dual problems consist of looking for small values of linear forms and small linear relations.

More precisely, given a norm $||\cdot||$ on $\mathbb{R}^d$, let $|||\cdot|||$ stand for the distance to the nearest integer. The usual norms that are considered are the supremum and the Euclidean norm. The quality of the approximation is measured by $\frac{1}{q^{(n)}}|||q^{(n)}\boldsymbol{\alpha}|||$, to be compared with Dirichlet's bound, i.e., $|||q^{(n)}\boldsymbol{\alpha}|||$ has to be compared with $(q^{(n)})^{-1/d}$. One can thus consider the approximation exponent

$$\limsup_n -\frac{\log \|\boldsymbol{\alpha} - \frac{\mathbf{p}_i^{(n)}}{q_i^{(n)}}\|}{\log q^{(n)}}$$

and compare it to the Dirichlet's bound $1 + 1/d$.

3.2 How to produce Rational Approximations

We focus here on two main approaches for producing rational approximations. The first one is based on the generation of infinite convergent sequences of matrices obtained dynamically by iteration of a map acting on a compact space, as illustrated with the Gauss map T_G for usual continued fractions in Sect. 2. This will be discussed further in Sect. 4. The second one is based on the existence of small vectors picked in well chosen lattices; we will discuss it in Sect. 5.

The first strategy associates with some element $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ a sequence of square matrices $(A^{(n)})_{n\in\mathbb{N}}$ of size $d+1$ with integer entries. It can be produced for instance via a dynamical system (X, T) with a map A as follows (see also (1)):

$$T\colon X \to X,\ A\colon X \to GL(d+1,\mathbb{Z}),\ \text{and } A^{(n)} = A(T^n(x)). \tag{2}$$

If the matrices belong to $GL(d+1,\mathbb{Z})$, then the corresponding algorithm is called *unimodular*. Matrices $A^{(n)}$ play the role of partial quotients and the product matrices $A^{(1)}\cdots A^{(n)}$ produce convergents. Convergents aim at providing Diophantine approximations (via their column vectors) of the direction $(\boldsymbol{\alpha}, 1)$. We write

$$A^{(1)}\cdots A^{(n)} = \begin{bmatrix} p_{1,1}^{(n)} & \cdots & p_{1,d+1}^{(n)} \\ \vdots & \ddots & \vdots \\ p_{d,1}^{(n)} & \cdots & p_{d,d+1}^{(n)} \\ q_1^{(n)} & \cdots & q_{d+1}^{(n)} \end{bmatrix}. \tag{3}$$

The last element of each column of $A^{(1)}\cdots A^{(n)}$ is a denominator for the associated simultaneous rational approximation. The first d rows of the convergent matrices are meant to provide the numerators of the simultaneous approximations, i.e., one considers the following vector

$$\left(\frac{p_{j,1}^{(n)}}{q_j^{(n)}}, \cdots, \frac{p_{j,d}^{(n)}}{q_j^{(n)}}\right),$$

whose entries are the *nth convergents* of $\boldsymbol{\alpha}$.

The convergence of these matrices means that they contract in the direction of the vector $(\boldsymbol{\alpha}, 1)$. We discuss more precisely their convergence in Sect. 3.3 by considering products $A^{(1)}\cdots A^{(n)}$ as n goes to infinity.

We now describe the second approach based on the existence of small vectors in well chosen lattices, as described in the seminal paper [108]. This approach yields a very fruitful compromise between the quality of approximation (a good approximation is deduced from a small vector) and the efficiency (this small vector is obtained in polynomial time). Let $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d) \in \mathbb{R}^d$ be a vector to

approximate. One works here again in a $d+1$-dimensional space, by introducing a one-parameter family of lattices $(\Lambda_t)_{t>0}$ with positive parameter t tending to 0. More precisely, let Λ_t be the lattice generated by the columns of the matrix

$$M_t := \begin{bmatrix} 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_d \\ 0 & 0 & \cdots & 0 & t \end{bmatrix}.$$

Note that $\det(M_t) = t$, hence, the lattice Λ_t changes at each step of the algorithm. Let us stress the fact that this strategy differs from the previous one where, in the unimodular case, one works with bases of the fixed lattice $\mathbb{Z}^{d+1}$. We will take t small, the parameter Q of Dirichlet's theorem being connected to t as follows: $Q = t^{-\frac{1}{d+1}}$.

One of the main features of the LLL algorithm is that it produces in polynomial time a non-zero vector $\mathbf{b} = (b_1, \cdots, b_{d+1})$ of the lattice Λ_t such that

$$||\mathbf{b}||_2 \leq 2^{d/4} \det(M_t)^{1/(d+1)} = 2^{d/4} t^{1/(d+1)}. \tag{4}$$

Note that the geometry of numbers, and more precisely Minkowski's first theorem, guarantees the existence of a "small" non-zero vector $\mathbf{x} \in \Lambda_t$, i.e., such that

$$||\mathbf{x}||_2 \leq \sqrt{(d+1)(d+5)/4}\, (\mathrm{vol}(\Lambda_t))^{1/(d+1)} = \sqrt{(d+1)(d+5)/4}\, t^{1/(d+1)}. \tag{5}$$

Let $(\mathbf{e}_i)_{i=1,\ldots,d+1}$ stand for the canonical basis of $\mathbb{Z}^{d+1}$. There exist integers $p_1, \ldots, p_d, q$ such that

$$\begin{aligned} \mathbf{b} &= p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + \cdots + p_d\mathbf{e}_d + q(-\alpha_1\mathbf{e}_1 - \cdots - \alpha_d\mathbf{e}_d + t\mathbf{e}_{d+1}) \\ &= (p_1 - q\alpha_1)\mathbf{e}_1 + \cdots + (p_d - q\alpha_d)\mathbf{e}_d + qt\mathbf{e}_{d+1}. \end{aligned}$$

One deduces from (4) that for all $1 \leq i \leq d$

$$|p_i - \alpha_i q| \leq 2^{d/4} t^{1/(d+1)}$$

and

$$qt \leq 2^{d/4} t^{1/(d+1)}, \text{ i.e., } t^{\frac{1}{d+1}} \leq \frac{2^{1/4}}{q^{1/d}}.$$

For all i, we deduce that

$$|p_i - \alpha_i q| \leq \frac{2^{(d+1)/4}}{q^{1/d}},$$

with

$$|q| \leq 2^{d/4} t^{-d/(d+1)} = 2^{d/4} Q^d.$$

The quality of approximation is the quality that is expected (according to Dirichlet's theorem) up to a multiplicative factor $2^{(d+1)/4}$ which depends exponentially on the dimension. We could have used the inequality (5) which would have given a different multiplicative factor, but the same quality ($q^{1/d}$). Nevertheless, the interest of a lattice reduction algorithm such as the LLL algorithm is that the small vector that is used is found in polynomial time.

For a given $t > 0$, the smallest vector of the lattice Λ_t (or a small vector in the sense of (5)) produces a good approximation. One can ask whether it is possible to devise a continued fraction algorithm from this. Note that one has to recompute everything from the beginning when one changes t. This will be discussed further in Sect. 5. Algorithms defined dynamically with maps such in (2) are on the contrary called memory-less (see Sect. 4).

3.3 Convergence and Lyapunov Exponents

One important feature of the first strategy (based on the generation of infinite products of matrices) from Sect. 3.2 is that we are still able to measure the quality of approximation that is produced in view of Dirichlet's theorem. To do this, we need to measure the quality of convergence of infinite products of matrices; this can be done using Lyapunov exponents. One can either measure the convergence of a given product of matrices or consider the generic behavior of products of matrices generated by a dynamical system. Lyapunov exponents allow, among other things, the description of the growth of the logarithm of the angles between column vectors of products of matrices $M_0 \cdots M_n$.

We state now the definitions concerning convergence. The norm $\| \cdot \|$ refers to the Euclidean norm. Let $M = (M_n)_{n\in\mathbb{N}}$ be a sequence of square matrices of size d and let $\boldsymbol{\ell} \in \mathbb{R}^d$ be a non-zero vector. Let $(\mathbf{e}_1, \cdots, \mathbf{e}_d)$ stand for the canonical basis of $\mathbb{R}^d$. We say that M is *weakly convergent* to $\boldsymbol{\ell}$ if

$$\lim_{n\to+\infty} \left\| \frac{M_0 \cdots M_{n-1}\mathbf{e}_i}{\|M_0 \cdots M_{n-1}\mathbf{e}_i\|} - \frac{\boldsymbol{\ell}}{\|\boldsymbol{\ell}\|} \right\| = 0 \quad \text{for all } i \in \{1, \ldots, d\}.$$

We say that M is *strongly convergent* to $\boldsymbol{\ell}$ if

$$\lim_{n\to+\infty} \mathrm{d}(M_0 \cdots M_{n-1}\mathbf{e}_i, \mathbb{R}\boldsymbol{\ell}) = 0 \quad \text{for all } i \in \{1, \ldots, d\},$$

where the distance d refers to the usual associated distance of a point to a line, i.e., the distance between a point $\mathbf{x}$ and the set $\mathbb{R}\boldsymbol{\ell}$ is the infimum of $\{\|\mathbf{x} - r\boldsymbol{\ell}\| \mid r \in \mathbb{R}\}$.

Lastly, we say that M is *exponentially convergent* to $\boldsymbol{\ell}$ if there exist $C, \gamma > 0$ such that

$$\mathrm{d}(M_0 \cdots M_{n-1}\mathbf{e}_i, \mathbb{R}\boldsymbol{\ell}) < Ce^{-\gamma n} \quad \text{for all } n \in \mathbb{N} \text{ and for all } i \in \{1, \ldots, d\}.$$

Exponential convergence is a first step in the direction of good rational approximations. The constant γ then has to be compared with Dirichlet's bound.

Theorem 2 below gives a sufficient condition for the sequence of cones $M_0 \cdots M_n \mathbb{R}_+^d$ to nest down to a single line as n tends to infinity for square matrices M_i with non-negative entries. It can be seen as a generalization of the classical Perron–Frobenius theorem. This statement is particularly useful in the present context since multidimensional continued fraction algorithms often generate non-negative matrices. Hence, weak convergence is usually not an issue for multidimensional continued fraction algorithms.

Theorem 2 ([64, pp. 91–95]) *Let $(M_n)_n$ be a sequence of non-negative integer matrices of size d. Assume that there exist a strictly positive matrix B and indices $j_1 < k_1 \leq j_2 < k_2 \leq \cdots$ such that $B = M_{j_1} \cdots M_{k_1-1} = M_{j_2} \cdots M_{k_2-1} = \cdots$. Then,*

$$\bigcap_{n\in\mathbb{N}} M_0 \cdots M_{n-1}\mathbb{R}_+^d = \mathbb{R}_+\boldsymbol{\ell} \quad \textit{for some positive vector } \boldsymbol{\ell} \in \mathbb{R}_+^d.$$

Let us now focus on strong convergence by first recalling the definition of the Lyapunov exponents of a sequence of matrices. For a matrix M in $\mathrm{GL}(d, \mathbb{R})$, the singular values $\delta_1, \ldots, \delta_d$ are the eigenvalues of the matrix $\left({}^tMM\right)^{1/2}$. Let us order these (positive and real) values as $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_d$. Given a sequence $\mathbf{M} = (M_n)_{n\in\mathbb{N}}$ of matrices in $\mathrm{GL}(\mathrm{d}, \mathbb{R})$, the *$i$-th Lyapunov* exponent θ_i is then defined as the limit

$$\theta_i := \lim_{n\to\infty} \frac{1}{n} \log(\delta_i(n)),$$

if this limit exists, with $\delta_i(n)$ being the i-th singular value of $M_0 \cdots M_{n-1}$. The Lyapunov exponents can also be defined recursively using exterior powers (see for example [11, Proposition 3.2.7]) by

$$\theta_1 + \cdots + \theta_k = \lim_{n\to\infty} \frac{1}{n} \log\|\wedge^k(M_0 \cdots M_{n-1})\|, \quad k = 1, \ldots, d,$$

provided that the limit exists. One has $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_d$. A sufficient condition for strong convergence can then be stated as follows (see [11, Proposition 3.4.2 (ii)] and [2]): Let $\mathbf{M} = (M_n)_{n\in\mathbb{N}}$ be a sequence of non-negative matrices in $\mathrm{GL}(d, \mathbb{N})$ for which the Lyapunov exponents exist; if $\mathbf{M}$ satisfies the growth condition $\limsup_{n\to\infty} \frac{1}{n} \log\|M_n\| \leq 0$ together with the condition $\theta_1 > 0 > \theta_2$, then strong convergence holds.

Lyapunov exponents can be also defined for random products of matrices, where the randomness can be provided by putting some distribution on the set of matrices or by iterating measurable dynamical systems that produce matrices with a cocyle map such as in (1). The first results in this direction were stated for sequences of independent random matrices with a given distribution function, with the Furstenberg–Kesten theorem (see [61, 65]); these results have then been refined via Kingman's subadditive ergodic theorem and lastly via Oseledets' multiplicative ergodic theorem (see for instance [10, 11, 151]) proving that the limits (involved in the definition of Lyapunov exponents) exist almost surely and take almost everywhere the same value.

Let us give a flavor of such results. We will come back to this also in Sect. 4.4. We first recall a few elements from ergodic theory. The (left) *shift* S acts on a sequence $(M_n)_{n\in\mathbb{N}}$ as $S((M_n)_n) = (M_{n+1})_n$ (i.e., the first term of the sequence $(M_n)_n$ is deleted). Let $\mathcal{M}$ be a finite set of matrices in $\mathrm{GL}(d,\mathbb{Z})$. Let $D \subset \mathcal{M}^{\mathbb{N}}$ be a closed shift-invariant subset of $\mathrm{GL}(d,\mathbb{Z})^{\mathbb{N}}$. A probability measure ν on D is called shift-invariant if $\nu(S^{-1}B) = \nu(B)$ for every measurable set $B \subset D$. A shift-invariant probability measure on D is *ergodic* if any shift-invariant measurable set has either measure 0 or 1. We have seen (see (2)) that such a set of matrices can be obtained by considering a measurable map T acting on some compact metric space and a measurable map $A : X \to \mathrm{GL}(d,\mathbb{Z})$. The sequences of matrices are then of the form $A(T^n(x))$ and one studies the existence and the almost everywhere behavior of limits of the form

$$\lim_{n\to\infty} \frac{1}{n} \log \|A(x)\cdots A(T^{n-1}(x))\|.$$

Theorem 3 ([11]) *Let $D \subset \mathcal{M}^{\mathbb{N}}$ be a closed shift-invariant subset of $\mathrm{GL}(d,\mathbb{Z})^{\mathbb{N}}$ together with a shift-invariant measure ν. Assume that (D, S, ν) is ergodic. Let $A : D \to \mathrm{GL}(d,\mathbb{Z})$, $\mathbf{M} = (M_n)_n \mapsto M_0$. Assume that A is log-integrable, i.e.,*

$$\int_D \log \max\{\|A(\mathbf{M})\|_\infty, \|A(\mathbf{M})^{-1}\|_\infty\} d\nu(\mathbf{M}) < \infty. \tag{6}$$

Then the quantities θ_i which are recursively defined by

$$\theta_1 + \cdots + \theta_k = \lim_{n\to\infty} \frac{1}{n} \log\|\wedge^k (M_0 \cdots M_{n-1})\|, \quad k = 1, \ldots, d, \tag{7}$$

exist and do not depend on $\mathbf{M}$ for almost all $\mathbf{M} \in D$. Here, $\wedge^k$ stands for the k-fold wedge product.

4 Higher-Dimensional Dynamical Continued Fractions

We now focus on multidimensional continued fraction algorithms. Their non-canonicity is discussed in Sect. 4.1. Section 4.2 recalls what is usually expected from a continued fraction algorithm. In Sect. 4.3 we focus on algorithms obtained by iteration of a dynamical system, which yield the so-called memory-less algorithms. The interest of their dynamical description is highlighted in Sect. 4.4. Lastly, we give examples of such algorithms in Sect. 4.5.

4.1 Non-Canonicity in Higher Dimension

The aim of this section is to present several facts sustaining the claim that there is no canonical multidimensional continued fractions algorithm.

Firstly, (usual) continued fractions rely on Euclid's algorithm: starting with two numbers, one subtracts the smallest from the largest (see Sect. 2). If we start with at least three numbers, it is not clear how to decide which operation has to be performed, hence the diversity of existing generalizations. For instance, Brun's algorithm can be described as subtracting the second largest entry from the largest one. See Sect. 4.5 for more details.

Moreover, the specific algebraic structure of $SL(2,\mathbb{N})$ plays an important role for one-dimensional continued fractions algorithms. Indeed, the matrices produced in the case of usual continued fractions are unimodular matrices with non-negative integer coefficients (see Sect. 2). The algebraic structure of $SL(2,\mathbb{N})$ is particularly simple: $SL(2,\mathbb{N})$ is a free and finitely generated monoid; it admits

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{8}$$

as generators; any matrix in $SL(2,\mathbb{N})$ thus admits a unique decomposition in terms of the matrices given in (8). This decomposition is a matricial translation of Euclid's algorithm, and the continued fraction expansion of α can be recovered from the unique decomposition of matrices

$$(-1)^n \begin{bmatrix} p_{n+1} & q_{n+1} \\ p_n & q_n \end{bmatrix}, \quad n \geq 0$$

in the free monoid $SL(2,\mathbb{N})$. This explains why most one-dimensional continued fraction algorithms are closely related.

The situation is completely different for $SL(3,\mathbb{N})$ which is not finitely generated. Consider, e.g. the family of matrices

$$M_n := \begin{bmatrix} 1 & 0 & n \\ 1 & n-1 & 0 \\ 1 & 1 & n-1 \end{bmatrix}.$$

According to [131, Chap. 12] these matrices are undecomposable for $n \geq 3$: they are not equal to an even permutation matrix, and, for any pair of matrices $A, B \in SL(3, \mathbb{N})$ such that $M_n = AB$, A or B is an even permutation matrix.

Another approach for generalizing continued fractions could rely on properties of best approximation.

Definition 1 A rational number p/q is said to be a *best approximation* of a real number α if every p'/q' with $1 \leq q' \leq q$, $p/q \neq p'/q'$ satisfies

$$|q\alpha - p| < |q'\alpha - p'|.$$

Convergents in the continued fraction expansion of α and best approximations are known to coincide [41, 89]. Nevertheless, this notion is not so satisfying for defining continued fractions in higher dimensions as stressed in [101, 102]. Firstly, best approximations depend on the choice of a norm [101], and secondly, the unimodularity property is lost. More precisely, one has the following.

Definition 2 Let $\boldsymbol{\alpha} \in [0, 1]^d$. Let $||\cdot||$ be a given norm in $\mathbb{R}^d$ and let $|||\cdot|||$ denote the distance to the nearest integer. The *sequence of best approximations* of $\boldsymbol{\alpha}$ with respect to the norm $||\cdot||$ is defined as the increasing sequence of non-negative integers $(q^{(n)})_{n\in\mathbb{N}}$ such that $|||q^{(n)}\boldsymbol{\alpha}||| < |||q\boldsymbol{\alpha}|||$ for any q with $1 \leq q < q^{(n)}$.

The existence of an infinite sequence of best approximations can be derived in a classic way from Dirichlet's theorem or from Minkowski's first theorem. Best approximations fail to be unimodular [102]. More precisely, consider the square matrix M_n of size $d+1$ whose rows are given by successive best approximations vectors $\mathbf{v}_n = (p_1^{(n)}, \cdots, p_d^{(n)}, q^{(n)})$ providing $|||q^{(n)}\boldsymbol{\alpha}|||$. Let D_n stand for the determinant of this matrix. It is proved in [102] that for any norm in dimension $d \geq 2$, there exists $\boldsymbol{\alpha} \in \mathbb{R}^d$, with $\dim_{\mathbb{Q}}[1, \alpha_1, \cdots, \alpha_d] = d+1$, such that for any positive integer N, there exists an n for which $D_n = D_{n+1} = \cdots = D_{n+N} = 0$. Arbitrarily large determinants can even occur in dimension $d = 2$ with the supremum norm. For more on best approximations and multidimensional continued fractions, see the survey [46]; see also [115] and Sect. 5.

4.2 What is expected?

We briefly recall here the main properties expected from a continued fraction expansion. For a general discussion on the quest for suitable higher-dimensional continued fractions, see [31]. See also [67] for a discussion on their limitations.

The original motivation for multidimensional continued fractions (and in particular the Jacobi–Perron algorithm) arises from a question posed by Hermite to Jacobi in 1848: how to generalize continued fractions to higher dimensions in order to characterize algebraic irrationals as having a periodic expansion. This is still a very challenging problem in number theory. For more details, see, e.g., [20, 88, 138]. Using such a characterization, one could hope to determine fundamental units (e.g. of a cubic number field) and to solve Diophantine equations (as the ones described in Sect. 6). The formalism of multidimensional continued fractions based on the Klein polyhedra and sails developed in [9, 12, 95, 98] is well-suited for the detection of periodic expansions in terms of algebraic number fields by providing generalized Lagrange's theorem. For more on the subject and its history, see, e.g., [99, 100], the references in [86], and the book [87], which also includes a review of various generalizations of continued fractions. More generally, still from an arithmetic viewpoint, a multidimensional continued fraction algorithm is also expected to detect linear relations between $1, \alpha_1, \cdots, \alpha_d$.

We already discussed the fact that a continued fraction algorithm is expected to yield simultaneous better and better rational approximations with the same denominator for d-tuples $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d)$ in $[0,1]^d$, in an effective way and with a good approximation quality. More precisely, it has to produce a sequence of positive integers $(q^{(n)})_n$ such that the distance to the nearest integer $|||q^{(n)}\boldsymbol{\alpha}|||$ converges exponentially fast to 0 with respect to $q^{(n)}$, and ideally in $(q^{(n)})^{-\frac{1}{d}}$ (as predicted by Dirichlet's theorem). We have seen in Sect. 4.1 that the sequence of best approximations (see Definition 2) depends heavily on the chosen norm and that the associated transformations are no longer unimodular [102, 115]. However, it is possible to use the action of the diagonal flow on the space of unimodular lattices to better understand their behavior [42, 92, 106]. See also Sect. 5.

From a dynamical viewpoint, continued fraction algorithms are also expected to have reasonable ergodic properties such as the ones described in Sect. 4.4. For instance, we would like to have control over the (almost sure) behaviors concerning the growth of the convergents, the distribution of the partial quotients, or the speed of convergence via Lyapunov exponents. Famous examples of algorithms expressed dynamically as piecewise linear fractional transformations are the Jacobi–Perron, Brun or Selmer algorithms. Their description is the object of Sects. 4.3 and 4.5 below.

4.3 Dynamical Continued Fraction Algorithms

We recall here the main concepts related to dynamical continued fraction algorithms, expanding the brief description from Sect. 3.2. This dynamical formalism covers the most classical multidimensional continued fraction algorithms discussed in the classical references [31, 138, 146]; see also [105] for results concerning their Lyapunov exponents.

We consider here algorithms that produce the sequence of matrices $(A^{(n)})_{n\in\mathbb{N}}$ in a dynamical way. We take mostly a measure-theoretical viewpoint: the algorithms will be defined almost everywhere with respect to the Lebesgue measure on $[0,1]^d$. We focus on the unimodular case, since this allows a geometric interpretation in terms of bases of the integer lattice $\mathbb{Z}^{d+1}$.

Let $X \subset [0,1]^d$. (Usually X is simply $[0,1]^d$ but some algorithms can be also defined on sets of the form $\{x = (x_1, \cdots, x_d) \in [0,1]^d \mid 0 \le x_1 \le \cdots \le x_d \le 1\}$.) A d-dimensional unimodular *continued fraction map* over X is given by measurable maps

$$T\colon X \to X, \quad A\colon X \to GL(d+1,\mathbb{Z}), \quad \theta\colon X \to \mathbb{R}$$

such that for a.e. $\boldsymbol{\alpha} \in X$:

$$\begin{bmatrix} \boldsymbol{\alpha} \\ 1 \end{bmatrix} = \theta(\boldsymbol{\alpha}) A(\boldsymbol{\alpha}) \begin{bmatrix} T(\boldsymbol{\alpha}) \\ 1 \end{bmatrix}. \tag{9}$$

The maps A and T play the main role in the algorithm; the role played by the map θ is minor: it serves as a renormalization.

The associated *continued fraction algorithm* consists of iteratively applying the map T to a vector $\boldsymbol{\alpha} \in X$. This yields a sequence of matrices $(A(T^n(\boldsymbol{\alpha})))_{n\ge 1}$, called the *continued fraction expansion* of $\boldsymbol{\alpha}$.

One has

$$\begin{bmatrix} \boldsymbol{\alpha} \\ 1 \end{bmatrix} = \theta(\boldsymbol{\alpha})\theta(T(\boldsymbol{\alpha}))\ldots\theta(T^{n-1}(\boldsymbol{\alpha}))A(\boldsymbol{\alpha})A(T(\boldsymbol{\alpha}))\ldots A(T^{n-1}(\boldsymbol{\alpha})) \begin{bmatrix} T^n(\boldsymbol{\alpha}) \\ 1 \end{bmatrix}.$$

Such an algorithm is said to be "without memory." Indeed, the $(n+1)$-th step of the algorithm depends only on the map T and on the value $T^n(\boldsymbol{\alpha})$. This is in contrast with the algorithms based on lattice reduction that we will discuss in Sect. 5.

In most classical examples of such algorithms the continued fraction map T is piecewise continuous, or even a piecewise homography (see Sect. 4.5). Further, many of them are *positive*, that is, all the matrices appearing as the images of the map A are non-negative.

We illustrate this formalism with the regular continued fraction case. Here θ is the identity. Indeed considering

$$\begin{aligned} \begin{bmatrix} \alpha \\ 1 \end{bmatrix} &= \alpha T_G(\alpha)\cdots T_G^{n-1}(\alpha) \begin{bmatrix} 0 & 1 \\ 1 & a_1 \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & a_n \end{bmatrix} \begin{bmatrix} T_G^n(\alpha) \\ 1 \end{bmatrix} \\ &= \alpha T_G(\alpha)\cdots T_G^{n-1}(\alpha) \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix} \begin{bmatrix} T_G^n(\alpha) \\ 1 \end{bmatrix}, \end{aligned}$$

one gets

$$\alpha\, T_G(\alpha) \cdots T_G^{n-1}(\alpha) = |q_{n-1}\alpha - p_{n-1}|,$$

where p_n/q_n stands for the n-th convergent of α, and T_G stands for the Gauss map. We see with this example that the map θ, which allows the renormalization with respect to the last coordinate (set to 1), is of an arithmetic nature, in the sense that the multiplicative cocycle associated with θ, i.e., $\theta(T_G^{n-1}(\alpha)) \cdots \theta(T_G(\alpha))\theta(\alpha)$, yields the arithmetic quantity $|q_{n-1}\alpha - p_{n-1}|$.

We now revisit this formalism more geometrically in terms of lattice bases following [31]. One approximates the vectorial line in $\mathbb{R}^{d+1}$ directed by the non-zero vector $(\boldsymbol{\alpha}, 1)$ by a sequence of integer lattice bases $(\mathbf{b}^{(n)})_{n\in\mathbb{N}}$ of $\mathbb{Z}^{d+1}$, namely the $d+1$ columns of the matrices $A(\boldsymbol{\alpha}) \cdots A(T^{n-1}(\boldsymbol{\alpha}))$. The lattice bases generate cones that are expected to converge toward the line directed by $(\boldsymbol{\alpha}, 1)$. Moreover, if the algorithm is positive, then these cones are nested and $(\boldsymbol{\alpha}, 1)$ belongs to the positive cone generated by the vectors $\mathbf{b}_i^{(n)}$, $i = 1, \ldots, d+1$, i.e.,

$$(\boldsymbol{\alpha}, 1) \in \left\{ \sum_{1\le i\le d+1} \lambda_i \mathbf{b}_i^{(n)} \mid \lambda_i \ge 0 \ \text{ for all } i = 1, \cdots, d+1 \right\}.$$

In fact, (9) implies that the coefficients $(\lambda_i)_{1\le i\le d+1}$ are proportional to the vector $(T^n(\boldsymbol{\alpha}), 1)$.

The algorithm thus produces a sequence of bases of the lattice $\mathbb{Z}^{d+1}$ that all determine a homogeneous cone in $\mathbb{R}^{d+1}$ that contains the ray $\{\lambda(\boldsymbol{\alpha}, 1) \mid \lambda \ge 0\}$. This is the viewpoint developed for instance in [31]; in fact, the algorithms there are designed in this way.

We recall that the continued fraction map T acts on parameters $\boldsymbol{\alpha}$ living in an ambient d-dimensional space. In view of what precedes, it is also natural to start directly with a general line $\boldsymbol{\ell} = (\ell_1, \cdots, \ell_{d+1})$ in $\mathbb{R}^{d+1}$ and to have an algorithm with such a line as an input. But then, there is no canonical way to "projectivize" the algorithm, i.e., to go from the line $\boldsymbol{\ell}$ to a d-dimensional vector $\boldsymbol{\alpha}$ working in a compact set X on which a dynamical system T acts. One can set, e.g., $\alpha_i = \ell_i/\ell_{d+1}$ for $i = 1, \cdots, d$ (e.g., if the line belongs to the positive cone of $\mathbb{R}^{d+1}$ and if ℓ_{d+1} is the largest entry). Usual ways to go from some $\boldsymbol{\ell} \in \mathbb{R}_+^{d+1}$ to some $\boldsymbol{\alpha} \in [0,1]^d$ consists in setting $\ell_{d+1} = 1$, and working with entries $\ell_i \in [0,1]^d$ for $1 \le i \le d$, or else, in working with the simplex $\sum_{i=1}^{d+1} \ell_i = 1$, with $\ell_i \ge 0$ for all i. See for instance [138] for more details. One can also choose to work directly on the projective space $\mathbb{P}(\mathbb{R}^d)$ by associating with each element $[y_1 : y_2 : \cdots : y_{d-1} : y_d]$ the representative defined by $\max y_i = 1$ and by working with projectivizations of matrices in $GL(d, \mathbb{Z})$. This possibility of having different choices for a same piecewise $d+1$-dimensional linear map explains the abundance of existing algorithms (as illustrated in Sect. 4.5).

Let us discuss now the possible steps of the algorithms, i.e., the operations that can be performed on the bases together with the choices allowed for partial quotient matrices $A^{(n)}$. An algorithm is said to be *additive* if all the matrices belong to a finite

set, i.e., the map A from (2) takes finitely many values. The terminologies additive vs. multiplicative or division vs. subtractive algorithm are also commonly used; see, e.g., [31] where an algorithm is said to be subtractive if $c^{(n)} = 1$ in (10), and additive if $c^{(n)}$ is chosen as the maximal possible number allowing the line to stay within the positive cone generated by the convergent vectors $\mathbf{b}_i^{(n)}$ for $i = 1, \ldots, d$.

As an illustration, the additive version of the Gauss map is given by the Farey map

$$x \mapsto \begin{cases} \dfrac{x}{1-x} & \text{for } 0 \le x \le 1/2, \\ \dfrac{1-x}{x} & \text{for } 1/2 \le x \le 1. \end{cases}$$

Dynamically, this creates non-trivial changes; indeed, the Gauss map is known to have a finite ergodic invariant measure (see (11) below), which is not the case for the Farey map. More generally, additive and multiplicative versions for a same type of rule can lead to very distinct behaviors. See for instance [31], where it is shown that the multiplicative form of Selmer's algorithm is not an acceleration of its additive version. The underlying cause of this is the fact that the group of matrices generated by positive transvections and permutations is not commutative.

One convenient way to get an additive algorithm is to restrict the range of the map A to the set of elementary and permutation matrices with entries in $\{0, 1\}$. A matrix is called *elementary* if it has 1's on the diagonal, one entry equal to 1 elsewhere, and all other entries equal to 0. In geometric terms (still following the geometric formalism from [31]) the allowed operations on the bases at each step n are of elementary types (they correspond to integer transvections): for every n, there exist $i \neq j$ (with i, j depending on n) and $c^{(n)} \in \mathbb{N}$ such that

$$\mathbf{b}_i^{(n+1)} = \mathbf{b}_i^{(n)} + c^{(n)}\mathbf{b}_j^{(n)}, \ \mathbf{b}_k^{(n+1)} = \mathbf{b}_k^{(n)} \text{ for } k \neq i. \tag{10}$$

This restriction is not a severe one and most of the algorithms discussed in the present survey enter this framework, by allowing also permutation rules between the vectors. Algorithms for which the choice of the coefficients i, j and $c^{(n)}$ depend only on the cofactors of $\boldsymbol{\ell}$ with respect to $\mathbf{b}^{(n)}$, i.e., the integers $a_i^{(n)}$ such that

$$\boldsymbol{\ell} = a_1^{(n)}\mathbf{b}_1^{(n)} + \cdots + a_{d+1}^{(n)}\mathbf{b}_{d+1}^{(n)},$$

are called *vectorial* in [31]. They are memory-less algorithms. In particular, with the notation of (2), if, for some $\boldsymbol{\alpha}$ the matrix $A(\boldsymbol{\alpha})$ is equal to an elementary matrix, then the entries of the vector $(T(\boldsymbol{\alpha}), 1)$ are obtained (modulo the renormalization produced by the map θ) by subtracting an entry from another one; it is obtained by performing subtractions.

4.4 The Effectiveness of Ergodic Theory

Having an underlying dynamical system offers a wide range of mathematical tools, including ergodic theory and thermodynamic formalism via transfer operators. Ergodic theorems describe the limiting behavior of ergodic sums of the form $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k$. In probabilistic terms, the random variables $f \circ T^n$ satisfy the strong law of large numbers under the ergodic hypothesis on T. Ergodic sums $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k$ allow the expression of a wide set of algorithmic and arithmetic parameters, and ergodic theorems allow the understanding of their mean behaviors. Besides ergodic theory, transfer operators provide a description of the evolution of probability density functions, by transporting the action of the map T from a dynamical system to the densities. If initial conditions (x_i) for trajectories are distributed according to a probability density function, then the new collection of points $(T(x_i))$ is distributed according to a new probability density function, obtained by applying a transfer operator.

Ergodic theory has been quite an effective tool in understanding the typical behavior of various expansions of numbers and their quality of approximation. To name a few, we mention β-expansions, Lüroth series and multidimensional continued fractions.

We first illustrate it with the Gauss map. For general ergodic aspects of the Gauss map, see, e.g., [24, 51, 57]. The statistics of occurrences of partial quotients in continued fraction expansions are deduced from the ergodic theorem applied to the Gauss map, and having an ergodic absolutely continuous invariant measure (a so-called *a.c.i.m.*) then provides metric results that hold almost everywhere with respect to the Lebesgue measure.

To explain this in more detail, let us recall basic ergodic properties of the Gauss map. See also Sect. 3.3 where first definitions have been given for shift spaces. We endow the dynamical system $([0, 1], T_G)$ with a structure of a measure-theoretic dynamical system. A *measure-theoretic dynamical system* is defined as a system $(X, T, \mu, \mathcal{B})$, where μ is a probability measure defined on the σ-algebra $\mathcal{B}$ of subsets of X, and $T : X \to X$ is a measurable map which preserves the measure μ, i.e. $\mu(T^{-1}(B)) = \mu(B)$ for all $B \in \mathcal{B}$. In this case one also says that the measure μ is *T-invariant*. Here, we endow $([0, 1], T_G)$ with the Gauss measure μ_G which is a Borel probability measure absolutely continuous with respect to the Lebesgue measure defined via the following density function

$$\mu_G = \frac{1}{\log 2}\int \frac{1}{1+\alpha} d\alpha. \tag{11}$$

One checks that this measure is T_G-invariant. The Gauss map is *ergodic* with respect to the Gauss measure, that is, every Borel subset B of $[0, 1]$ such that $T_G^{-1}(B) = B$ has either zero or full measure. This implies that almost all orbits are dense in $[0, 1]$.

Ergodicity yields furthermore the following striking convergence result. Indeed, measure-theoretic ergodic dynamical systems satisfy *Birkhoff's ergodic theorem*,

also called *individual ergodic theorem*, which relates spatial means to temporal means.

Theorem 4 (Birkhoff's Ergodic Theorem) *Let $(X, T, \mu, \mathcal{B})$ be an ergodic measure-theoretic dynamical system. Let $f \in L^1(X, \mathbb{R})$. Then the sequence $(\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k)_{n\geq 0}$ converges a.e. to $\int_X f\, d\mu$:*

$$\forall f \in L^1(X, \mathbb{R}) , \quad \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \xrightarrow[n\to\infty]{\mu-a.e.} \int_X f\, d\mu .$$

For example in the case of continued fractions, using this theorem one is able to describe the distribution of the digits.

Theorem 5 *For a.e. $\alpha \in [0, 1]$ the digit j occurs in the continued fraction expansion of α with density* $\dfrac{1}{\log 2}(2\log(1 + j) - \log j - \log(2 + j))$.

One can even show that the digits have the mixing property; one can as well give a short and elegant proof of Lévy's theorem which states that

$$\lim_{n\to\infty} \frac{1}{n} \log q_n(\alpha) = \frac{\pi^2}{12 \log 2} \tag{12}$$

for Lebesgue almost every point α. With the help of a dynamical construction called the *natural extension* (which is basically a way to make the dynamics invertible), one can describe the asymptotic behavior of the approximation coefficients

$$\Theta_n(\alpha) = q_n^2 |\alpha - \frac{p_n}{q_n}|.$$

As an illustration, we mention the ergodic proof of Bosma, Jager and Wiedijk [26] of the famous Doeblin-Lenstra conjecture which says that for almost all α the limit

$$\lim_{n\to\infty} \frac{1}{n} \#\{1 \leq j \leq n \mid \Theta_j(\alpha) \leq z\}, \quad \text{where } 0 \leq z \leq 1,$$

exists and equals the distribution function $F(z)$ given by

$$F(z) = \begin{cases} \dfrac{z}{\log 2} & 0 \leq z \leq \frac{1}{2}, \\ \dfrac{1}{\log 2}(1 - z + \log 2z) & \frac{1}{2} \leq z \leq 1. \end{cases} \tag{13}$$

In [77], Jager used ergodic tools to describe the simultaneous distribution of two consecutive Θ's by determining for almost all α the exact value of

$$\lim_{n\to\infty} \frac{1}{n} \#\{1 \le j \le n \,:\, \Theta_{j-1}(\alpha) \le z_1,\ \Theta_j(\alpha) \le z_2\}, \quad \text{where } 0 \le z_1, z_2 \le 1.$$

As a third example we mention that it can be proven with the help of ergodic theory that for each real irrational number α and each integer $n \ge 1$ one has

$$\min(\Theta_{n-1}, \Theta_n, \Theta_{n+1}) \ < \ \frac{1}{\sqrt{a_{n+1}^2 + 4}} \tag{14}$$

and

$$\max(\Theta_{n-1}, \Theta_n, \Theta_{n+1}) \ > \ \frac{1}{\sqrt{a_{n+1}^2 + 4}}\,. \tag{15}$$

For a generalization of these results in the setting of best approximations, see [42, 43], and also Sect. 5. Inequality (14) is a generalization of a result by Borel [30], which states that

$$\min(\Theta_{n-1}, \Theta_n, \Theta_{n+1}) \ < \ \frac{1}{\sqrt{5}}\,.$$

A great number of people independently found (14), see for example [28, 122, 143]. Inequality (15) is due to Tong [149]. In fact, ergodic theoretic methods yield easy proofs of generalizations of a great number of classical results by Fujiwara, Segre and others like LeVeque, Szüzs, and Segre; see [51, 78] for more results and details.

In a different direction, ergodic theoretic methods can also be used to prove (generalizations of) Lochs' theorem which compares the amount of information given by the digits of different types of number expansions. The original statement of Lochs [109] compared the regular continued fraction digits to decimal digits. More precisely, for an $\alpha \in (0, 1)$, let $(a_k)_{k\ge1}$ denote its regular continued fraction digits and $(d_k)_{k\ge1}$ its decimal digits and, for each $n \ge 1$, let $m_n(\alpha)$ denote the largest number of digits a_k that can be determined from knowing $d_1, \ldots, d_n$. Then Lochs proved the following statement.

Theorem 6 *For a.e.* $\alpha \in (0, 1)$, $\lim_{n\to\infty} \frac{m_n(\alpha)}{n} = \frac{6 \log 2 \log 10}{\pi^2}$.

Lochs' theorem was placed in a dynamical setting in [18] and developed further in [48]. It became apparent that the number $\frac{6 \log 2 \log 10}{\pi^2}$ is related to the measure-theoretic entropies $h(T_G) = \frac{\pi^2}{6 \log 2}$ and $h(T_{10}) = \log 10$ of the transformations T_G and $T_{10}(x) = 10x \pmod 1$ that generate regular continued fractions and decimal expansions, respectively. See also [23, 110] for similar results for other types of expansions. Later Lochs' theorem was extended and generalized in several directions, e.g., in [97] for random dynamical systems and in [17] to include systems with zero entropy.

Let us consider now the case of multidimensional continued fraction algorithms. For a general description of their ergodic properties, see [138]. A first step in the ergodic study of a continued fraction algorithm consists of proving the existence of an a.c.i.m. (absolutely continuous invariant measure), playing a similar role as the Gauss measure defined in (11). Note that a very efficient way to find a.c.i.m.'s for continued fraction transformations is via the natural extensions; see, e.g., [6, 7, 13, 55, 93].

Let us take a closer look at the convergence properties from a dynamical viewpoint. The a.e. exponential (strong) convergence of the Brun [60, 113, 139] and Jacobi–Perron algorithm [14] (see also [105, 137]) holds in dimension $d = 2$: there exists $\delta > 0$ s.t. for a.e. (α_1, α_2), there exists $n_0 = n_0(\alpha, \beta)$ s.t. for all $n \geq n_0$

$$|\alpha_1 - p_1^{(n)}/q^{(n)}| < \frac{1}{(q^{(n)})^{1+\delta}}, \ |\alpha_2 - p_2^{(n)}/q^{(n)}| < \frac{1}{(q^{(n)})^{1+\delta}},$$

where $p_1^{(n)}, p_2^{(n)}, q^{(n)}$ are given by the Brun (resp. Jacobi–Perron) algorithm.

We have seen already in Sect. 3.3 that the quality of rational approximations provided by continued fraction algorithms obtained dynamically by an iteration of a map can be expressed in terms of the two first Lyapunov exponents λ_2 and λ_1: the coefficient δ corresponds to $-\lambda_2/\lambda_1$. There is, in fact, a strong link with the uniform approximation exponent $\eta_A^*(\boldsymbol{\alpha})$ (whose definition is recalled below) as shown in [105]; see also [70, Theorem 1], [15, Proposition 4] and [38, Section 2]. One considers as in (2) the maps $T : X \to X$, $A : X \to \mathrm{GL}(d+1, \mathbb{Z})$. For fixed $\boldsymbol{\alpha} \in X$ and $i \in \{1, \ldots, d+1\}$, we have

$$\eta_A^*(\boldsymbol{\alpha}, i) =$$

$$\sup \left\{ \delta > 0 \ : \ \exists\, n_0 = n_0(\boldsymbol{\alpha}, i, \delta) \in \mathbb{N} \text{ s. t. } \forall\, n \geq n_0, \ \left\| \boldsymbol{\alpha} - \frac{\mathbf{p}_i^{(n)}}{q_i^{(n)}} \right\| < (q_i^{(n)})^{-\delta} \right\},$$

where $\| \cdot \|$ is an arbitrary norm in $\mathbb{R}^d$. The quantity

$$\eta_A^*(\boldsymbol{\alpha}) = \min_{1 \leq i \leq d+1} \eta_A^*(\boldsymbol{\alpha}, i)$$

is called the *uniform approximation exponent* for $\boldsymbol{\alpha}$ using the algorithm A. Now, following [105, Theorem 4.1], we consider a d-dimensional multidimensional continued fraction algorithm A satisfying some mild ergodic conditions (called (H1) to (H5) in [105]). Let η_A^* be the uniform approximation exponent of A. We have $\lambda_1(A) > \lambda_2(A)$ and

$$\eta_A^*(\boldsymbol{\alpha}) = 1 - \frac{\lambda_2(A)}{\lambda_1(A)}$$

holds for almost all $\boldsymbol{\alpha} \in X$. In particular, if $\lambda_2(A) < 0$ then A is a.e. strongly convergent. Lyapunov exponents of classical algorithms such as the Jacobi–Perron algorithm or the Brun algorithm have been thoroughly studied in [14, 32]; see also [63] for results on the simplicity of the Lyapunov spectrum.

The computation of Lyapunov exponents is a challenging problem. For numerical results, see [38]. In practice, ergodic theorems provide efficient ways of estimating Lyapunov exponents numerically by following trajectories and then taking averages over truncated trajectories. This has been developed, e.g., in [29, 72] for continued fractions or in [157] for interval exchanges. Transfer operators are efficient ways to reach them, such as developed, e.g., in [81, 128], by getting in some cases exact computations for the top Lyapunov exponents of random products of matrices using transfer operators. Indeed, transfer operators come with the analogue of Perron–Frobenius theory for non-negative matrices. They are well suited to provide approximations by working in finite dimension, for instance by truncating Taylor expansions of analytic functions [49, 107]. See also [144] for lower and upper bounds for the Lyapunov exponents for random products of square matrices with determinant 1 having real and non-negative entries.

4.5 Classical Examples

With this section we want to illustrate the variety of continued fractions algorithms defined according to the formalism of Sect. 4.4, as well as to focus on two classical algorithms, namely the Brun and the Jacobi–Perron algorithms.

Consider algorithms based on elementary matrices, i.e., on subtractions, such as described in Sect. 4.4. In order to stress the simple rules that govern them, we express them in dimension $d + 1 = 3$. We thus start with parameters $\boldsymbol{\ell} = (\ell_1, \ell_2, \ell_3) \in \mathbb{R}_+^3$. We have to decide which number has to be subtracted, and with respect to which number it has to be done. Usually numbers ℓ_1, ℓ_2, ℓ_3 are sorted in increasing (or decreasing) order. We stress the subtraction rule but it is usually preceded and followed by a sorting operation.

For Jacobi–Perron, we subtract the second entry from the other ones, with the entries being ordered in such a way that the first entry is the largest one (but with no further order restriction on the two other entries). For Brun, we subtract the second largest from the largest one. For Poincaré, we subtract the second largest entry from the largest one, the third largest from the second largest, etc. For Selmer, we subtract the smallest (positive) entry from the largest one. For the fully subtractive algorithm, we subtract the smallest (positive entry) from all the larger ones.

We also recall that given an algorithm described at the level of a $d + 1$-dimensional space, there exist several possible projectivizations. This is the case of the various forms taken by the Brun algorithm. In particular, the Brun algorithm is also called modified Jacobi–Perron algorithm: the modified Jacobi–Perron algorithm introduced by Podsypanin in [127] is a two-point extension of the Brun algorithm.

Consider now the Jacobi–Perron algorithm (that will be handled in more details with respect to the nearest integer part in Sect. 7). The linear form of the Jacobi–Perron algorithm is defined on $\{(y_0, y_1, y_2) \in \mathbb{R}^3 \setminus \{\mathbf{0}\} : \ y_0 \geq y_1, y_2 \geq 0\}$ by

$$(y_0, y_1, y_2) \mapsto (y_1, y_2 - \lfloor y_2/y_1 \rfloor y_1, y_0 - \lfloor y_0/y_1 \rfloor y_1).$$

If we set

$$x_1 := y_1/y_0, \ x_2 := y_2/y_0,$$

we recover its projective version defined on $[0, 1]^2$ as

$$(x_1, x_2) \mapsto (\{x_2/x_1\}, \{1/x_1\}) .$$

Let us stress some differences between Brun's and Jacobi–Perron's rule. The Brun algorithm is a space-ordering algorithm according to the terminology introduced in [71]. (Note that it is called ordered Jacobi–Perron in [70].) Furthermore, each step of the Brun algorithm produces only one digit. This helps in computing the natural extension and the invariant measure of the Brun algorithm (see, e.g., [7]). Contrary to the Brun algorithm, the role played by y_1 and y_2 (in the description above) is not determined by a comparison between both parameters in the Jacobi–Perron case; this might explain the fact that an explicit expression of the natural extension of this algorithm is still not known. Nevertheless, the framework of S-expansions and the so-called techniques of Insertion and Singularization (see [75]) allow one to relate both algorithms as shown in [140]; see also [141]. Both algorithms (Brun and Jacobi–Perron) are known to have an invariant ergodic probability measure equivalent to the Lebesgue measure (an a.c.i.m.); see for instance [135] and [138]. However, this measure is not known explicitly for Jacobi–Perron (the density of the measure is shown to be a piecewise analytic function in [32]), whereas it is known explicitly for Brun [7, 60]. Note that the Brun algorithm can be considered as an additive algorithm [27], where it is proved that partial quotients tend to be equal to 1. Lastly, let us quote [119] for Borel-Bernstein type theorems on the growth of partial quotients.

5 Lattice Reduction and Rational Approximations

We now focus on the second approach discussed in Sect. 3.2 based on lattice reduction. Lattice reduction methods induce indeed a particularly fruitful way of exhibiting good simultaneous approximations, or else small values for linear forms. Algorithms based on lattice reduction theory are based on the following idea: lattice reduction algorithms do not produce a priori the smallest vector of a lattice, but a reasonably small vector. That is, a vector that is small enough for guarantying Diophantine approximation properties that can be compared with Dirichlet's quality

up to an approximation factor exponential in the dimension. We can thus consider these algorithms as providing effective versions of Dirichlet's theorem, yielding a satisfying compromise between efficient computation and sharpness of the obtained bounds, that is, between algorithmic issues and Diophantine quality.

The range of applications of lattice reduction is quite wide for the following reasons. Firstly, lattice reduction algorithms play a central algorithmic role in cryptology, computer algebra, integer linear programming and algorithmic number theory. Secondly, they are particularly versatile in terms both of existing variants and algebraic contexts where they can be developed, see, e.g., [40, 62, 117]. Thirdly, they are efficient: LLL has a polynomial runtime with respect to the dimension. In particular, in the present context, they produce efficient gcd algorithms (see, e.g., [73]), and there exist promising attempts in order to devise continued fractions upon them (see [21, 37, 106]). However, there remains much to understand concerning their executions and the geometry of the outputs [120, 155].

Lattice reduction is based on the following elementary basis transformations: they can be described in terms of size reduction (the vector $\mathbf{b}_i$ of the basis $(\mathbf{b}_1, \ldots, \mathbf{b}_{d+1})$ is replaced by $\mathbf{b}_i - \lambda \mathbf{b}_j$ with $1 \leq j < i$), and of exchange steps, also called swaps (one exchanges $\mathbf{b}_i$ and $\mathbf{b}_{i+1}$). These operations are decided with respect to the Gram–Schmidt orthogonalization of the basis $\mathbf{b}$.

More precisely, let $(\mathbf{b}_i^*)$ stand for the basis obtained via the Gram–Schmidt orthogonalization from the basis $(\mathbf{b}_1, \ldots, \mathbf{b}_{d+1})$, i.e., $\mathbf{b}_i^*$ is the orthogonal projection of $\mathbf{b}_i$ on the orthogonal of the space generated by $\mathbf{b}_i, \cdots, \mathbf{b}_{i-1}$. One writes $\mathbf{b}_i^* = \mathbf{b}_i - \sum_{k=1}^{i-1} \mu_{ik} \mathbf{b}_k^*$ with $\mu_{ik} = \frac{\langle \mathbf{b_i}, \mathbf{b}_k^* \rangle}{\langle \mathbf{b}_\mathbf{k}^*, \mathbf{b}_k^* \rangle}$ for $k \leq i - 1$. A basis $(\mathbf{b}_1, \cdots, \mathbf{b}_d)$ is said to be LLL-reduced if

- $|\mu_{ik}| \leq 1/2$ for all i, k with $1 \leq i \leq d + 1$ and $k \leq i - 1$ (the basis is said *proper*).
- $3/4 \|\mathbf{b}_i^*\| \leq \|\mu_{i+1,i} \mathbf{b}_i^* + \mathbf{b}_{i+1}^*\|$ for all i (this condition is called *Lovasz' condition*).

The factor $3/4$ can be replaced by a parameter t with $3/4 < t < 1$. The LLL algorithm consists of two steps.

- First, make the basis proper by replacing $\mathbf{b}_i$ by $\mathbf{b}_i - \|\mu_{ij}\| \mathbf{b}_j$, for $j < i$, where $\|\mu_{ij}\|$ stands for the distance to the nearest integer.
- If for some i, Lovasz' condition is not satisfied, then swap $\mathbf{b}_i$ and $\mathbf{b}_{i+1}$ and go the previous step.

Recall that we have sketched the basic strategy underlying the use of lattice reduction in this framework in Sect. 3.2. One starts with the lattice Λ_t generated by the columns of the matrix

$$M_t := \begin{bmatrix} 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_d \\ 0 & 0 & \cdots & 0 & t \end{bmatrix}.$$

Note that $\det(M_t) = t$, hence, the lattice Λ_t changes at each step of the algorithm. One lets t tend to 0. Let us stress the fact that this strategy differs from the one discussed in Sect. 4.5 where one worked with bases of the fixed lattice $\mathbb{Z}^{d+1}$.

Lattice reduction algorithms such as LLL then perform a succession of permutations and subtractions on the matrix M_t, i.e., they multiply the matrix M_t by elementary matrices and permutation matrices. This is a common feature between unimodular continued fraction algorithms and algorithms based on lattice reduction, namely that they are made of a succession of permutations and subtractions. The decisions are taken for classical unimodular continued fractions by comparing entries, whereas lattice reduction involves quadratic comparisons in the sense that they depend on the Gram–Schmidt orthogonalization.

Lattice reduction provides best approximations of a real number, see [121, p.226,267], and for a survey on the overall strategy for getting constructive type results in Diophantine approximation based on LLL, see [121, p. 222]. Nevertheless, note that even in dimension 2, when using the Gauss algorithm whose efficiency has been largely proved, one has "little control on the convergent which is returned; in particular, this is *not* the largest convergent with denominator less than $2\sqrt{C/3}$", as quoted in [121, p.226 Example 1]; here the bound $2\sqrt{C/3}$ comes from Theorem 7 of [121, Chapter 6].

As an illustration of the efficiency of lattice reduction in algorithmic number theory, see, e.g., [129] where algorithms for solving the dual problems of approximating linear forms and of simultaneous approximation in number fields are developed by applying the LLL algorithm. A system of independent units of full rank is generated in any algebraic number field, which also leads to an algorithmic generalization of Lagrange's theorem.

Several attempts already exist in order to use lattice reduction to get simultaneous approximations. Let us quote [58, 59, 82–84, 103, 104] for strongly convergent algorithms; however they do not present the same advantages as more classical memory-less algorithms. Let us quote also [104] and [106] based on Minkowski lattice reduction; this approach is not effective, but it produces best approximations. This study is also extended in [66], [46] and [92].

Some approaches are built specifically upon LLL; see [21, 37]. In [37] an iterated LLL algorithm is designed that is obtained by decreasing the parameter t from the lattice Λ_t, by dividing it by a given constant (including also experimental data). In [21], the conditions that occur in LLL are proved to be linear in the parameter t (tending to 0). The idea at step k is to consider the smallest parameter t_k for which Λ_{t_k} is reduced, and then perform a reduction with $t_k - \varepsilon$.

An efficient way to input some dynamics with this reduction viewpoint is to rely on homogeneous dynamics with the (left) action of the diagonal flow (g_t) defined by

$$\begin{bmatrix} e^t I_d & 0 \\ 0 & e^{-dt} \end{bmatrix}$$

on the space of unimodular lattices, i.e., on the homogeneous space $SL(d+1,\mathbb{R})/SL(d+1,\mathbb{Z})$. This is a very fruitful way to combine lattice reduction with the strength of dynamical methods such as highlighted in the survey [46]. This amounts to changing the parameter t. See in particular [92] for a strongly convergent multidimensional continued fraction algorithm with effective hyperbolicity estimates.

This is particularly relevant for continued fractions defined in terms of best approximations. In [42, 43], Levy's result about the almost sure growth rate of the denominators of the convergents from (12) is extended to the best Diophantine approximations (see Definition 2). The value of the limit is given by an integral over a codimension one submanifold in the space of lattices $SL(d+1,\mathbb{R})/SL(d+1,\mathbb{Z})$. An analogue of the Doeblin-Lenstra discussed in Sect. 4.4 is also given in [43]. As highlighted in [43], the section is provided by lattices whose first two minima of lattices are equal and the first return map of the geodesic flow in the transversal plays the role of an invertible extension of the missing Gauss map. The main idea is to relate shortest and so-called minimal vectors of the lattice Λ_t with best approximations (see [46, Lemma 8]) together with a suitable choice of a norm (see [46, Section 3.2] and [45]).

6 Some Applications of Continued Fractions

In this section we give a diverse range of applications of continued fractions in arithmetics, cryptography and symbolic dynamics. We then propose possible hints for improvements of dynamical unimodular continued fraction algorithms.

We start with arithmetical applications. Continued fractions play an important role in the **arithmetic of algebraic curves**. The relation between the geometry of the plane quartic model of an elliptic curve, given by the equation $y^2 = x^4 - 6ax^2 - 8bx + c$, and the continued fraction of y with respect to the x^{-1}-adic valuation is given in [8] and [1]. The given elliptic curve has two rational points at infinity, call them P and O, such that O is the origin of the Mordell-Weil group of the elliptic curve. Then the order of the image of P in the Jacobian (called the infinity divisor of the curve) is finite if and only if the continued fraction of y is periodic. Moreover, the order depends on the period of the continued fraction. This result has been generalized for the hyperelliptic curves in [126]. The torsion order of the infinity divisor is given as the sum of the degrees of all partial fractions from 0 to the period of the continued fraction of y in [8, 130, 150]. Furthermore, thanks to the periodic continued fraction of y, for a given even degree hyperelliptic curve

$y^2 = f(x)$, the Pell equation for f is proved to have a non-trivial solution, i.e., there exist p, q in $k[x]$, p not a constant, such that $p^2 - q^2 f = 1$, in [8, 126].

Cryptography is another field where continued fractions occur in various places. First of all, continued fractions can be used in cryptanalysis, e.g., to attack the RSA cryptosystem. This cryptosystem is based on the mathematical problem of factoring a natural number that is a product of two large prime numbers. It is an *asymmetric cryptosystem*, i.e., a publicly known key is used for encryption and a secret key is used for decryption. If the secret key is chosen too small, one can use continued fraction expansions to efficiently compute the private key and break the cryptosystem. This attack is known as Wiener's attack (see [152]).

Continued fractions can be used to directly factorize integers. The continued fraction factorization method (CFRAC) and Shanks's square forms factorization (SQUFOF) are two such algorithms. For example, the main idea behind the CFRAC algorithm is to compute the continued fraction expansion of $\sqrt{n}$, where n is the number we want to factorize. This continued fraction expansion is then used to obtain a factor dividing n. However, these factorization algorithms are not efficient enough to factorize numbers in the magnitude of RSA parameters.

Apart from cryptanalysis, there is a connection between stream ciphers and continued fraction expansions. Stream ciphers are *symmetric cryptosystems*, i.e., the same key is used for encryption and decryption. A stream cipher generates a pseudorandom bit string. To encrypt a message, each bit of the pseudorandom bit string is combined with a bit of the secret message (e.g., with the XOR operation). Niederreiter uses continued fraction expansions of generating functions to analyze the randomness of pseudorandom sequences (see [118]). In [85], Kane constructs a stream cipher using continued fractions. The cryptosystem KronCrypt is a block cipher that is based on a constructive proof of Kronecker's approximation theorem (see [56]). Block ciphers encrypt blocks of the secret message, instead of generating a bit string that is combined with the secret message bitwise. KronCrypt has the structure of a Feistel cipher with key-dependent S-boxes. The continued fraction expansion is used for the key generation.

A rather indirect connection to cryptography is the shared interest in lattice reduction algorithms in the setting of post-quantum cryptography. As described in Sect. 5, these algorithms compute relatively short vectors of lattices. Various cryptosystems such as Frodo, Dilithium, and Kyber are based on the problem of finding short vectors in lattices. Hence, the study of lattice reduction algorithms such as the LLL algorithm is important for the security analysis of lattice-based cryptosystems.

We also note that due to its simplicity, the Brun algorithm appears in various application fields in cryptography and beyond. See [54], where efficient exponentiation using addition chains is used. Note that continued fractions were already used for addition chains; see, e.g., [16]. See also the survey [154] for application of lattice reduction and of the Brun algorithm in wireless communications and statistical signal processing.

We now consider some applications in **(symbolic) dynamics** with the seminal example of Sturmian dynamical systems, introduced by Morse and Hedlund in [114]. There is an impressive literature devoted to their study and to their possible generalizations in word combinatorics, and in digital geometry [133]. This is due to several factors. They provide symbolic codings for the simplest arithmetic systems, namely the irrational translations on the circle, they also code discrete lines and are unidimensional models of quasicrystals [22]. Moreover, the scale invariance of Sturmian dynamical systems allows their description using a renormalization scheme governed by usual continued fractions via the geodesic flow acting on the modular surface. More generally, renormalization schemes can often be interpreted as continued fractions [156]; a striking illustration comes from the study of interval exchanges in relation with the Teichmüller flow, through the work, among others, of Veech, Masur, Yoccoz, and Avila.

Similarly as for continued fractions, there is no canonical generalization of Sturmian words. Episturmian words, also called Arnoux-Rauzy words, have attracted a lot of attention in this direction. See in particular [4, 5, 44] for the study of associated continued fractions on the Rauzy gasket. More generally, it has been a long-standing problem to find good symbolic codings for translations on the d-dimensional torus that enjoy the beautiful properties of Sturmian sequences. This is the object of [39] which defines symbolic codings of toral translations based on exponentially convergent multidimensional continued fraction algorithms, leading to renormalization schemes. However these results rely on the positiveness of the second Lyapunov exponent, and they thus hold only in small dimensions.

We end this section by discussing possible ways to improve existing continued fractions algorithms. Indeed, in view of the non-positivity of the second Lyapunov exponent, there is a need to design strongly convergent continued fractions algorithms in high dimensions. In fact, positivity of the second Lyapunov exponent has only been proved in dimension $d = 2$ for classical algorithms, and $d = 3$ for the Brun algorithm. One natural approach for the design of continued fractions consists in trying to derive continued fraction algorithms from lattice reduction algorithms using the fact that they compute short vectors and that they reach Dirichlet's bound (up to a constant depending exponentially on the dimension). The design of continued fractions will go along with the dynamical modeling of lattice reduction algorithms (and more specifically LLL) for their probabilistic analysis. There is also a need to improve existing algorithms of a dynamical nature by taking advantage of their dynamical properties.

For other types of number expansions, in particular for β-expansions, it turned out to be very fruitful to introduce a random dynamical system to produce the expansions. Where a deterministic system is defined as a pair (X, T), a *random dynamical system* makes use of a family of transformations $(T_i : X \to X)_{i\in I}$ (for some index set I) all defined on the same domain X, where each map is chosen with a certain probability. One then studies the compositions of the form

$$T_{i_n} \circ \cdots \circ T_{i_1}(x), \quad i_j \in I,$$

instead of $T^n(x)$. For β-expansions, i.e., expansions of real numbers x of the form $\sum_{k\geq 1} \frac{b_k}{\beta^k}$, with $\beta > 1$ a non-integer and each b_i an integer between 0 and β, a corresponding random dynamical system was first introduced in [52]. Applications of random β-expansions with respect to rational approximations of real numbers were then described by Daubechies et al. [47], see also [50, 90], in relation to analog-to-digital conversion and in [79, 80] in relation to random number generation. Random one-dimensional continued fraction algorithms and their invariant measures were studied in [33, 53, 91, 94, 147]. Often, in terms of dynamics and the corresponding approximation properties, such a random system performs comparably to the best performing deterministic system present in the family $(T_i : X \to X)_{i\in I}$, but with the added flexibility that real numbers now have many different expansions assigned to them. Even though it is not quite clear at the moment whether placing multidimensional continued fraction algorithms in a random framework could yield similar advantages, it might be of interest to study this further.

7 Improving Jacobi–Perron Algorithm

In 1981, both Ito and Tanaka in [148] and Nakada in [116] introduced the notion of α-continued fractions for $\alpha \in [1/2, 1]$. This was done by replacing $\lfloor \frac{1}{x} \rfloor$ in the definition of the Gauss map T_G by $\lfloor \frac{1}{x} + 1 - \alpha \rfloor$ in [148] and by $\lfloor \frac{1}{|x|} + 1 - \alpha \rfloor$ in [116]. For the setup of Nakada, the sequence of the convergents of the corresponding α-continued fraction of a point form a subsequence of its regular sequence of convergents, and hence provide better approximations. In particular, for $\alpha \in [\frac{1}{2}, \frac{\sqrt{5}-1}{2}]$, all the corresponding α-continued fraction algorithms are isomorphic and provide better approximations than the ones given by the regular continued fractions. See in particular [96] for results on the entropy. One can use a similar idea to improve the convergence properties of the Jacobi–Perron algorithm. To keep the exposition simple, we will consider the nearest integer case, corresponding to $\alpha = \frac{1}{2}$, and we will not take absolute values inside the floor function (so the digits generated could be negative).

It is well known that when it comes to convergence speed, in one dimension the nearest integer algorithm performs best, see, e.g., [19]. For further results on nearest integer continued fractions see [3, 76, 153]. Dynamically the algorithm is given by the map $T(x) = \frac{1}{x} - |||\frac{1}{x}|||$, where $||| \cdot |||$ denotes the distance to the nearest integer as before. The map is well defined on the interval $[-\frac{1}{2}, \frac{1}{2}]$. The possible continued fraction digits for this algorithm are all integers n with $|n| \geq 2$. In higher dimensions a nearest integer version of the Jacobi–Perron algorithm has attracted some attention. See Sect. 7.2 for numerical values for the nearest integer Jacobi–Perron algorithm from [145]. We also recall that the Markov conditions on the digits produced by the classical Jacobi–Perron algorithm have a very simple form and that

the piecewise analyticity of the density of its invariant measure has been proved in [32].

7.1 First Dynamical Properties of the Algorithm

Let $d \geq 2$. The definition of the *d-dimensional nearest integer Jacobi–Perron map (NIJP)* on $C = \left[-\frac{1}{2}, \frac{1}{2}\right]^d$ is given by

$$T_0(x_1, \ldots, x_d) = \left(\frac{x_2}{x_1} - \left\lfloor \frac{x_2}{x_1} + \frac{1}{2} \right\rfloor, \ldots, \frac{x_d}{x_1} - \left\lfloor \frac{x_d}{x_1} + \frac{1}{2} \right\rfloor, \frac{1}{x_1} - \left\lfloor \frac{1}{x_1} + \frac{1}{2} \right\rfloor\right)$$

and can be used to create for each d-tuple $(x_1, \ldots, x_d) \in C$ a sequence of continued fraction approximations with the same denominator.

The matrix version of T_0 is therefore defined on

$$\Lambda = \{(y_0, y_1, \ldots, y_d) \in \mathbb{R}^{d+1} \setminus \{\mathbf{0}\} : y_0 \geq y_1, \cdots, y_d \geq 0\}$$

as

$$(y_0, y_1, \ldots, y_d) \mapsto \left(y_1, y_2 - \left\lfloor \frac{y_2}{y_1} + \frac{1}{2} \right\rfloor y_1, \ldots, y_d - \left\lfloor \frac{y_d}{y_1} + \frac{1}{2} \right\rfloor y_1, y_0 - \left\lfloor \frac{y_0}{y_1} + \frac{1}{2} \right\rfloor y_1\right),$$

and we have

$$\begin{aligned} &{}^t(y_0, y_1, \ldots, y_d) = {}^tA_0(y_0, y_1, \ldots, y_d) \cdot \\ &{}^t\left(y_1, y_2 - \left\lfloor \frac{y_2}{y_1} + \frac{1}{2} \right\rfloor y_1, \ldots, y_d - \left\lfloor \frac{y_d}{y_1} + \frac{1}{2} \right\rfloor y_1, y_0 - \left\lfloor \frac{y_0}{y_1} + \frac{1}{2} \right\rfloor y_1\right) \end{aligned}$$

with

$${}^tA_0(y_0, y_1, \ldots, y_d) = \begin{bmatrix} \left\lfloor \frac{1}{y_1} + \frac{1}{2} \right\rfloor & 1 & \left\lfloor \frac{y_2}{y_1} + \frac{1}{2} \right\rfloor & \cdots & \left\lfloor \frac{y_d}{y_1} + \frac{1}{2} \right\rfloor \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}.$$

Iterates of T_0 then produce a matrix

$$\begin{bmatrix} q_0^{(n)} & p_{0,1}^{(n)} & \cdots & p_{0,d}^{(n)} \\ q_1^{(n)} & p_{1,1}^{(n)} & \cdots & p_{1,d}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ q_d^{(n)} & p_{d,1}^{(n)} & \cdots & p_{d,d}^{(n)} \end{bmatrix}.$$

Note that this is not exactly the same matrix as in (3). This is mainly due to the fact that we use a different normalization for technical reasons: the largest coordinate is the first one (namely y_0), and not the last one, as before when working with vectors of the form $(\boldsymbol{\alpha}, 1)$. The $q_i^{(n)}$ are thus given here in the first column.

For $\mathbf{x} = (x_1, \ldots, x_d) \in C = \left[-\frac{1}{2}, \frac{1}{2}\right]^d$, set

$$a = a(\mathbf{x}) = \left|\left|\left|\frac{1}{x_1}\right|\right|\right|, \quad b^{(i)} = b^{(i)}(\mathbf{x}) = \left|\left|\left|\frac{x_i}{x_1}\right|\right|\right|, \ 2 \le i \le d.$$

The functions a and $b^{(i)}$ are piecewise constant on C. To be precise, for each $\mathbf{x} \in C$, it holds that $a(\mathbf{x}) = k$, $k \in \mathbb{Z}$, if and only if $\frac{2}{2k+1} < x_1 \le \frac{2}{2k-1}$, and, for each $2 \le i \le d$, it holds that $b^{(i)}(\mathbf{x}) = k$ if and only if

$$\begin{cases} x_1\left(k-\frac{1}{2}\right) \le x_2 < x_1\left(k+\frac{1}{2}\right), & \text{if } x_1 > 0, \\ x_1\left(k+\frac{1}{2}\right) < x_2 \le x_1\left(k-\frac{1}{2}\right), & \text{if } x_1 < 0. \end{cases}$$

Hence, if we let

$$C_{a,b_2,\ldots,b_d} = \left\{\mathbf{x} \in C \,:\, a(\mathbf{x}) = a,\ b^{(i)}(\mathbf{x}) = b_i,\ 2 \le i \le d\right\},$$

then the collection $\mathcal{C} = \{C_{a,b_2,\ldots,b_d}\}$ yields a partition of C, the elements of which are called the *cylinder sets* of T_0. Figure 1 shows the partition $\mathcal{C}$ of C for $d = 2$.

Note that $C_{a,b_2,\ldots,b_d} \neq \emptyset$ if and only if $|a| \ge 2$ and $0 \le |b_i| \le \lceil \frac{a}{2} \rceil$ for each $2 \le i \le d$. This implies that for the linear version of the NIJP algorithm not all matrices of the form

$$A_0 = \begin{bmatrix} a & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ b_2 & 1 & 0 & \cdots & 0 & 0 \\ b_3 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ b_d & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{16}$$

are allowed. It needs to hold that $|a| \ge 2$ and $0 \le |b_i| \le \lceil \frac{a}{2} \rceil$ for each $2 \le i \le d$, but there are more restrictions. Below we describe these restrictions in detail for the case $d = 2$ by showing that in this case T_0 admits a *Markov partition*, i.e., there

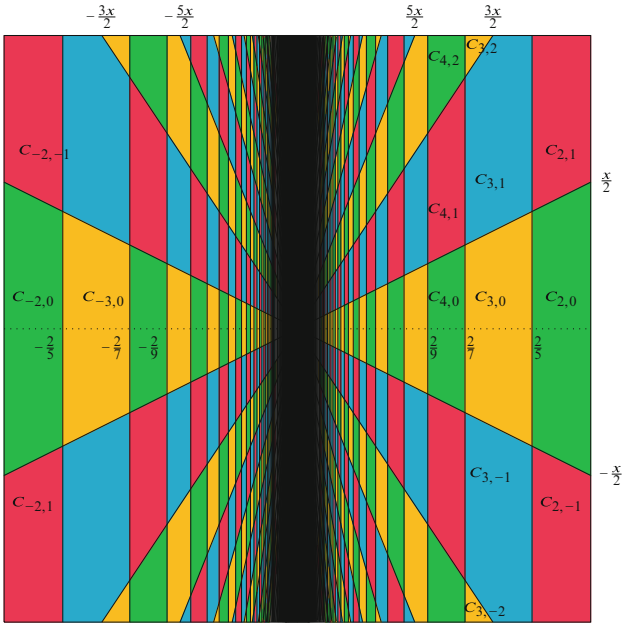
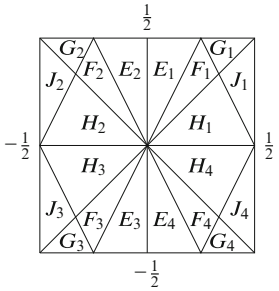


Fig. 1 The cylinder sets of the NIJP. The colors just serve to distinguish the different cylinder sets

Fig. 2 The Markov partition for the NIJP map for $d = 2$



exists a finite collection $\mathcal{P}$ of polygonal subsets of $[-\frac{1}{2}, \frac{1}{2}]^2$ that have the property that for any set $P \cap A$ with $P \in \mathcal{P}$ and $A \in C$, there exist $P_1, \ldots, P_N \in \mathcal{P}$ such that $T_0(P \cap A) = \bigcup_{i=1}^{N} P_i$ up to sets of zero Lebesgue measure. Figure 2 shows the 20 sets that are in $\mathcal{P}$ for the map T_0.

Theorem 7 *Let $d = 2$. Let $\mathcal{P}$ be the collection of disjoint subsets of C bounded by the ten lines $x_1 = 0$, $x_2 = 0$, $x_2 = \pm x_1$, $x_2 = \pm 2x_1$, $x_2 = \pm 1 \pm 2x_1$. Then $\mathcal{P}$ is a Markov partition for T_0.*

Proof To show that $\mathcal{P}$ from Fig. 2 is a Markov partition, we need to consider the image under T_0 of all sets $P \cap C_{a,b}$ for $P \in \mathcal{P}$ and the cylinder set $C_{a,b} \in C$. Due to symmetry, it is enough to only consider the sets in the first quadrant. We label

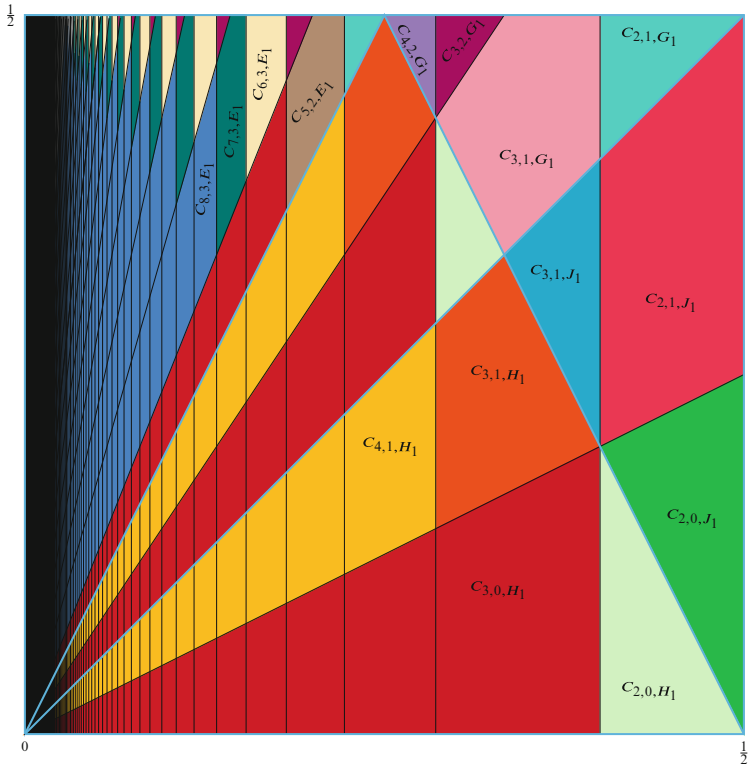


Fig. 3 The black lines indicate the boundaries of cylinder sets. The blue lines indicate the boundaries of the Markov partition elements. The colors of the polygons indicate the different types of sets $C_{a,b,P}$ according to their images under the NIJP map. Sets with the same color have the same type

the sets of $\mathcal{P}$ by E_i, F_i, G_i, H_i, J_i, $1 \le i \le 4$, as shown in Fig. 2 and use $C_{a,b,P}$ to denote the set $P \cap C_{a,b}$ with $P \in \mathcal{P}$ and $C_{a,b} \in C$. Figure 3 shows several of these sets in the first quadrant. Based on the images of the sets $C_{a,b,P}$ we distinguish 15 types, indicated by the different colors in Fig. 3. Hence, to prove that T_0 admits a Markov partition for $d = 2$ we need to compute the image under T_0 of each of these types of sets. We describe what happens to the set $C_{2,0,H_1}$.

On $C_{2,0,H_1}$ the map T_0 is given by $T_0(x_1, x_2) = (\frac{x_2}{x_1}, \frac{1}{x_1} - 2)$. The boundary of $C_{2,0,H_1}$ is the union of three sets (we take all intervals open, because the endpoints of the intervals have no impact on the Lebesgue measure of the sets):

Table 1 Images of the sets $C_{a,b,P}$ for $a, b \geq 0$ and $P \in \mathcal{P}$ under T_0

	$C_{a,b,P}$	$T(C_{a,b,P})$
Type 1	$C_{2,0,H_1}, C_{3,1,F_1}$	$E_1 \cup F_1 \cup G_1$
Type 2	$C_{2,0,J_1}$	$H_1 \cup J_1$
Type 3	$C_{2,1,J_1}$	$E_2 \cup F_2 \cup G_2 \cup H_2 \cup J_2$
Type 4	$C_{2,1,G_1}, C_{4,2,E_1}$	E_1
Type 5	$C_{a,0,H_1}$ for $a \geq 3$, $C_{a,1,F_1}$ for $a \geq 4$, $C_{a,2,E_1}$ for $a \geq 6$	$\bigcup_{i=1,4}(E_i \cup F_i \cup G_i \cup H_i \cup J_i)$
Type 6	$C_{3,1,H_1}, C_{4,2,F_1}$	$E_2 \cup F_2 \cup G_2 \cup H_2 \cup J_2 \cup H_3 \cup J_3$
Type 7	$C_{3,1,J_1}$	$E_3 \cup F_3 \cup G_3$
Type 8	$C_{3,1,G_1}$	$H_1 \cup J_1 \cup E_4 \cup F_4 \cup H_4$
Type 9	$C_{3,2,G_1}, C_{2k-1,k,E_1}$ for $k \geq 3$	$G_2 \cup J_2$
Type 10	$C_{a,1,H_1}$ for $a \geq 4$, $C_{a,2,F_1}$ for $a \geq 5$	$\bigcup_{i=2,3}(E_i \cup F_i \cup G_i \cup H_i \cup J_i)$
Type 11	$C_{4,2,G_1}$	$F_3 \cup G_3$
Type 12	$C_{5,2,E_1}$	$E_1 \cup F_1 \cup G_1 \cup H_1 \cup J_1 \cup E_4 \cup F_4 \cup H_4$
Type 13	C_{2k,k,E_1} for $k \geq 3$	$E_1 \cup E_2 \cup F_2 \cup G_2 \cup H_2 \cup J_2 \cup F_3 \cup G_3 \cup H_3 \cup J_3$
Type 14	C_{2k+1,k,E_1} for $k \geq 3$	$C \setminus (G_4 \cup J_4)$
Type 15	C_{a,b,E_1} for $b \geq 3$ and $a \geq 2b+2$	C

$$\partial_1 = \left\{(x_1, x_2) \in C \,:\, x_2 = 0,\ x_1 \in \left(\frac{2}{5}, \frac{1}{2}\right)\right\},$$
$$\partial_2 = \left\{(x_1, x_2) \in C \,:\, x_1 = \frac{2}{5},\ x_2 \in \left(0, \frac{1}{5}\right)\right\},$$
$$\partial_3 = \left\{(x_1, x_2) \in C \,:\, x_1 \in \left(\frac{2}{5}, \frac{1}{2}\right),\ x_2 = 1 - 2x_1\right\}.$$

Then

$$T_0(\partial_1) = \left\{(x_1, x_2) \in C \,:\, x_1 = 0,\ x_2 \in \left(0, \frac{1}{2}\right)\right\},$$
$$T_0(\partial_2) = \left\{(x_1, x_2) \in C \,:\, x_2 = \frac{1}{2},\ x_1 \in \left(0, \frac{1}{2}\right)\right\},$$
$$T_0(\partial_3) = \left\{(x_1, x_2) \in C \,:\, x_2 = x_1,\ x_1 \in \left(0, \frac{1}{2}\right)\right\}.$$

From this we can conclude that $T_0(C_{2,0,H_1}) = E_1 \cup F_1 \cup G_1$. A similar computation can be done for each of the sets $C_{a,b,P}$. Table 1 lists the images of each type of set $C_{a,b,P}$ in the first quadrant. See also Fig. 4. By symmetry, similar results are obtained for sets $C_{a,b,P}$ in the other quadrants, from which we can deduce that the collection $\mathcal{P}$ is a Markov partition for T_0 with $d = 2$.

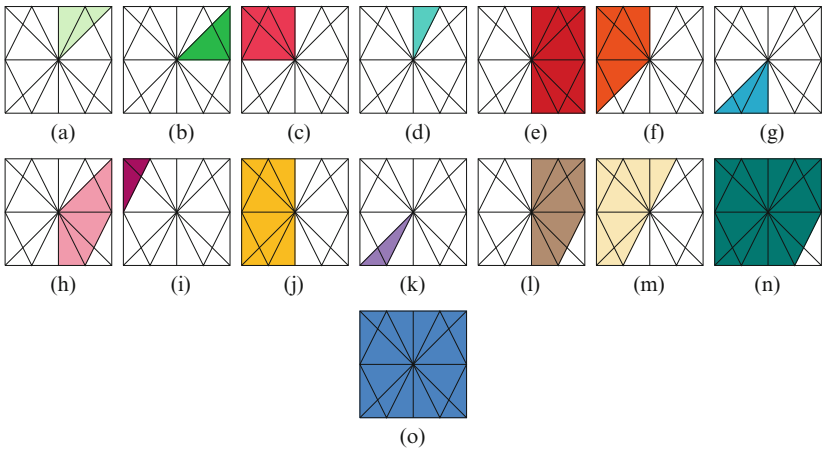


Fig. 4 The images of the sets from Fig. 3. (**a**) type 1 (**b**) type 2. (**c**) type 3. (**d**) type 4. (**e**) type 5. (**f**) type 6. (**g**) type 7. (**h**) type 8. (**i**) type 9. (**j**) type 10. (**k**) type 11. (**l**) type 12. (**m**) type 13. (**n**) type 14. (**o**) type 15

From Table 1 we can deduce the restrictions that apply to applications of the matrices from (16). For example, from the first two lines we read that the digit $(2, 0)$ can only be followed by a digit (a, b) with $a, b \geq 0$. Some restrictions are more complicated to describe and carry further. For example, the digit $(3, 2)$ can be followed by the digits $(-2, 0)$, $(-2, -1)$, $(-3, -1)$, $(-3, -2)$ and $(-4, -2)$. The digit $(-2, 0)$ can in principle be followed by any digit (a, b) with $a, b \leq 0$, but if one sees $(3, 2)$ followed by $(-2, 0)$, then this can only be followed by those digit (a, b) with $a, b \leq 0$ such that $C_{a,b} \in H_3 \cup J_3$. Knowing which sequences of digits are allowed tells us which matrix products involving matrices of the form (16) we have to analyze in order to obtaining numerical information about the approximation properties of the nearest integer Jacobi–Perron algorithm. Giving a fuller description of the allowed digit sequences from the results in Table 1 would be a first step in this direction.

Having a Markov partition can also help in finding invariant measures for the dynamical system given by T_0. Here we describe a different approach that might lead to an invariant measure that is absolutely continuous with respect to the Lebesgue measure. To prove the existence of an a.c.i.m., we follow the book of Schweiger [138, Chapter 4], where he analyzed the ergodic properties of the Jacobi–Perron algorithm (see also [14, 112]). The analysis is similar; we will be brief. First notice that only cylinders of type 15 are full, i.e., $T_0([a, b)]) = C$ up to sets of zero Lebesgue measure. All other types are non-full. However, after at most three iterations, any non-full cylinder is mapped to a region containing at least a fixed positive proportion, say q, of full cylinders. This implies that every non-full cylinder can be written as a countable union of disjoint full cylinders of higher rank (i.e., where more digits are specified), so that the collection of full cylinders generate

the Borel σ-algebra. This then allows one to define a jump transformation with full cylinders and satisfying the conditions of Rényi [132] (see also Theorem 8 in [138]). From this one concludes that the jump transformation admits an absolutely continuous invariant ergodic measure. Then, Theorem 11 (see also Theorem 18) of [138] implies that T_0 admits a finite absolutely continuous invariant ergodic measure.

7.2 Experimental Data

We now provide some experimental data due to Steiner [145] indicating a better behavior of the nearest integer Jacobi–Perron algorithm in terms of Lyapunov exponents than of its usual version.

For the nearest integer Jacobi–Perron algorithm (on the left) and the usual Jacobi–Perron algorithm (on the right), one gets the following experimental data: the usual Jacobi-Perron algorithm stops being strongly convergent in dimension 10, while the nearest integer Jacobi-Perron algorithm is strongly convergent up to dimension 13.

d	λ_1	λ_2	$1-\frac{\lambda_2}{\lambda_1}$
d = 2	1.72241	−0.691444	1.40144
d = 3	1.72394	−0.388217	1.22520
d = 4	1.72400	−0.24779	1.14372
d = 5	1.72408	−0.16873	1.09786
d = 6	1.72417	−0.11892	1.06897
d = 7	1.72413	−0.08522	1.049430
d = 8	1.72417	−0.06122	1.03551
d = 9	1.72409	−0.04347	1.02522
d = 10	1.72414	−0.02995	1.01737
d = 11	1.72413	−0.01939	1.01125
d = 12	1.72414	−0.01102	1.00640
d = 13	1.72409	−0.00425	1.00247
d = 14	1.72414	0.001304	0.99924

d	λ_1	λ_2	$1-\frac{\lambda_2}{\lambda_1}$
d = 2	1.20052	−0.448404	1.37351
d = 3	1.18560	−0.227877	1.19220
d = 4	1.17295	−0.13064	1.11138
d = 5	1.16579	−0.07882	1.067614
d = 6	1.16224	−0.04798	1.041279
d = 7	1.16068	−0.028202	1.02430
d = 8	1.15992	−0.014708	1.01268
d = 9	1.15956	−0.00505	1.004358
d = 10	1.159476	0.00217	0.99813
d = 11	1.159360	0.00776	0.993308
d = 12	1.159364	0.01221	0.98946
d = 13	1.159401	0.01586	0.986320
d = 14	1.15930	0.01889	0.983705

Acknowledgments We are greatly indebted to Wolfgang Steiner for the numerical data he provided us in Sect. 7.2. We would like to thank Women in Numbers Europe-4 (WINE4) for bringing us together and for giving us the opportunity to work on this project. The research of A.T. was supported by the DFG-project MO 1884/2-1 and the Collaborative Research Centre TRR 326 "Geometry and Arithmetic of Uniformized Structures". The research of V.B. was supported by the Agence Nationale de la Recherche through the project CODYS (ANR-18-CE40-0007)

References

1. N.H. Abel, Über die Integration der Differential-Formel $\frac{\varrho d.x}{R}$, wenn R und ϱ ganze Functionen sind. J. Reine Angew. Math. **1826**(1), 185–221 (1826)
2. P. Arnoux, V. Berthé, M. Minervino, W. Steiner, J. Thuswaldner, *Nonstationary Markov partitions and multidimensional continued fractions*. Preprint (2023)
3. W.W. Adams, On a relationship between the convergents of the nearest integer and regular continued fractions. Math. Comp. **33**(148), 1321–1331 (1979)
4. A. Avila, P. Hubert, A. Skripchenko, Diffusion for chaotic plane sections of 3-periodic surfaces. Invent. Math. **206**(1), 109–146 (2016)
5. A. Avila, P. Hubert, A. Skripchenko, On the Hausdorff dimension of the Rauzy gasket. Bull. Soc. Math. France **144**, 539–568 (2016)
6. P. Arnoux, S. Labbé, On some symmetric multidimensional continued fraction algorithms. Ergodic Theory Dynam. Syst. **38**(5), 1601–1626 (2018)
7. P. Arnoux, A. Nogueira, Mesures de Gauss pour des algorithmes de fractions continues multidimensionnelles. Ann. Sci. École Norm. Sup. **26**(6), 645–664 (1993)
8. W.W. Adams, M.J. Razar, Multiples of points on elliptic curves and continued fractions. Proc. Lond. Math. Soc. **s3-41**(3), 481–498 (1980)
9. V.I. Arnold, A-graded algebras and continued fractions. Commun. Pure Appl. Math. **42**(7), 993–1000 (1989)
10. L. Arnold, *Random Dynamical Systems*. Dynamical systems (Montecatini Terme, 1994). Lecture Notes in Math., vol. 1609 (Springer, Berlin, 1995), pp. 1–43
11. L. Arnold, *Random Dynamical Systems*. Springer Monographs in Mathematics (Springer, Berlin, 1998)
12. V.I. Arnold, Higher-dimensional continued fractions. Regul. Chaotic Dyn. **3**(3), 10–17 (1998)
13. P. Arnoux, T.A. Schmidt, Natural extensions and Gauss measures for piecewise homographic continued fractions. Bull. Soc. Math. France **147**(3), 515–544 (2019)
14. A. Broise-Alamichel, Y. Guivarc'h, Exposants caractéristiques de l'algorithme de Jacobi-Perron et de la transformation associée. Ann. Inst. Fourier (Grenoble) **51**(3), 565–686 (2001)
15. P.R. Baldwin, A convergence exponent for multidimensional continued-fraction algorithms. J. Statist. Phys. **66**(5-6), 1507–1526 (1992)
16. F. Bergeron, J. Berstel, S. Brlek, C. Duboc, Addition chains using continued fractions. J. Algorithms **10**(3), 403–412 (1989)
17. V. Berthé, E. Cesaratto, P. Rotondo, M.D. Safe, Lochs-type theorems beyond positive entropy. Monatshefte für Mathematik **200**, 737–779 (2023)
18. W. Bosma, K. Dajani, C. Kraaikamp, *Entropy and Counting Correct Digits*. Tech. Report 9925, University of Nijmegen, 1999. http://www-math.sci.kun.nl/math/onderzoek/reports/reports1999.html
19. J. Bourdon, B. Daireaux, B. Vallée, *Dynamical Analysis of α-Euclidean Algorithms*, vol. 44, 2002. Analysis of Algorithms, pp. 246–285
20. L. Bernstein, *The Jacobi-Perron Algorithm—Its Theory and Application*. Lecture Notes in Mathematics, vol. 207 (Springer, Berlin, New York, 1971)
21. F. Beukers, Geodesic continued fractions and LLL. Indag. Math. (N.S.) **25**, 632–645 (2014)
22. M. Baake, U. Grimm, *Aperiodic Order, vol. 1*. Encyclopedia of Mathematics and Its Applications, vol. 149 (Cambridge University Press, Cambridge, 2013)
23. L. Barreira, G. Iommi, Partial quotients of continued fractions and β-expansions. Nonlinearity **21**(10), 2211–2219 (2008)
24. P. Billingsley, *Ergodic Theory and Information* (Robert E. Krieger Publishing Co., Huntington, NY, 1978). Reprint of the 1965 original
25. L. Babai, B. Just, F. Meyer auf der Heide, On the limits of computations with the floor function. Inform. Comput. **78**(2), 99–107 (1988)
26. W. Bosma, H. Jager, F. Wiedijk, Some metrical observations on the approximation by continued fractions. Nederl. Akad. Wetensch. Indag. Math. **45**(3), 281–299 (1983)

27. V. Berthé, L. Lhote, B. Vallée, The Brun gcd algorithm in high dimensions is almost always subtractive. J. Symbolic Comput. **85**, 72–107 (2018)
28. F. Bagemihl, J.R. McLaughlin, Generalization of some classical theorems concerning triples of consecutive convergents to simple continued fractions. J. Reine Angew. Math. **221**, 146–149 (1966)
29. V. Baladi, A. Nogueira, Lyapunov exponents for non-classical multidimensional continued fraction algorithms. Nonlinearity **9**, 1529–1546 (1996)
30. É. Borel, Contribution à l'analyse arithmétique du continu, J. Math. Pures et Appl. (5) **9**, 329–375 (1903)
31. A.J. Brentjes, *Multidimensional Continued Fraction Algorithms* (Mathematisch Centrum, Amsterdam, 1981)
32. A. Broise, Fractions continues multidimensionnelles et lois stables. Bull. Soc. Math. France **124**(1), 97–139 (1996)
33. W. Bahsoun, M. Ruziboev, B. Saussol, Linear response for random dynamical systems. Adv. Math. **364**, 107011, 44 (2020)
34. V. Brun, En generalisation av kjedebrøken I. Skr. Vidensk.-Selsk. Christiana Math.-Nat. Kl. **6**, 1–29(1919)
35. V. Brun, En generalisation av kjedebrøken II. Skr. Vidensk.-Selsk. Christiana Math.-Nat. Kl. **6**, 1–24 (1920)
36. V. Brun, *Algorithmes euclidiens pour trois et quatre nombres*, Treizième congrès des mathèmaticiens scandinaves, tenu à Helsinki 18-23 août 1957 (Mercators Tryckeri, Helsinki, 1958), pp. 45–64
37. W. Bosma, I. Smeets, *Finding Simultaneous Diophantine Approximations with Prescribed Quality*, in *ANTS X—Proceedings of the Tenth Algorithmic Number Theory Symposium*. Open Book Ser., vol. 1 (Math. Sci. Publ., Berkeley, CA, 2013), pp. 167–185
38. V. Berthé, W. Steiner, J.M. Thuswaldner, On the second Lyapunov exponent of some multidimensional continued fraction algorithms. Math. Comput. **90**(328), 883–905 (2021)
39. V. Berthé, W. Steiner, J.M. Thuswaldner, Multidimensional continued fractions and symbolic codings of toral translations. J. Eur. Math. Soc. (JEMS) **25**, 4997–5057 (2023)
40. T. Camus, *Méthodes algorithmiques pour les réseaux algébriques. (algorithmic methods for algebraic lattices)*, Ph.D. thesis, Grenoble Alpes University, France, 2017
41. J.W.S. Cassels, *An Introduction to Diophantine Approximation* (Hafner Publishing Co., New York, 1972). Facsimile reprint of the 1957 edition, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45.
42. Y. Cheung, N. Chevallier, *About the value of the two dimensional levy's constant*. Preprint (2021). https://arxiv.org/pdf/2107.01907.pdf
43. Y. Cheung, N. Chevallier, *Levy-khintchin theorem for best simultaneous diophantine approximations*. Preprint (2019). arXiv:1906.11173
44. J. Cassaigne, S. Ferenczi, A. Messaoudi, Weak mixing and eigenvalues for Arnoux-Rauzy sequences. Ann. Inst. Fourier (Grenoble) **58**(6), 1983–2005 (2008)
45. Y. Cheung, Hausdorff dimension of the set of singular pairs. Ann. Math. (2) **173**(1), 127–167 (2011)
46. N. Chevallier, Best simultaneous Diophantine approximations and multidimensional continued fraction expansions. Mosc. J. Comb. Number Theory **3**, 3–56 (2013)
47. I. Daubechies, R.A. DeVore, C.S. Güntürk, V.A. Vaishampayan, A/D conversion with imperfect quantizers. IEEE Trans. Inform. Theory **52**(3), 874–885 (2006)
48. K. Dajani, A. Fieldsteel, Equipartition of interval partitions and an application to number theory. Proc. Am. Math. Soc. **129**(12), 3453–3460 (2001)
49. H. Daudé, P. Flajolet, B. Vallée, An average-case analysis of the Gaussian algorithm for lattice reduction. Combin. Probab. Comput. **6**, 397–433 (1997)
50. I. Daubechies, S. Güntürk, Y. Wang, Ö. Yı lmaz, The golden ratio encoder. IEEE Trans. Inform. Theory **56**(10), 5097–5110 (2010)
51. K. Dajani, C. Kraaikamp, *Ergodic Theory of Numbers*. Carus Mathematical Monographs, vol. 29 (Mathematical Association of America, Washington, DC, 2002)

52. K. Dajani, C. Kraaikamp, Random β-expansions. Ergodic Theory Dynam. Syst. **23**(2), 461–479 (2003)
53. K. Dajani, C. Kalle, M. Maggioni, Matching for random systems with an application to minimal weight expansions. Nonlinearity **34**(6), 3676–3708 (2021)
54. P. de Rooij, *Efficient Exponentiation Using Precomputation and Vector Addition Chains*. Advances in Cryptology—EUROCRYPT '94 (Perugia). Lecture Notes in Comput. Sci., vol. 950 (Springer, Berlin, 1995), pp. 389–399
55. H. Ei, S. Ito, H. Nakada, R. Natsui, On the construction of the natural extension of the Hurwitz complex continued fraction map. Monatsh. Math. **188**(1), 37–86 (2019)
56. C. Elsner, M. Schmidt, *Kroncrypt—a new symmetric cryptosystem based on kronecker's approximation theorem*. Cryptology ePrint Archive. Paper 2009/416, 2009. https://eprint.iacr.org/2009/416
57. M. Einsiedler, T. Ward, *Ergodic Theory with a View Towards Number Theory*. Graduate Texts in Mathematics, vol. 259 (Springer, London, 2011)
58. H.R.P. Ferguson, A noninductive GL(n, $\mathbf{Z}$) algorithm that constructs integral linear relations for n $\mathbf{Z}$-linearly dependent real numbers. J. Algorithms **8**(1), 131–145 (1987)
59. H.R.P. Ferguson, R.W. Forcade, Generalization of the Euclidean algorithm for real numbers to all dimensions higher than two. Bull. Am. Math. Soc. (N.S.) **1**(6), 912–914 (1979)
60. T. Fujita, S. Ito, M. Keane, M. Ohtsuki, On almost everywhere exponential convergence of the modified Jacobi-Perron algorithm: a corrected proof. Ergodic Theory Dynam. Syst. **16**(6), 1345–1352 (1996)
61. H. Furstenberg, H. Kesten, Products of random matrices. Ann. Math. Statist. **31**, 457–469 (1960)
62. C. Fieker, D. Stehlé, *Short Bases of Lattices over Number Fields*. Algorithmic Number Theory. Lecture Notes in Comput. Sci., vol. 6197 (Springer, Berlin, 2010), pp. 157–173
63. C. Fougeron, A. Skripchenko, Simplicity of spectra for certain multidimensional continued fraction algorithms. Monatsh. Math. **194**(4), 767–787 (2021)
64. H. Furstenberg, *Stationary Processes and Prediction Theory*. Annals of Mathematics Studies, vol. 44 (Princeton University Press, Princeton, NJ, 1960)
65. H. Furstenberg, Noncommuting random products. Trans. Am. Math. Soc. **108**, 377–428 (1963)
66. D.J. Grabiner, J.C. Lagarias, Cutting sequences for geodesic flow on the modular surface and continued fractions. Monatsh. Math. **133**(4), 295–339 (2001)
67. D.J. Grabiner, Farey nets and multidimensional continued fractions. Monatsh. Math. **114**(1), 35–61 (1992)
68. E. Heine, C.G.J. Jacobi, Allgemeine Theorie der kettenbruchähnlichen Algorithmen, in welchen jede Zahl aus drei vorhergehenden gebildet wird. J. Reine Angew. Math. **69**, 29–64 (1868)
69. J. Håstad, B. Just, J.C. Lagarias, C.-P. Schnorr, Polynomial time algorithms for finding integer relations among real numbers. SIAM J. Comput. **18**(5), 859–881 (1989)
70. D.M. Hardcastle, K. Khanin, On almost everywhere strong convergence of multi-dimensional continued fraction algorithms. Ergodic Theory Dynam. Syst. **20**(6), 1711–1733 (2000)
71. D.M. Hardcastle, K. Khanin, Continued fractions and the d-dimensional Gauss transformation. Commun. Math. Phys. **215**(3), 487–515 (2001)
72. D.M. Hardcastle, K. Khanin, The d-dimensional Gauss transformation : strong convergence and Lyapunov exponents. Exp. Math. **11**, 119–129 (2002)
73. G. Havas, B.S. Majewski, K.R. Matthews, Extended GCD and Hermite normal form algorithms via lattice basis reduction. Exp. Math. **7**, 125–136 (1998)
74. G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers* (Oxford Science Publications, Oxford, 1979)
75. M. Iosifescu, C. Kraaikamp, *Metrical Theory of Continued Fractions* (Dordrecht Kluwer Academic Publishers, 2002)
76. H. Jager, Metrical results for the nearest integer continued fraction. Nederl. Akad. Wetensch. Indag. Math. **47**(4), 417–427 (1985)

77. H. Jager, The distribution of certain sequences connected with the continued fraction. Nederl. Akad. Wetensch. Indag. Math. **48**(1), 61–69 (1986)
78. H. Jager, C. Kraaikamp, On the approximation by continued fractions. Nederl. Akad. Wetensch. Indag. Math. **51**(3), 289–307 (1989)
79. Y. Jitsumatsu, K. Matsumura, A β-ary to binary conversion for random number generation using a β encoder. IEICE (NOLTA), 38–55 (2016)
80. Y. Jitsumatsu, K. Matsumura, T. Kohda, K. Aihara, *Pseudo-random number generator using beta-encoder cmos circuit*, in *The 3rd Int. Symp. Innovative Mathematical Modelling* (2013), p. 107
81. O. Jenkinson, M. Pollicott, P. Vytnova, Rigorous computation of diffusion coefficients for expanding maps. J. Stat. Phys. **170**, 221–253 (2018)
82. B. Just, *Integer Relations Among Algebraic Numbers*. Mathematical Foundations of Computer Science, 1989 (Porąbka-Kozubnik, 1989). Lecture Notes in Comput. Sci., vol. 379 (Springer, Berlin, 1989), pp. 314–320
83. B. Just, Integer relations among algebraic numbers. Math. Comput. **54**(189), 467–477 (1990)
84. B. Just, Generalizing the continued fraction algorithm to arbitrary dimensions. SIAM J. Comput. **21**(5), 909–926 (1992)
85. A.M. Kane, On the use of continued fractions for stream ciphers. IACR Cryptol. **2013**, 319 (2009). ePrint Arch.
86. O.N. Karpenkov, Constructing multidimensional periodic continued fractions in the sense of Klein. Math. Comput. **78**(267), 1687–1711 (2009)
87. O.N. Karpenkov, *Geometry of Continued Fractions*. Algorithms and Computation in Mathematics, vol. 26 (Springer, Heidelberg, 2013)
88. O.N. Karpenkov, On Hermite's problem, Jacobi-Perron type algorithms, and Dirichlet groups. Acta Arith. **203**(1), 27–48 (2022)
89. A.Ya. Khintchine, *Continued Fractions*. Translated by Peter Wynn (P. Noordhoff Ltd., Groningen, 1963)
90. T. Kohda, Y. Horio, Y. Takahashi, K. Aihara, Beta encoders: symbolic dynamics and electronic implementation. Int. J. Bifur. Chaos Appl. Sci. Eng. **22**(9), 1230031, 55 (2012)
91. C. Kalle, T. Kempton, E. Verbitskiy, The random continued fraction transformation. Nonlinearity **30**(3), 1182–1203 (2017)
92. K. Khanin, J. Lopes Dias, J. Marklof, Multidimensional continued fractions, dynamical renormalization and KAM theory. Commun. Math. Phys. **270**(1), 197–231 (2007)
93. C. Kalle, N. Langeveld, M. Maggioni, S. Munday, Matching for a family of infinite measure continued fraction transformations. Discrete Contin. Dyn. Syst. **40**(11), 6309–6330 (2020)
94. C. Kalle, V. Matache, M. Tsujii, E. Verbitskiy, Invariant densities for random continued fractions. J. Math. Anal. Appl. **512**(2), Paper No. 126163, 28 (2022)
95. E. Korkina, La périodicité des fractions continues multidimensionnelles. C. R. Acad. Sci. Paris Sér. I Math. **319**(8), 777–780 (1994)
96. C. Kraaikamp, T.A. Schmidt, W. Steiner, Natural extensions and entropy of α-continued fractions. Nonlinearity **25**(8), 2207–2243 (2012)
97. C. Kalle, E. Verbitskiy, B. Zeegers, Random Lochs' theorem. Studia Math. **267**(2), 201–239 (2022)
98. G. Lachaud, Polyèdre d'Arnol' d et voile d'un cône simplicial: analogues du théorème de Lagrange. C. R. Acad. Sci. Paris Sér. I Math. **317**(8), 711–716 (1993)
99. G. Lachaud, *Klein Polygons and Geometric Diagrams*. Number theory (Tiruchirapalli, 1996). Contemp. Math., vol. 210 (Amer. Math. Soc., Providence, RI, 1998), pp. 365–372
100. G. Lachaud, *Sails and Klein Polyhedra*. Number theory (Tiruchirapalli, 1996). Contemp. Math., vol. 210 (Amer. Math. Soc., Providence, RI, 1998), pp. 373–385
101. J.C. Lagarias, Best simultaneous Diophantine approximations. I. Growth rates of best approximation denominators. Trans. Am. Math. Soc. **272**, 545–554 (1982)
102. J.C. Lagarias, Best simultaneous Diophantine approximations. II. Behavior of consecutive best approximations. Pacific J. Math. **102**, 61–88 (1982)

103. J.C. Lagarias, *The computational complexity of simultaneous Diophantine approximation problems*, in *23rd Annual Symposium on Foundations of Computer Science* (Chicago, Ill., 1982) (IEEE, New York, 1982), pp. 32–39
104. J.C. Lagarias, The computational complexity of simultaneous Diophantine approximation problems. SIAM J. Comput. **14**, 196–209 (1985)
105. J.C. Lagarias, The quality of the Diophantine approximations found by the Jacobi-Perron algorithm and related algorithms. Monatsh. Math. **115**(4), 299–328 (1993)
106. J.C. Lagarias, Geodesic multidimensional continued fractions. Proc. Lond. Math. Soc. (3) **69**, 464–488 (1994)
107. L. Lhote, *Computation of a Class of Continued Fraction Constants*. Analytic Algorithmics and Combinatorics ANALCO 2014 (SIAM, 2004), pp. 199–210
108. A.K. Lenstra, H.W. Lenstra, Jr., L. Lovász, Factoring polynomials with rational coefficients. Math. Ann. **261**(4), 515–534 (1982)
109. G. Lochs, Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch. Abh. Math. Sem. Univ. Hamburg **27**, 142–144 (1964)
110. B. Li, J. Wu, Beta-expansion and continued fraction expansion. J. Math. Anal. Appl. **339**(2), 1322–1331 (2008)
111. R. Mönkemeyer, Über Fareynetze in n Dimensionen. Math. Nachr. **11**, 321–344 (1954)
112. D.H. Mayer, Approach to equilibrium for locally expanding maps in $\mathbf{R}^k$. Commun. Math. Phys. **95**(1), 1–15 (1984)
113. R. Meester, A simple proof of the exponential convergence of the modified Jacobi-Perron algorithm. Ergodic Theory Dynam. Syst. **19**(4), 1077–1083 (1999)
114. M. Morse, G.A. Hedlund, Symbolic dynamics II. Sturmian trajectories. Am. J. Math. **62**, 1–42 (1940)
115. N.G. Moshchevitin, *Best Diophantine Approximations: The Phenomenon of Degenerate Dimension*. Surveys in geometry and number theory: reports on contemporary Russian mathematics. London Math. Soc. Lecture Note Ser., vol. 338 (Cambridge University Press, Cambridge, 2007), pp. 158–182
116. H. Nakada, Metrical theory for a class of continued fraction transformations and their natural extensions. Tokyo J. Math. **4**(2), 399–426 (1981)
117. H. Napias, A generalization of the LLL-algorithm over Euclidean rings or orders. J. Théor. Nombres Bordeaux **8**, 387–396 (1996)
118. H. Niederreiter, *Sequences with Almost Perfect Linear Complexity Profile*, in *Advances in Cryptology—EUROCRYPT' 87* (Berlin, Heidelberg), ed. by D. Chaum, W.L. Price (Springer, Berlin, Heidelberg, 1988), pp. 37–51
119. A. Nogueira, The Borel-Bernstein theorem for multidimensional continued fractions. J. Anal. Math. **85**, 1–41 (2001)
120. P.Q. Nguyen, D. Stehlé, *LLL on the Average*. Algorithmic number theory. Lecture Notes in Comput. Sci., vol. 4076 (Springer, Berlin, 2006), pp. 238–256
121. P.Q. Nguyen, B. Vallée (ed.) *The LLL Algorithm*. Survey and Applications. Information Security and Cryptography (Springer, Dordrecht, 2010)
122. N. Obrechkoff, Sur l'approximation des nombres irrationnels par des nombres rationnels. C. R. Acad. Bulgare Sci. **3**(1), 1–4 (1950,1951)
123. G. Panti, Multidimensional continued fractions and a Minkowski function. Monatsh. Math. **154**(3), 247–264 (2008)
124. O. Perron, Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus. Math. Ann. **64**(1), 1–76 (1907)
125. O. Perron, Über diophantische approximationen. Math. Annalen **83**, 77–84 (1921)
126. V. Platonov, G. Fedorov, On the periodicity of continued fractions in hyperelliptic fields. Doklady Math. **95**, 254–258 (2017)
127. E.V. Podsypanin, A generalization of the continued fraction algorithm that is related to the Viggo Brun algorithm. Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **67**, 184–194, 227 (1977). Studies in number theory (LOMI), 4

128. M. Pollicott, Maximal Lyapunov exponents for random matrix products. Invent. Math. **181**(1), 209–226 (2010)
129. A. Pethő, M.E. Pohst, C. Bertók, On multidimensional Diophantine approximation of algebraic numbers. J. Number Theory **171**, 422–448 (2017)
130. F. Pappalardi, A.J. Van Der Poorten, Pseudo-elliptic integrals, units, and torsion. J. Aust. Math. Soc. **79**(3), 335–347 (2005). Cited by: 4; All Open Access, Bronze Open Access, Green Open Access
131. N. Pytheas Fogg, in *Substitutions in Dynamics, Arithmetics and Combinatorics*, ed. by V. Berthé, S. Ferenczi, C. Mauduit, A. Siegel. Lecture Notes in Mathematics, vol. 1794 (Springer, 2002)
132. A. Rényi, Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hungar. **8**, 477–493 (1957)
133. A. Rosenfeld, R. Klette, Digital straightness. Electron. Notes Theor. Comput. Sci. **46**, 1–32 (2001). IWCIA 2001, 8th International Workshop on Combinatorial Image Analysis
134. F. Schweiger, *The Metrical Theory of Jacobi-Perron Algorithm*. Lecture Notes in Mathematics, vol. 334 (Springer, Berlin, New York, 1973)
135. F. Schweiger, On the invariant measure for Jacobi-Perron algorithm. Math. Pannon. **1**(2), 91–106 (1990)
136. W.S. Schmidt, *Diophantine Approximation*. Lecture Notes in Mathematics, vol. 785 (Springer, 1996)
137. F. Schweiger, *The Exponent of Convergence for the* 2*-Dimensional Jacobi-Perron Algorithm*, in *Proceedings of the Conference on Analytic and Elementary Number Theory* (Vienna), ed. by W.G. Nowak, J. Schoissengeier (1996), pp. 207–213
138. F. Schweiger, *Multidimensional Continued Fractions*. Oxford Science Publications (Oxford University Press, Oxford, 2000)
139. B. R. Schratzberger, The quality of approximation of Brun's algorithm in three dimensions. Monatshefte für Mathematik **134**, 143–157 (2001)
140. B. Schratzberger, On the singularization of the two-dimensional Jacobi-Perron algorithm. Exp. Math. **16**(4), 441–454 (2007) (English)
141. B.R. Schratzberger, A conversion algorithm based on the technique of singularization. Theor. Comput. Sci. **391**(1-2), 138–149 (2008) (English)
142. E.S. Selmer, Continued fractions in several dimensions. Nordisk Nat. Tidskr. **9**, 37–43, 95 (1961)
143. B. Sendov, Der Zahlensche Satz über die singulären Kettenbrüche und die Kettenbrüche nach nächsten Ganzen. Annuaire Univ. Sofia Fac. Sci. Phys. Math. Livre 1 Math. **54**, 251–258 (1959/1960)
144. R. Sturman, J.-L. Thiffeault, Lyapunov exponents for the random product of two shears. J. Nonlinear Sci. **29**(2), 593–620 (2019)
145. W. Steiner, Personnal communication
146. G. Szekeres, Multidimensional continued fractions. Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **13**, 113–140 (1970/1971)
147. T. Taylor-Crush, On the regularity and approximation of invariant densities for random continued fractions. Dyn. Syst. **36**(1), 1–18 (2021)
148. S. Tanaka, S. Ito, On a family of continued-fraction transformations and their ergodic properties. Tokyo J. Math. **4**(1), 153–175 (1981)
149. J.C. Tong, The conjugate property of the Borel theorem on Diophantine approximation. Math. Z. **184**(2), 151–153 (1983)
150. A.J. Van der Poorten, X.C. Tran, Quasi-elliptic integrals and periodic continued fractions. Monatshefte für Mathematik **131**(2), 155–169 (2000)
151. M. Viana, *Lectures on Lyapunov Exponents*. Cambridge Studies in Advanced Mathematics, vol. 145 (Cambridge University Press, Cambridge, 2014)
152. M. Wiener, *Cryptanalysis of Short RSA Secret Exponents*, in *Advances in Cryptology—EUROCRYPT '89* (Berlin, Heidelberg), ed. by J.-J. Quisquater, J. Vandewalle (Springer, Berlin, Heidelberg, 1990), pp. 372–372

153. H.C. Williams, Some results concerning the nearest integer continued fraction expansion of $\sqrt{D}$. J. Reine Angew. Math. **315**, 1–15 (1980)
154. D. Wübben, D. Seethaler, J. Jaldén, G. Matz, Lattice reduction. IEEE Signal Process. Mag. **28**(3), 70–91 (2011)
155. Y. Yu, L. Ducas, *Second Order Statistical Behavior of LLL and BKZ*. Selected areas in cryptography—SAC 2017. Lecture Notes in Comput. Sci., vol. 10719 (Springer, 2018), pp. 3–22
156. J.-C. Yoccoz, *Continued Fraction Algorithms for Interval Exchange Maps: an Introduction*. Frontiers in Number Theory, Physics, and Geometry. I (Springer, Berlin, 2006), pp. 401–435
157. A. Zorich, Deviation for interval exchange transformations. Ergodic Theory Dynam. Syst. **17**, 1477–1499 (1997)