Dynamical Analysis of a Class of Euclidean Algorithms

Brigitte VALLÉE
GREYC, Université de Caen, F-14032 Caen (France)
Brigitte.Vallee@info.unicaen.fr

Abstract

We develop a general framework for the analysis of algorithms of a broad Euclidean type. The average-case complexity of an algorithm is seen to be related to the analytic behaviour in the complex plane of the set of elementary transformations determined by the algorithm. The methods rely on properties of transfer operators suitably adapted from dynamical systems theory. As a consequence, we obtain precise average-case analyses of algorithms for evaluating the Jacobi symbol of computational number theory fame, thereby solving conjectures of Bach and Shallit. These methods also provide a unifying framework for the analysis of an entire class of gcd-like algorithms together with new results regarding the probable behaviour of their cost functions.

1 Introduction

In this paper, we provide new analyses of several classical and semi-classical variants of the Euclidean algorithm. A strong motivation of our study is a group of gcd-like algorithms that compute the Jacobi symbol. These algorithms all share a common structure: they perform a sequence of iterations, and each iteration involves a division and an exchange. In order to compute the Jacobi Symbol, the classical Euclidean division is replaced by a slightly modified division that is named a pseudo-Euclidean division and whose cost is very close to that of a classical division.

There are (at least) six kinds of Euclidean divisions that can be defined, and each of them gives rise to an algorithm for computing the GCD. Euclid's Algorithm based on the usual division was discovered as early as 300BC, and is "the grandfather of all the algorithms", as Knuth says. It is called here the Classical Euclidean Algorithm and is denoted by (G). The other five Euclidean Algorithms are called By-Excess (L), Centered (K), Even (E), Odd (O) and Subtractive (T). In addition, each Euclidean division can be slightly modified by removing the powers of two from the remainder. This modification corresponds to a pseudo-Euclidean division that gives rise to an algorithm for computing the Jacobi symbol. We study five algorithms of this type that are modifications of the previous ones. The first four pseudo-Euclidean Algorithms are called Pseudo-Classical $(\widehat{G})$, Pseudo-By-Excess $(\widehat{L})$, Pseudo-Centered $(\widehat{K})$, Pseudo-Odd $(\widehat{O})$. The last one, that is the pseudo-version $(\widehat{T})$ of the Subtractive algorithm is the well–known Binary Algorithm, and is also denoted by (B).

Motivations. The complexity analysis of each algorithm aims to evaluate the number of iterations that are performed during execution and we are mainly interested here in the average-case complexity. The complexity of the first four algorithms is now known: Euclid's algorithm was analysed in the worst case in 1733 by de Lagny, then in the average-case around 1969 independently by Heilbronn [19] and Dixon [10], and finally in distribution by Hensley [20] who proved in 1994 that the Euclidean algorithm has Gaussian behaviour; see Knuth's and Shallit's vivid accounts [26, 39]. The Centered algorithm (K) has been analysed by Rieger [35]. The Subtractive algorithm (T) was studied by Knuth and Yao [51], and Vardi [49] analysed the By-Excess Algorithm (L) by comparing it to the Subtractive Algorithm. The methods employed so far are rather disparate, and their applicability to new situations is somewhat unclear. Here, we design a unifying framework that additionally provides new results on the distribution of costs.

Three of the remaining Algorithms, the Pseudo-Classical ($\widehat{G}$), the Pseudo-Centered ($\widehat{K}$), and the Even algorithm (E), have been studied by Shallit [38] who limited himslef to a worst-case analysis. The present paper solves completely a conjecture of Bach and Shallit. Indeed, in [38], Shallit writes: "Bach has also suggested that one could investigate the average number of division steps in [these three algorithms] [...]. This analysis is probably feasible to carry out for the Even Algorithm, and it seems likely that the average number of division steps is $\Theta(\log^2 N)$. However, determining the average behaviour for the two other algorithms seems quite hard."

Methods. The first methods used in the average-case analysis of Euclidean Algorithms range from combinatorial (Heilbronn) to probabilistic (Dixon). In parallel, studies by Lévy, Khinchin, Kuzmin and Wirsing had established the metric theory of continued fractions (that can be viewed as continuous versions of Euclidean algorithms) by means of a specific density transformer. The more recent works rely for a good deal on *transfer operators*, a far-reaching generalization of density transformers, originally introduced by Ruelle [36, 37] in connection with the thermodynamic formalism and dynamical systems theory [4]. Examples are Mayer's studies on the continued fraction transformation [32], Hensley's work [20] and several papers of the author [45, 46] including her analysis of the Binary GCD Algorithm [47].

Our approach here consists in viewing an algorithm of the broad gcd type as a dynamical system, where each iterative step is a linear fractional transformation (LFT) of the form $(az+b)/(cz+d)$. A specific set of transformations is then associated to each algorithm. The introduction of pseudo-divisions, (that involve dyadic valuation of integers) leads to complex dynamical systems termed random since they involve random choices. Furthermore, the control structure that is simple for dynamical systems relative to Euclidean algorithms may become multimodal (what we call Markovian) in the case of pseudo-Euclidean algorithms. It will appear from our treatment that the computational complexity of an algorithm is in fact dictated by the collective dynamics of an associated set of transformations. More precisely, two factors intervene: (i) the characteristics of the LFT's in the complex domain; (ii) their contraction properties, notably near fixed points. There results a classification of gcd-like algorithms in terms of the average number of iterations: some of them are "fast", that is, of logarithmic complexity $\Theta(\log N)$, both on average and in the worst case while others are "slow", that is, of the "log-squared" type $\Theta(\log^2 N)$ on average. (The worst-case complexity of the slow algorithms is not even polynomial in $\log N$, being in fact of order $\Theta(N)$.)

It is established here that strong contraction properties of the elementary transformations that build up a gcd-like algorithm entail logarithmic cost, while the presence of an indifferent fixed-point leads to log-squared behaviour. In the latter case, the analysis requires a special twist that takes its inspiration from the study of intermittency phenomena in physical systems that was introduced by Bowen [6] and is nicely exposed in a paper of Prellberg and Slawny [34]. An additional benefit of our approach is to open access to characteristics of the distribution of costs, including information on moments: the fast algorithms appear to have concentration of distribution—the cost converges in probability to its mean—while the slow ones exhibit an extremely large dispersion of costs.

Technically, this paper relies on a description of relevant parameters by means of generating functions, a by now common tool in the average-case of algorithms [13, 14]. As is usual in number theory contexts, the generating functions are Dirichlet series. They are first proved to be algebraically related to specific operators that encapsulate all the important informations relative to the "dynamics" of the algorithm. Their analytical properties depend on spectral properties of the operators, most notably the existence of a "spectral gap" that separates the dominant eigenvalue from the remainder of the spectrum. This determines the singularities of the Dirichlet series of costs. The asymptotic extraction of coefficients is then achieved by means of Tauberian theorems, one of the many ways to derive the prime number theorem. Average complexity estimates finally result. The main thread of the paper is thus adequately summarized by the chain:

$$\begin{gathered}\text{Euclidean algorithm} \rightsquigarrow \text{Associated transformations} \rightsquigarrow \text{Transfer operator}\\ \rightsquigarrow \text{Dirichlet series of costs} \rightsquigarrow \text{Tauberian inversion} \rightsquigarrow \text{Average-case complexity.}\end{gathered}$$

This chain then leads to effective and simple criteria for distinguishing slow algorithms from fast ones, for establishing concentration of distribution, for analysing various cost parameters of algorithms, etc. The constants that intervene in the analysis are in all cases computable numbers closely related to the entropy of the associated dynamical system. In the the case of the six Euclidean algorithms, they are even explicit. However, in the remaining cases, they do not seem be related to classical constants of analysis.

Results and plan of the paper. In Section 2, we present the ten algorithms to be analyzed, then, in Section 3, the relevant dynamical systems. Sections 4 and 5 are the central technical sections of the paper. There, we develop the line of attack outlined earlier and introduce successively Dirichlet generating functions, transfer operators of the Ruelle type, and the basic elements of Tauberian theory that are adequate for our purposes. The main results of these sections are summarized in Theorem 1. Theorems 2 and 4 of Section 6 imply a general criterion for logarithmic versus log-squared behaviour, while providing a framework for higher moment analyses.

Concerning our ten favorite algorithms—six Euclidean algorithms and four pseudo-Euclidean algorithms—, the corresponding analyses are summarized in Theorems 3 and 5 where we list our main results, some old and some new, that fall as natural consequences of the present framework. It results from the analysis (Theorem 3) that the Fast Class contains three Euclidean algorithms, the Classical Algorithm (G), the Centered Algorithm (K), and the Odd Algorithm (O) together with three pseudo-Euclidean algorithms, the Pseudo-Classical Algorithm $(\widehat{G})$, the Pseudo-Centered Algorithm $(\widehat{K})$ and the Pseudo-Odd Algorithm $(\widehat{O})$. Their respective average-case complexities on pairs of integers less than N are of the form

$$H_N \sim A_H \log N, \quad \text{for} \qquad H \in \{G, K, O, \widehat{G}, \widehat{K}, \widehat{O}\}.$$

The constants A_H are effectively characterized in terms of entropies of the associated dynamical system, and the constants related to the three Euclidean algorithms are easily obtained,

$$A_G = \frac{12 \log 2}{\pi^2}, \qquad A_K = \frac{12 \log \phi}{\pi^2}, \qquad A_O = \frac{18 \log \phi}{\pi^2}.$$

We also prove that concentration of distribution holds in the case of the fast algorithms, since, in this case, the moments of order k of the "number of iterations" (cost) $H_N^{[k]}$ are shown to be of the form

$$H_N^{[k]} \sim A_H^k \log^k N, \quad \text{for} \qquad H \in \{G, K, O, \widehat{G}, \widehat{K}, \widehat{O}\}.$$

Theorem 5 establishes that the Slow Class contains the remaining four algorithms, three Euclidean Algorithms, the By-Excess Algorithm (L), the Subtractive Algorithm (T), the Even Algorithm (E), and one of the pseudo-Euclidean algorithms, the pseudo-By-Excess Algorithm $(\widehat{L})$. These all have a complexity of the log-squared type,

$$H_N \sim B_H \log^2 N, \quad \text{for} \qquad H \in \{L, T, E, \widehat{L}\},$$

and the constants related to the three Euclidean algorithms are easily obtained, and related to the value $\zeta(2)$ of the Riemann Zeta function,

$$B_L = \frac{3}{\pi^2}, \qquad B_T = \frac{6}{\pi^2}, \qquad B_E = \frac{2}{\pi^2}.$$

In the case of slow algorithms, the k–th moment of the cost function is of order N^{k-1} for $k \geq 2$, with a constant that involves integer values of the Riemann Zeta function $\zeta(k)$ and $\zeta(k+1)$. In particular the standard deviation is $\Theta(\sqrt{N})$.

In summary, apart from specific analyses, our main contributions are the following:

(a) We show how transfer operator method may be extended to cope with complex situations where the associated dynamical system may be either random or Markovian (or both!).

(b) An original feature in the context of analysis of algorithms is the encapsulation of the method of inducing (related to intermittency as evoked above).

(c) Our approch opens access to information on higher moments of the distribution of costs, which appears to be new.
(d) It is quite clear that a pseudo-Euclidean algorithm is faster (in the sense of number of iterations) than its associated Euclidean algorithm. Consequently, the pseudo-version of a fast algorithm is always in the Fast Class. However, the pseudo-version of a slow algorithm may either remain slow (this is the case for the By-Excess Algorithm) or become fast. The latter situation is illustrated by the Subtractive algorithm that gives rise to the Binary Algorithm, a member of the Fast Class.

2 Ten variations of the Euclidean algorithm

We present here the ten algorithms to be analysed; the first six are classical variants of the Euclidean Algorithm, while the last four are pseudo-Euclidean algorithms where powers of two are removed from the remainder whenever they occur. The latter algorithms are well-suited for computing the Jacobi Symbol.

2.1. Six Euclidean algorithms. There are two divisions between u and v ($v > u$), that produce a positive remainder r such that $0 \le r < u$: the classical division (by-default) of the form $v = cu + r$, and the division by-excess, of the form $v = cu - r$. The centered division between u and v ($v > u$), of the form $v = cu + \varepsilon r$, with $\varepsilon = \pm 1$ produces a positive remainder r such that $0 \le r < u/2$. The even division produces an even quotient while the odd division produces an odd quotient; they are of the form $v = cu + \varepsilon r$, with $\varepsilon = \pm 1, 0 \le r < u$, and c even (resp. odd). There are five Euclidean algorithms, each of them being associated to each type of division, respectively called the Classical Algorithm (G), the By-Excess Algorithm (L), the Centered Algorithm (K), the Even Algorithm (E) and the Odd Algorithm (O). Finally, the Subtractive Algorithm (T) uses only subtractions and no divisions, since it replaces the classical division $v = cu + r$ by exactly c subtractions of the form $v := u + r$.

2.2. Main properties of the Jacobi symbol. The Jacobi symbol, introduced by Jacobi in 1846 [22], is a very important tool in algebra, since it is related to quadratic characteristics of modular arithmetics. Interest in its efficient computation has been reawakened by its utilisation in primality tests [43] or more generally in cryptography. The Jacobi symbol intervenes in the definition of the Quadratic Residuality Problem, and many cryptographic primitives are based on the computational difficulty of this problem, such as the pseudo-random generator of Blum, Blum, Shub [5] or the probabilistic encryption scheme proposed by Goldswasser and Micali [16].
Let u and v be two integers. First, Legendre's symbol is defined for an odd prime number v as

$$(\frac{u}{v}) = \begin{cases} 0, & \text{if } u \equiv 0 \bmod v; \\ 1, & \text{if } v \text{ is a square modulo } v; \\ -1, & \text{if } v \text{ is not a square modulo } v. \end{cases}$$

Next, the Jacobi symbol $J(u, v)$ extends the Legendre symbol multiplicatively to the general case when v is any odd integer,

$$J(u, v) := \prod_{i \in I} (\frac{u}{v_i})^{e_i} \qquad \text{for} \quad v := \prod_{i \in I} v_i^{e_i} \qquad \text{with odd primes } v_i.$$

2.3. Pseudo-Euclidean algorithms. First, one must note that the Jacobi symbol can be directly computed from the classical Euclidean algorithm, thanks to a formula due to Hickerson [21], quoted in [49]. This formula involves the classical continued fraction expansion $(a_1, a_2, a_3, \ldots a_{2r})$ of even length of the rational u/v and the inverse w of u mod v (that can be also computed by the classical extended GCD algorithm), under the form

$$J(u, v) = 2v + \frac{1}{2}\,[1 - w - u + v \sum_{i=1}^{2r} (-1)^i a_i] \quad \bmod 4.$$

However, this formula is mostly of theoretical interest, since specific algorithms that we shall study run faster than the classical (extended) GCD algorithm. These algorithms are fundamentally based on the following properties,

Quadratic Reciprocity law: $J(u,v) = (-1)^{(u-1)(v-1)/4}\, J(v,u)$ for u, v odd positive integers,
Modulo law: $J(v,u) = J(v-bu, u)$, Multiplicativity law: $J(vw,u) = J(v,u)\, J(w,u)$,
Special values: $J(2,v) = (-1)^{(v^2-1)/8}$, $J(\varepsilon, u) = \varepsilon^{(u-1)/2}$ for $\varepsilon = \pm 1$,

and they all have a structure similar to the Euclidean algorithms, since they perform a sequence of successive Euclidean-like divisions and exchanges. However, the Quadratic Reciprocity law being only true for a pair of odd integers, the Euclidean division has to be changed to a pseudo-Euclidean division where pseudo-remainders will always be odd. A pseudo-Euclidean division on a pair of positive odd integers is thus a modification of the Euclidean division where powers of two are removed from the remainder r; then, the decomposition $r := 2^k s$ creates an odd integer s which is called the pseudo-remainder. Finally, a pseudo-Euclidean division on a pair (u,v) of positive odd integers

$$v = bu + \varepsilon 2^k s \qquad \text{with } \varepsilon = \pm 1,\ s \text{ odd},\ k \geq 0, \text{ and } 2^k s \text{ strictly less than } u \text{ (or than } \tfrac{u}{2}\text{)}, \tag{1}$$

creates another pair (s,u) for the following step. Then the Jacobi Symbol $J(u,v)$ is easily computed from the Jacobi symbol $J(s,u)$ by means of the following properties ($\delta(i,j)$ denotes the Kronecker symbol):

$$J(u,v) = (-1)^e\ J(s,u) \qquad \text{with } e = \tfrac{1}{4}(u-1)(v-1) + \tfrac{k}{8}(u^2-1) + \tfrac{1}{2}\delta(\varepsilon,-1)\,(u-1) \mod 2\ .$$

To each of the six Euclidean divisions that have been defined in 2.1, there corresponds a pseudo-Euclidean division, and thus a pseudo-Euclidean Algorithm. The pseudo-Euclidean algorithm associated to a Euclidean Algorithm (H) is denoted by $(\widehat{H})$. There are thus a priori six pseudo-Euclidean algorithms $(\widehat{G}), (\widehat{L}), (\widehat{K}), (\widehat{E}), (\widehat{O}), (\widehat{T})$. Since the Even division between two odd integers always produces an odd remainder, the Pseudo-Even Algorithm $(\widehat{E})$ coincides with the Even Algorithm (E). Note that what appears in the present context as the Pseudo-Subtractive Algorithm $(\widehat{T})$ is well-known since it coincides with the Binary GCD algorithm (B). The binary algorithm will not be studied in detail here since the analytic treatment requires additional developments.

2.4. The sets of linear fractional transformations. Instead of the integer pair (u,v), we consider the rational u/v; then, the division that expresses the pair (u,v) as a function of the following pair (r,u) is replaced by a linear fractional transformation h that expresses the rational u/v as a function of r/u. In the same way, the pseudo-division (1) is replaced by a linear fractional transformation h that expresses the rational u/v as a function of s/u. When performing ℓ Euclidean or pseudo-Euclidean divisions on the input (u,v), each of the ten algorithms builds a specific continued fraction of height ℓ for the rational u/v and decomposes the rational u/v as

$$\frac{u}{v} = h_1 \circ h_2 \circ \ldots \circ h_\ell(a).$$

Here, each h_i is a linear fractional transformation (LFT) and a is the stopping value of the rational. Note that the Subtractive Algorithm (T) uses up to two LFTs at each step, depending on whether the subtraction is followed by an exchange or not:

$$p(x) := \frac{x}{1+x}, \qquad q(x) := \frac{1}{1+x}.$$

Since the Even Algorithm (E) and the By-Excess Algorithms (L), $(\widehat{L})$ stop when the remainder r (or the pseudo-remainder s) equals u, the stopping value a equals 1 for these algorithms. It equals 0 for the other algorithms. For the Centered algorithms (K), $(\widehat{K})$, the rational x belongs to $\mathcal{I} = [0, 1/2]$. In the other cases, the rational x belongs to $\mathcal{I} = [0,1]$. Finally, the rational inputs of each algorithm belong to the basic interval

Alg., Type	*Division*	*Set of LFT's*	*Conditions on $\mathcal{J}$ or $\mathcal{F}$.*
(G) (1, all, 0)	$v = cu + r$ $0 \le r < u$	$\mathcal{G} = \{\frac{1}{c+x},\ c \ge 1\}$	$\mathcal{F} = \mathcal{G} \bigcap \{c \ge 2\}$
(L) (1, all, 1)	$v = cu - r$ $0 \le r < u$	$\mathcal{L} = \{\frac{1}{c-x},\ c \ge 2\}$	$\mathcal{F} = \mathcal{L} \bigcap \{c \ge 3\}$
(K) ($\frac{1}{2}$, all, 0)	$v = cu + \varepsilon r$ $\varepsilon = \pm 1, (c, \varepsilon) \ge (2, +1)$ $0 \le r < \frac{u}{2}$	$\mathcal{K} = \{\frac{1}{c+\varepsilon x}, \varepsilon = \pm 1,$ $(c, \varepsilon) \ge (2, +1)\}$	$\mathcal{F} = \mathcal{K} \bigcap \{\varepsilon = 1\}$
(E) (1, all, 1)	$v = cu + \varepsilon r$ c even, $\varepsilon = \pm 1,\ 0 < r < u$	$\mathcal{E} = \{\frac{1}{c+\varepsilon x}, c \text{ even}, \varepsilon = \pm 1\}$	$\mathcal{F} = \mathcal{E} \bigcap \{\varepsilon = 1\}$
(O) (1, all, 0)	$v = cu + \varepsilon r$ c odd, $\varepsilon = \pm 1, 0 \le r < u$	$\mathcal{O} = \{\frac{1}{c+\varepsilon x}, c \text{ odd}, \varepsilon = \pm 1,$ $(c, \varepsilon) \ge (1, 1)\}$	$\mathcal{F} = \mathcal{O} \bigcap \{c \ge 3, \varepsilon = 1\}$
(T) (1, all, 0)	$v = u + r$	$\mathcal{T} = \{q = \frac{1}{1+x}, p = \frac{x}{1+x}\}$	Finishes with pq

Figure 1: The six Euclidean algorithms.

Alg., Type	*Division*	*Set of LFT's*	*Conditions on $\mathcal{J}$ or $\mathcal{F}$.*
$(\widehat{O})$ (1, odd, 0)	$v = cu + \varepsilon 2^k s$ c odd, $\varepsilon = \pm 1$, s odd $k \ge 1, 0 \le 2^k s < u$	$\widehat{\mathcal{O}} = \{\frac{2^k}{c+\varepsilon x}, \varepsilon = \pm 1, k \ge 1, c \text{ odd},$ $(c, \varepsilon) \ge (2^k, +1)\}$	$\mathcal{J} = \mathcal{O}$
$(\widehat{G})$ (1, odd, 0)	$v = cu + 2^k s$ $s = 0$ or s odd, $k \ge 0$ $0 \le 2^k s < u$	$\widehat{\mathcal{G}}_0 = \mathcal{G}$ $\widehat{\mathcal{G}}_1 = \{\frac{2^k}{c+x}, k \ge 1, c \ge 2^k\}$ $\widehat{\mathcal{G}}_{i\vert j} = \widehat{\mathcal{G}}_j \cap \{c \equiv i \bmod 2\}$	$\mathcal{I} = \widehat{\mathcal{G}}_0$ $\mathcal{F} = \widehat{\mathcal{G}}_1$
$(\widehat{L})$ (1, odd, 1)	$v = cu - 2^k s$ s odd, $k \ge 0$ $0 \le 2^k s < u$	$\widehat{\mathcal{L}}_0 = \mathcal{L}$ $\widehat{\mathcal{L}}_1 = \{\frac{2^k}{c-x}, k \ge 1, c > 2^k\}$ $\widehat{\mathcal{L}}_{i\vert j} = \widehat{\mathcal{L}}_j \cap \{c \equiv i \bmod 2\}$	$\mathcal{I} = \widehat{\mathcal{L}}_0$ $\mathcal{F} = \widehat{\mathcal{L}}_1$
$(\widehat{K})$ ($\frac{1}{2}$, odd, 0)	$v = cu + \varepsilon 2^k s$ $s = 0$ or s odd, $k \ge 0$ $0 \le 2^k s < \frac{u}{2}$	$\widehat{\mathcal{K}}_0 = \mathcal{K}$ $\widehat{\mathcal{K}}_1 = \{\frac{2^k}{c+\varepsilon x}, k \ge 1,$ $\varepsilon = \pm 1, (c, \varepsilon) \ge (2^{k+1}, +1)\}$ $\widehat{\mathcal{K}}_{i\vert j} = \widehat{\mathcal{K}}_j \cap \{c \equiv i \bmod 2\}$	$\mathcal{I} = \widehat{\mathcal{K}}_0$ $\mathcal{F} = \widehat{\mathcal{K}}_1$

Figure 2: The four pseudo-Euclidean algorithms.

$\mathcal{I} = [0, \rho]$ with $\rho = 1$ or $\rho = 1/2$. For the Euclidean algorithms, the valid inputs are all the rationals of $\mathcal{I}$, while the valid inputs of the pseudo-Euclidean algorithms are only the odd rationals of $\mathcal{I}$. The variable valid has two possible values $\{\text{all}, \text{odd}\}$, and finally, the type of the algorithm is defined as the triple (ρ, valid, a).
The precise set of the LFT's that are used by the algorithm depends on the particular algorithm considered. It turns out that here are two classes of algorithms, the generic class and the Markovian class, as we now explain.
In the case of the six Euclidean algorithms and the Pseudo-Odd Algorithm $(\widehat{O})$, there may exist special sets of LFT's in the initial step $(\mathcal{J})$ or in the final step $(\mathcal{F})$. However, all the other steps are generic, in the sense that they use the same set of LFT's, that we call the generic set. These algorithms are called themselves generic.
On the contrary, the three pseudo-Euclidean algorithms $(\widehat{G})$, $(\widehat{L})$, $(\widehat{K})$ have a Markovian (i.e., finite-state) flavour. When applied to a pair (u, v) of odd integers, the pseudo-division (1) entails the following: if b is odd, then the remainder is even, and thus k satisfies $k \geq 1$; if b is even, then the remainder is odd, and thus k satisfies $k = 0$. This relation is of a Markovian type, and we consider two states: the 0–state, which means "the quotient of (v, u) is even" (or equivalently the remainder of (u, v) is odd), i.e., $k = 0$; the 1–state, which means "the quotient of (v, u) is odd" (or equivalently the remainder of (u, v) is even), i.e., $k \geq 1$. Denoting generally a Markovian algorithm by (U) and by $\mathcal{U}_j$ the set of LFT's which can be used in state j, we obtain four different sets, where $\mathcal{U}_{i|j}$ brings rationals from state j to state i. The initial state is always the 0 state and the final state is always the 1 state.
The last column of the arrays of Figures 1 and 2 describes the initial set $\mathcal{J}$ and the final set $\mathcal{F}$.

3 Euclidean dynamical systems.

In this section, we relate the ten variations of the Euclidean algorithm to continued fractions algorithms that can be viewed as continuous extensions of them. It proves fruitful to cast continued fractions algorithms in a dynamical systems framework. To this purpose, we present in this section the main tools that are used in the sequel, namely the Perron-Frobenius operator together with its generalized version, the Ruelle operator. The dynamical systems that are used here may be complex, since they are sometimes random, sometimes Markovian, and even sometimes both. Some of the dynamical systems give rise to chaotic behaviour, while the other ones present an intermittency phenomenon. In the latter case, we appeal to the method of induction explained below that will play an important rôle in the sequel.

3.1. Dynamical systems related to Continued Fraction Algorithms. A dynamical system (X, V) consists of a compact metric space together with a continuous map $V : X \to X$ which is called the shift mapping. Given an initial condition x in X, the sequence $(x, Vx, V^2x, \ldots)$ of iterates of x under the action of V forms the orbit (or the trajectory) of the initial point x. The main study in dynamical systems concerns itself with the interplay between properties of the transformation V and discrete properties of trajectories of points under iteration of the transformation. Here, Continued Fraction algorithms are important particular cases of what is usually called "piecewise analytic maps of the interval"

Definition [Piecewise analytic maps of the interval] *Let $\mathcal{I}$ be a real interval. A mapping $V : \mathcal{I} \to \mathcal{I}$ is piecewise analytic if there exists a (finite or denumerable) set $\mathcal{M}$, and a partition $\{\mathcal{I}_m\}_{m \in \mathcal{M}}$ of the interval $\mathcal{I}$ in subintervals $\mathcal{I}_m$ such that the function $x \to Vx$ maps analytically each $\mathcal{I}_m$ onto $\mathcal{I}$.*

A special rôle is played by the set $\mathcal{H}$ of branches of the inverse function V^{-1} of V that are also naturally numbered by the index set $\mathcal{M}$: we denote by $h_{[m]}$ the inverse of the restriction $V|_{\mathcal{I}_m}$, so that $\mathcal{I}_m$ is exactly the image $h_{[m]}(\mathcal{I})$. The set $\mathcal{H}^k$ is the set of the inverse branches of the iterate V^k; its elements are of the form $h_{[m_1]} \circ h_{[m_2]} \circ \cdots \circ h_{[m_k]}$ and are called the inverse branches of depth k. The set $\mathcal{H}^\star := \cup_k \mathcal{H}^k$ is the semi-group generated by $\mathcal{H}$.
The general framework of piecewise analytic maps of interval is very convenient for a discussion of continued

Alg.	*Function* $A(x)$	*Invariant function* ψ	*Entropy*	*Bad LFT* p	*Type of Class.*
(G)	Integer part of x	$\frac{1}{\log 2}\frac{1}{1+x}$	$\frac{\pi^2}{6\log 2}$	——-	Good
(L)	Smallest integer at least equal to x	$\frac{1}{1-x}$	Not defined.	$\frac{1}{2-x}$	Bad
(K)	Nearest integer to x	$\frac{1}{\log\phi}[\frac{1}{\phi+x}+\frac{1}{\phi^2-x}]$.	$\frac{\pi^2}{6\log\phi}$	——-	Good
(E)	Even integer nearest to x	$\frac{1}{1-x}+\frac{1}{1+x}$	Not defined	$\frac{1}{2-x}$	Bad
(O)	Odd integer nearest to x	$\frac{1}{\phi-1+x}+\frac{1}{\phi^2-x}$	$\frac{\pi^2}{9\log\phi}$	——-	Good

Figure 3: The five Euclidean dynamical systems. Here, the shift mapping V is defined from function A, with the formula $V(x) := |\frac{1}{x} - A(\frac{1}{x})|$.

Alg.	*Function* $V(x)$	*Invariant function* ψ	*Entropy*	*Bad LFT* p	*Type of Class.*
(T)	$V(x) := \begin{cases} \frac{x}{1-x} & \text{for } 0 \le x \le 1/2; \\ \frac{1-x}{x} & \text{for } 1/2 \le x \le 1 \end{cases}$	$\frac{1}{x}$	Not defined	$\frac{x}{1+x}$	Bad.

Figure 4: The dynamical system relative to the Subtractive Algorithm.

fraction algorithms that are extensions of the six Euclidean Algorithms. The interval $\mathcal{I}$ is then defined by the type of the algorithm (cf. Figure 1). In all the six cases, the shift mapping extends the map defined on rationals by the equality $V(u/v) = r/u$ where r is the remainder of the Euclidean division on (u, v). The shift mapping V relative to the first five Euclidean algorithms admits the common form

$$V(x) := |\frac{1}{x} - A(\frac{1}{x})|,$$

where the function A depends on the algorithm and is defined in Figure 3. The shift mapping used in the Subtractive continued Fraction Algorithm is defined in Figure 4.

These dynamical systems are represented in Figure 5. They are now well-known: The Classical Euclidean system was studied by Gauss himself, Lévy [30], Khinchin [24], Kuzmin [28], Wirsing [50] and Babenko [2], the Centered system by Rieger [35], the Even system by Schweiger, Bauer, Kraaicamp and Lopes [40, 3, 25], the Odd System by Schweiger [40].

The four continuous extensions of pseudo-Euclidean algorithms are also related to dynamical systems, but in a more complex way: the dynamical systems to which they are associated are random for the Pseudo-Odd Algorithm, and are both random and markovian for the last three algorithms. Randomness intervenes because the pseudo-divisions are defined with dyadic valuation, so that the continued fractions expansions are a priori defined only for rationals numbers. However, one can define random continued fraction for real numbers in these cases: The state 0 (that is attained with an even quotient) is deterministic, and in the state 1 (that is attained with an odd quotient and is the only state for the Pseudo-Odd Algorithm), one chooses at random the dyadic valuation k of a real number, according to the law $\Pr[k = d] = 2^{-d}$ (for $d \ge 1$), which extends the natural law on even integers. Then, the shift mapping $\widehat{V}$ of a pseudo-Euclidean Algorithm is

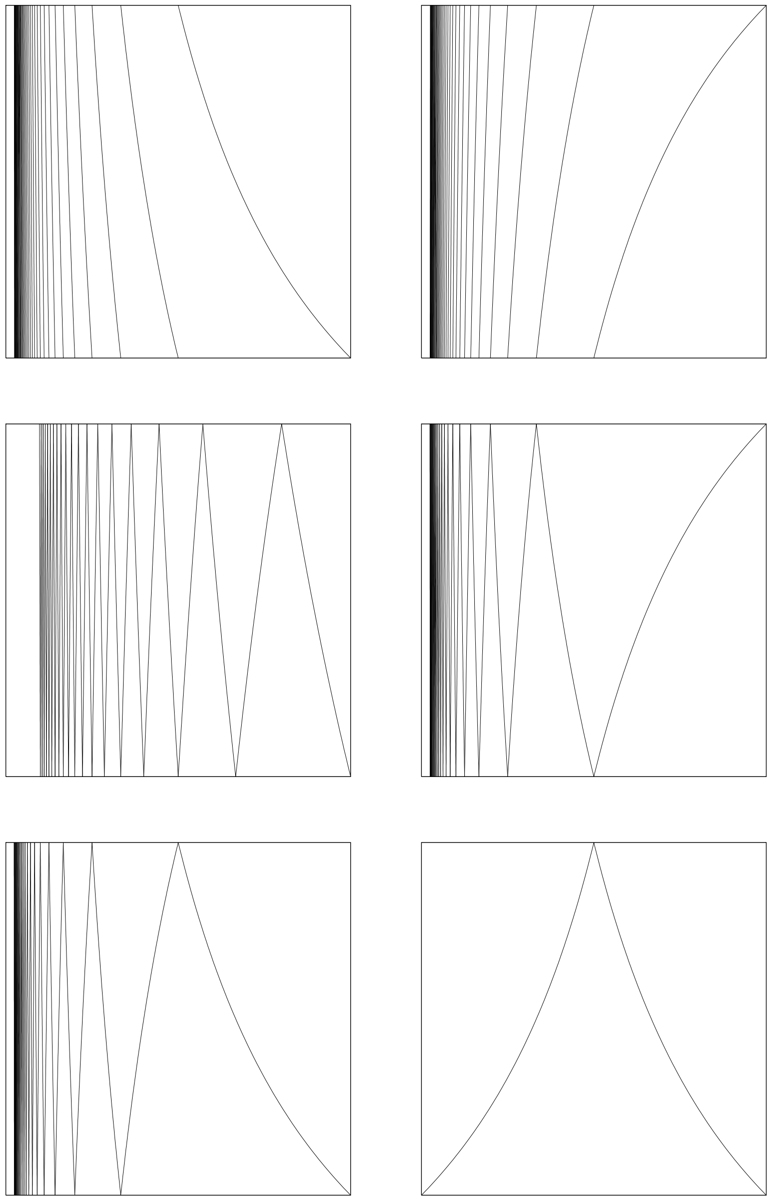

Figure 5: The six maps, from top left to bottom right: G, L, K, E, O, T.

defined from the shift mapping V of the related Euclidean Algorithm as:

$$\widehat{V}(x) = \begin{cases} V(x), & \text{in the state } 0; \\ V(\frac{x}{2^k}), & \text{in the state 1, with integer } k \text{ chosen with probability } 2^{-k} \ ; \end{cases}$$

If $V_{[k]}$ denotes the mapping defined by $V_{[k]}(x) := V(x/2^k)$, the set $\mathcal{H}_{[k]}$ of the inverse branches of $V_{[k]}$ is defined as

$$\mathcal{H}_{[k]} = \{g(x) := 2^k h(x) \qquad \text{with } h \in \mathcal{H} \text{ such that } \ h(\mathcal{I}) \subset [0, \frac{\rho}{2^k}] \ \},$$

and its elements are of determinant 2^k. The sets $\widehat{\mathcal{H}}_i$ of the LFTs used in the state i are defined as

$$\widehat{\mathcal{H}}_0 = \mathcal{H}_{[0]}; \qquad \widehat{\mathcal{H}}_1 = \bigcup_{k \geq 1} \mathcal{H}_{[k]}, \quad \text{the set } \mathcal{H}_{[k]} \text{ being chosen with probability } 2^{-k}.$$

Thus, all the Euclidean and pseudo-Euclidean algorithms can be extended into continuous dynamical systems where the rational inputs give rise to very particular trajectories that finish forever with $x = a$ (the final value a equals 0 or 1, according to the type of the algorithm defined in Figure 1 and 2).

3.2. The Perron-Frobenius operator and the Ruelle operator relative to a dynamical system. The behaviour of typical trajectories of dynamical systems is more easily explained by examining the flow of densities. If the set X is endowed with some initial distribution relative to some density $f = f_0$, the time evolution governed by the map V modifies the density, and the successive densities $f_1, f_2, \ldots, f_n, \ldots$ describe the global evolution of the system at time $t = 0$, $t = 1$, $t = 2$, ... Since the laws governing change do not change with time, there exists an operator $\mathbf{H}$ for which $f_1 = \mathbf{H}[f_0]$, $f_2 = \mathbf{H}[f_1]$, and more generally $f_n = \mathbf{H}[f_{n-1}] = \mathbf{H}^n[f_0]$ for all n. This operator is called the density transformer, or the Perron-Frobenius operator.

As previously, the set $\mathcal{H}$ denotes the set of inverse branches of a piecewise analytic map V of the interval $\mathcal{I}$. The Perron-Frobenius operator $\mathbf{H}$ relative to this dynamical system $(\mathcal{I}, V)$ is a functional operator defined by

$$\mathbf{H}[f](z) = \sum_{h \in \mathcal{H}} h'(z)\, f \circ h(z). \tag{2}$$

In the case when the dynamical system is random, with a family $\mathcal{H}_{[d]}, d \geq 1$ of sets of inverse branches, each set $\mathcal{H}_{[d]}$ being chosen with probability 2^{-d}, the Perron-Frobenius operator $\mathbf{H}$ is defined by

$$\mathbf{H}[f](z) = \sum_{d \geq 1} \frac{1}{2^d} \sum_{h \in \mathcal{H}_{[d]}} h'(z)\, f \circ h(z). \tag{3}$$

When the system is deterministic, the set $\mathcal{H}$ of the inverse branches is formed with LFT's of determinant 1. When the system is random, the set $\mathcal{H}_{[d]}$ of the inverse branches is chosen with probability 2^{-d} and is formed with LFT's of determinant 2^d. Then, in both cases, the Perron-Frobenius operator admits the common alternative form

$$\mathbf{H}[f](z) = \sum_{h \in \mathcal{H}} \frac{1}{D[h](z)^2}\, f \circ h(z) \tag{4}$$

that involves the denominator function $D[h]$ of the LFT h defined by

$$D[h](x) := |cx + d| = \sqrt{\frac{|\det h|}{|h'(x)|}} \qquad \text{for} \qquad h(x) = \frac{ax + b}{cx + d} \qquad \text{with } a, b, c, d \text{ coprime integers.}$$

We are interested here in studying rational trajectories that are very special trajectories that finish forever at $x = a$. It appears that the study involves an extension of the density transformer, that depends on some parameter s that plays the same rôle as the temperature in dynamical systems. This extension of the Perron-Frobenius operator is called the Ruelle operator. For a generic dynamical system (deterministic or random) relative to a set $\mathcal{H}$ of inverse branches, it is defined as

$$\mathbf{H}_s[f](z) = \sum_{h\in\mathcal{H}} \frac{1}{D[h](z)^s}\, f\circ h(z), \tag{5}$$

and extends the Perron-Frobenius operator $\mathbf{H}$ defined in (2), via the equality $\mathbf{H}_2 = \mathbf{H}$.
In the case when the dynamical system is Markovian with two states 0 and 1, the set $\mathcal{U}_i$ denotes the set of LFTs used in the state i, and the set $\mathcal{U}_{i|j}$ "brings" the system from state j to state i. If $\mathbf{U}_{i|j}$, resp $\mathbf{U}_i$, denotes the Perron-Frobenius operator associated to set $\mathcal{U}_{i|j}$, resp. $\mathcal{U}_i$, the Perron-Frobenius operator $\mathbf{U}$ relative to this Markovian system is a "matrix operator"

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_{0|0} & \mathbf{U}_{0|1} \\ \mathbf{U}_{1|0} & \mathbf{U}_{1|1} \end{pmatrix} \tag{6}$$

that operates on pairs $f = (f^{[0]}, f^{[1]})$ of functions so that

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{U}[f] = \begin{pmatrix} \mathbf{U}_0 & \mathbf{U}_1 \end{pmatrix} [f] = \mathbf{U}_0[f^{[0]}] + \mathbf{U}_1[f^{[1]}]. \tag{7}$$

Here, if $f^{[j]}$ denotes the density at x and in the state j, then $\mathbf{U}_{i|j}[f^{[j]}]$ is exactly the part of the density $f^{[i]}(x)$ that "comes from" density $f^{[j]}$. As previously, the Ruelle operator $\mathbf{U}_s$ relative to this Markovian system is a "matrix operator"

$$\mathbf{U}_s = \begin{pmatrix} \mathbf{U}_{s,0|0} & \mathbf{U}_{s,0|1} \\ \mathbf{U}_{s,1|0} & \mathbf{U}_{s,1|1} \end{pmatrix} \tag{8}$$

that involves the Ruelle operators $\mathbf{U}_{s,i}, \mathbf{U}_{s,i|j}$ relative to sets $\mathcal{U}_{i|j}$ and $\mathcal{U}_i$. It operates on pairs $f = (f^{[0]}, f^{[1]})$ of functions and

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \mathbf{U}_s[f] = \begin{pmatrix} \mathbf{U}_{s,0} & \mathbf{U}_{s,1} \end{pmatrix} [f] = \mathbf{U}_{s,0}[f^{[0]}] + \mathbf{U}_{s,1}[f^{[1]}]. \tag{9}$$

3.3. First properties of the Ruelle operators. The Ruelle operators satisfy two main properties:
(a) The ℓ-th iterate of the Ruelle operator describes what happens during ℓ iterations of the process. The multiplicative property of denominator D,

$$D[h\circ g](x) = D[h](g(x))\, D[g](x) \tag{10}$$

is translated into a multiplicative property on Ruelle operators: when given two sets of LFT's, $\mathcal{L}$ and $\mathcal{K}$ and their Ruelle operators $\mathbf{K}_s, \mathbf{L}_s$, the set $\mathcal{LK}$ is formed of all $h\circ g$ with $h \in \mathcal{L}$ and $g \in \mathcal{K}$, and the Ruelle operator relative to the set $\mathcal{LK}$ is exactly the operator $\mathbf{K}_s \circ \mathbf{L}_s$. In particular, the ℓ-th iterate of $\mathbf{H}_s$ involves all the inverse branches of depth ℓ,

$$\mathbf{H}_s^{\ell}[f](z) = \sum_{h\in\mathcal{H}^{\ell}} \frac{1}{D[h](z)^s} f\circ h(z), \tag{11}$$

and the Ruelle operator relative to the semi-group $\mathcal{H}^\star := \cup_{k\geq 0}\mathcal{H}^k$ is exactly $\sum_{k\geq 0} \mathbf{H}_s^k = (I-\mathbf{H}_s)^{-1}$. It is the quasi-inverse of the Ruelle operator $\mathbf{H}_s$ associated to the set $\mathcal{H}$.
In the same vein, the ℓ-th iterate of $\mathbf{U}_s$ generates all the elements of $\mathcal{U}^{<\ell>}$, i.e., all the possible LFT's of height ℓ. More precisely, the coefficient of index (i,j) of the matrix $\mathbf{U}_s^{\ell}$ is the Ruelle operator relative to the set $\mathcal{U}_{i|j}^{<\ell>}$ that brings the system from state j to state i in ℓ steps.

(b) At $s=2$, The Ruelle operator is a density transformer. In the generic (determinitic) case, one has

$$\int_{\mathcal{I}} \mathbf{H}[f](x)dx = \sum_{h\in\mathcal{H}} \int_{\mathcal{I}} |h'(x)| f\circ h(x) dx = \sum_{h\in\mathcal{H}} \int_{h(\mathcal{I})} f(x)dx = \int_{\mathcal{I}} f(x)dx; \tag{12}$$

in the same vein, for the generic (random) case, one has

$$\int_{\mathcal{I}} \mathbf{H}[f](x)dx = \sum_{d\geq 1} \frac{1}{2^d} \sum_{h\in\mathcal{H}_{[d]}} \int_{\mathcal{I}} |h'(x)| f\circ h(x)dx = \sum_{d\geq 1} \frac{1}{2^d} \sum_{h\in\mathcal{H}[d]} \int_{h(\mathcal{I})} f(x)dx = \int_{\mathcal{I}} f(x)dx; \tag{13}$$

and, in the Markovian case, the relation, valid for $f = (f^{[0]}, f^{[1]})$,

$$(1 \quad 1)\, \mathbf{U}[f](x) = \sum_{h\in\mathcal{U}_0} D[h](x)^{-2} f^{[0]} \circ h(x) + \sum_{h\in\mathcal{U}_1} D[h](x)^{-2} f^{[1]} \circ h(x) \tag{14}$$

proves the equality

$$\int_{\mathcal{I}} (1 \quad 1)\, \mathbf{U}[f](x)dx = \int_{\mathcal{I}} [f^{[0]}(x) + f^{[1]}(x)]dx = \int_{\mathcal{I}} (1 \quad 1)\, [f](x)dx. \tag{15}$$

3.4. Two classes of Euclidean dynamical systems. The most interesting properties of a dynamical system are connected with the long time behaviour of the sequence $\{f_n\}$ of successive densities. The basic question is: Does the sequence have a limit? (and in which sense?). Suppose that it is the case. Then this limit is the asymptotic density of the dynamical system . It is thus an invariant function for the operator $\mathbf{H}$—i.e., satisfies $\mathbf{H}[\psi] = \psi$—and gives rise to an (asymptotic) measure μ that is invariant under the action of V.

It is thus important to determine the invariant functions of the Perron-Frobenius operators relative to the six Euclidean dynamical systems. The results given in Figures 3 and 4 are well-known and exhibit two different classes of dynamical systems. The first one is the so-called Good Class and is formed with the systems for which the invariant function of the Perron-Frobenius operator is integrable on the interval $\mathcal{I}$. It contains the Algorithms $(G), (K), (O)$. The second one is the so-called Bad Class and is formed with the systems for which the invariant function of the Perron-Frobenius operator is not integrable on the interval $\mathcal{I}$. It contains the Algorithms $(L), (E), (T)$.

The Good Class. For the dynamical systems relative to the algorithms $(G), (K), (O)$, the invariant functions of the Perron-Frobenius operator are integrable on the interval $\mathcal{I}$, and the associated measure μ is invariant under the action of V. At the same time, the typical trajectories exhibit a chaotic behaviour and are a sort of replica of the interval $\mathcal{I}$ endowed with the measure μ. [see Figure $6(a)$]. The rational trajectories appear to quickly attain their final value a.

The notion of entropy will play an important rôle in this case. The entropy $h(\mathcal{H})$ of a dynamical system relative to a piecewise analytic map of the interval $\mathcal{I}$ with a set $\mathcal{H}$ of inverse branches is defined as the limit, if it exists, of a quantity that involves measures u_g of intervals $g(\mathcal{I})$, (for $g \in \mathcal{H}^\star$)

$$h(\mathcal{H}) := \lim_{n\to\infty} \frac{-1}{n} \sum_{g\in\mathcal{H}^n} u_g \log u_g.$$

For the "good" algorithms $(G), (K), (O)$, when the asymptotic measure is finite and relative to an invariant function ψ, a classical formula due to Rohlin shows that the entropy is related to the asymptotic mean value of $\log|V'|$. The mapping shift of all the Euclidean algorithms satisfies $V'(t) = -1/t^2$, so that

$$h(\mathcal{H}) = E_\infty[\log|V'|] := \int_{\mathcal{I}} \log|V'(t)|\, \psi(t)dt = -2\int_{\mathcal{I}} \log|t|\, \psi(t)dt. \tag{16}$$

The following relation involves the derivative $\Delta\mathbf{H}$ of the Ruelle operator with respect to s, taken at $s = 2$,

$$\int_{\mathcal{I}} \Delta\mathbf{H}[f](t)dt = \int_{\mathcal{I}} \log|t|\, f(t)dt \qquad \text{with } \Delta\mathbf{H} := \frac{d}{ds}\mathbf{H}_{s|_{s=2}},$$

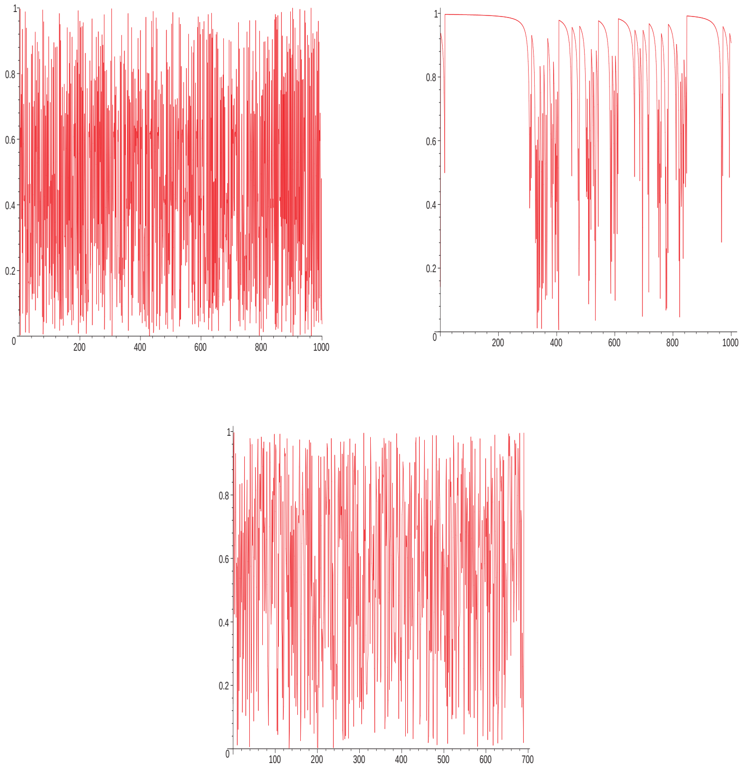

Figure 6: The iterates of $\pi - 3$ by the Classical continued fraction map (top left) and by the By-Excess map (top right); the trimmed version of the iterates of the By-Excess map (bottom).

and entails an alternative expression for the entropy

$$h(\mathcal{H}) = -2 \int_{\mathcal{I}} \mathbf{\Delta H}[\psi](t)dt.$$

The fourth column of Figure 3 gives the values of the entropy in the three cases.

The Bad Class. The situation is quite different for the other three Euclidean algorithms $(L), (E), (T)$. The invariant function is singular at the point a which is the final value for the algorithm; we recall that $a = 0$ for the Subtractive algorithm (T) and $a = 1$ for the By-Excess algorithm (L), or the Even algorithm (E). The reason is that this point a is indifferent under the action of V: it is a fixed point for V with a derivative equal to 1 (i. e. $V(a) = a, V'(a) = 1$). Then, the typical trajectory admits a quite different form, since, when it arrives near this point, it passes many times near it. [see Figure 6(b)]. As this point a is (weakly) repulsive, the rational trajectories that eventually finish at this point a, will attain it after a long time. Then it is (at least intuitively) clear that the relative Euclidean algorithms will be "slow".
However, the "induced" dynamical system which "forgets" all the sub-trajectories that stay near the indifferent point a admits typical trajectories that again exhibit a chaotic behaviour. [see Figure 6(c)]. Beginning with a dynamical system $(\mathcal{I}, V)$, the induced system $(\mathcal{I}, \widetilde{V})$ is defined in a more formal way as follows: If p denotes the "bad" inverse branch that contains the indifferent point a and $\mathcal{Q}$ denotes the set $\mathcal{H} \setminus p$ of good

Induced Alg.	*Invariant function* $\widetilde{\psi}$	*Entropy* $h(\widetilde{\mathcal{H}})$
$(\widetilde{L})$	$\frac{1}{\log 2}\frac{1}{2-x}$	$\frac{\pi^2}{3\log 2}$
$(\widetilde{E})$	$\frac{1}{\log 3}[\frac{1}{3-x}+\frac{1}{1+x}]$	$\frac{\pi^2}{2\log 3}$
$(\widetilde{T})= (G)$	$\frac{1}{\log 2}\frac{1}{1+x}$	$\frac{\pi^2}{6\log 2}$

Figure 7: The induced dynamical systems relative to the Bad Class.

inverse branches, the interval $\widetilde{\mathcal{I}}$ is defined as

$$\widetilde{\mathcal{I}} := \mathcal{I} \setminus p(\mathcal{I}) = \bigcup_{h\in\mathcal{Q}} h(\mathcal{I}).$$

The induced shift mapping $\widetilde{V} : \mathcal{I} \to \mathcal{I}$ is then defined from the first iterate of V that returns into $\widetilde{\mathcal{I}}$,

$$\widetilde{V}(x) := V^{n(x)+1}(x) \qquad \text{where } n(x) \text{ is the smallest integer } k \geq 0 \text{ such that } V^k(x) \in \widetilde{\mathcal{I}}.$$

This means that $\widetilde{V}(x)$ equals $V(x)$ for $x \in \widetilde{\mathcal{I}}$ while, for $x \notin \widetilde{\mathcal{I}}$, $\widetilde{V}(x)$ equals $V(y)$ where y is the first iterate of x that belongs to $\widetilde{\mathcal{I}}$. Then the trajectory $(x, \widetilde{V}x, \widetilde{V}^2x, \ldots)$ is exactly the trajectory $(x, Vx, V^2x, \ldots)$ which forgets all the sub-trajectories that stay near the indifferent point a, whereas the set $\widetilde{\mathcal{H}}$ of inverse branches is exactly the set $\widetilde{\mathcal{H}} = p^\star\mathcal{Q}$ where one groups a sequence of bad LFTs with a good one. The Ruelle operator $\widetilde{\mathbf{H}}_s$ of this dynamical system involves the Ruelle operators $\mathbf{Q}_s, \mathbf{P}_s$ relative to the sets $\{p\}, \mathcal{Q}$, under the form

$$\widetilde{\mathbf{H}}_s = \sum_{k\geq 0} \mathbf{Q}_s\ \mathbf{P}_s^k = \mathbf{Q}_s\ (I - \mathbf{P}_s)^{-1}, \qquad \text{so that} \qquad \widetilde{\mathbf{H}}_s \circ (I -\ \mathbf{P}_s) =\ \mathbf{Q}_s \qquad \text{and} \qquad \mathbf{P}_s + \mathbf{Q}_s = \mathbf{H}_s.$$

The operator $\mathbf{H}$ admits an invariant density denoted by ψ, that satisfies $\mathbf{H}[\psi] = \mathbf{P}[\psi] + \mathbf{Q}[\psi] = \psi$. Then, the function $g := \mathbf{Q}[\psi] = (I - \mathbf{P})[\psi]$ satisfies $\widetilde{\mathbf{H}}[g] = \widetilde{\mathbf{H}} \circ (I -\ \mathbf{P})[\psi] = \mathbf{Q}[\psi] = g$, so that $(I - \mathbf{P})[\psi]$ is an invariant function for $\widetilde{\mathbf{H}}$. With normalization condition, we obtain the expressions of $\widetilde{\psi}$, and, with Rohlin's formula, the expressions of the entropies of the induced dynamical system. The results are given in Figure 7. Remark that the induced system $(\widetilde{T})$ relative to the Subtractive algorithm (T) is exactly the Classical system (G).

3.5. The pseudo-Euclidean dynamical systems. These systems are defined in a more complex way than the Euclidean systems. It is not clear whether the invariant density $\widehat{\psi}$ of the pseudo-Euclidean system $(\widehat{H})$ is closely related to the invariant density ψ of the original Euclidean system (H). We do not succeed to obtain a closed form for the invariant density $\widehat{\psi}$ for the systems $(\widehat{G}), (\widehat{L}), (\widehat{O}), (\widehat{K})$. However, we will show in the sequel the following: If the original Euclidean algorithm belongs to the Good Class, it will be the same for its pseudo-version. This entails that the pseudo-Classical algorithm $(\widehat{G})$, the pseudo-Centered algorithm $(\widehat{K})$, and the pseudo-Odd algorithm $(\widehat{O})$ certainly belong to the Good Class. However, the pseudo-version of an algorithm that belongs to the Bad Class may remain in this class. This is the case for the pseudo-By-Excess algorithm $(\widehat{L})$. Since the bad LFT p is used in the 0 state, it is not modified and it remains bad. We will explain at the end why the situation is quite different for the pseudo-version $(\widehat{T})$ of the Subtractive algorithm that becomes good: The main reason is that the bad LFT p is modified by the random dyadic valuation.

For the three "good" pseudo-Euclidean dynamical systems, the Rohlin formula can be adapted, and involves the invariant density $\widehat{\psi}$ and the derivative of the Ruelle operator with respect to s. In the generic (random) case, one has

$$h(\widehat{\mathcal{H}}) = -2\int_{\mathcal{I}} \Delta\widehat{\mathbf{H}}\,[\widehat{\psi}](t)dt \qquad \text{with } \Delta\widehat{\mathbf{H}} := \frac{d}{ds}\widehat{\mathbf{H}}_{s|_{s=2}}$$

so that

$$h(\widehat{\mathcal{H}}) = 2\log 2 - 2\int_{\mathcal{I}} \log|t|\,\widehat{\psi}(t)dt.$$

In the same vein, for a Markovian random dynamical system, one has

$$h(\widehat{\mathcal{U}}) = -2\int_{\mathcal{I}} (1\quad 1)\,\Delta\widehat{\mathbf{U}}\,[\widehat{\psi}](t)dt = 2\log 2 - 2\int_{\mathcal{I}} \log|t|\,[\widehat{\psi}^{[0]}(t) + \widehat{\psi}^{[1]}(t)]dt.$$

In the next section, we come back to the analysis of the ten algorithms: we are then interested in the behaviour of rational trajectories, that finish at point $x = a$. The behaviour of such trajectories—a priori not at all typical—will be quite different according as the Algorithm belongs to the Good Class or to the Bad Class. We shall show that the algorithms of the Good Class are fast, while the algorithms of the Bad Class are slow.

4 Generating functions, dynamical operators and Tauberian Theorems

Here, we describe the general tools for analysing algorithms of the Euclidean type We first introduce the generating functions relative to the depth of the continued fraction and we relate them to the Ruelle operator which plays here the rôle of a generating operator. This operator can be generic or Markovian, according to the structure of the algorithm. In this way, the Dirichlet series that intervene in the analysis, called $F(s)$, $G(s)$, and more generally $G_k(s)$ are expressed in terms of the Ruelle operator. The average number of steps, and, more generally, the moments of order k of the variable "number of steps" involve partial sums of coefficients of these Dirichlet series, and Tauberian Theorems are a classical tool that transfers analytical properties of Dirichlet series into asymptotic behaviour of their coefficients.

4.1. Generating functions. The basic interval is $\mathcal{I} := [0, \rho]$, where ρ has two possible values $\rho = 1$ or $\rho = 1/2$. We consider the following sets

$$\bar{\Omega} := \{(u,v);\ u, v \text{ valid, } \frac{u}{v} \in \mathcal{I}\},\quad \Omega := \{(u,v);\ u, v \text{ valid}, \gcd(u,v) = 1, \frac{u}{v} \in \mathcal{I}\},$$

$$\bar{\Omega}_N := \{(u,v) \in \bar{\Omega}, v \le N\}, \qquad \Omega_N := \{(u,v) \in \Omega, v \le N\},$$

for the possible inputs of an algorithm, and we denote by $\Omega^{[\ell]}$, $\bar{\Omega}^{[\ell]}$, $\Omega_N^{[\ell]}$, $\bar{\Omega}_N^{[\ell]}$ the subsets of Ω, $\bar{\Omega}$, Ω_N, $\bar{\Omega}_N$ for which the algorithm performs exactly ℓ pseudo-divisions. Equivalently, the depth $X(u,v)$ of the continued fraction of u/v is equal to ℓ. We study the random variable X, and more precisely, its average values on Ω_N and on $\bar{\Omega}_N$, respectively denoted by $E_N[X]$ and $\bar{E}_N[X]$. More generally, we are interested in the distribution of X, and we also study its moments of order k, denoted by $E_N[X^k]$ and $\bar{E}_N[X^k]$. These quantities satisfy

$$E_N[X] := \frac{1}{|\Omega_N|}\sum_{\ell\ge 0} \ell\ |\Omega_N^{[\ell]}| \qquad \bar{E}_N[X] := \frac{1}{|\bar{\Omega}_N|}\sum_{\ell\ge 0} \ell\ |\bar{\Omega}_N^{[\ell]}| \tag{17}$$

$$E_N[X^k] := \frac{1}{|\Omega_N|}\sum_{\ell\ge 0} \ell^k\ |\Omega_N^{[\ell]}| \qquad \bar{E}_N[X^k] := \frac{1}{|\bar{\Omega}_N|}\sum_{\ell\ge 0} \ell^k\ |\bar{\Omega}_N^{[\ell]}| \tag{18}$$

and we wish to evaluate their asymptotic behaviour (for $N \to \infty$). We first consider pairs (u, v) with fixed $v = n$, and we denote by $\nu_n^{[\ell]}$ (resp. $\bar{\nu}_n^{[\ell]}$) the number of such elements of $\Omega^{[\ell]}$ (resp. $\bar{\Omega}^{[\ell]}$). We introduce the double generating functions $S(s,w)$ and $\bar{S}(s,w)$ of the sequences $(\nu_n^{[\ell]})$ and $(\bar{\nu}_n^{[\ell]})$,

$$S(s,w) := \sum_{\ell \geq 1} w^\ell \sum_{n > 1} \frac{\nu_n^{[\ell]}}{n^s}, \qquad \bar{S}(s,w) := \sum_{\ell \geq 1} w^\ell \sum_{n > 1} \frac{\bar{\nu}_n^{[\ell]}}{n^s}. \tag{19}$$

Since an algorithm performs the same steps on two pairs (u, v) and (ru, rv) for any valid integer r, the Riemann series $\bar{\zeta}$ relative to valid numbers $\bar{\zeta}(s) := \sum_{v \text{ valid}} v^{-s}$ that satisfies

$$\bar{\zeta}(s) = \zeta(s) \qquad \text{if valid= all} \quad \text{and} \qquad \bar{\zeta}(s) = (1 - \frac{1}{2^s})\zeta(s) \qquad \text{if valid= odd} \tag{20}$$

relates the two generating functions via the equality $\bar{S}(s,w) = \bar{\zeta}(s) S(s,w)$. It is then sufficient to study $S(s,w)$.

We introduce the two sequences (a_n) and (b_n), and more generally the sequences $(a_n^{[k]})$ (for $k \geq 0$),

$$a_n := \sum_{\ell \geq 1} \nu_n^{[\ell]} \qquad b_n := \sum_{\ell \geq 1} \ell \ \nu_n^{[\ell]}, \qquad a_n^{[k]} := \sum_{\ell \geq 1} \ell^k \ \nu_n^{[\ell]}, \tag{21}$$

that intervene, via partial sums of their coefficients, in the evaluation of the average $E_N[X]$,

$$E_N[X] = \frac{\sum_{n \leq N} \sum_{\ell \geq 0} \ell \ \nu_n^{[\ell]}}{\sum_{n \leq N} \sum_{\ell \geq 0} \nu_n^{[\ell]}} = \frac{\sum_{n \leq N} b_n}{\sum_{n \leq N} a_n}, \qquad E_N[X^k] = \frac{\sum_{n \leq N} \sum_{\ell \geq 0} \ell^k \ \nu_n^{[\ell]}}{\sum_{n \leq N} \sum_{\ell \geq 0} \nu_n^{[\ell]}} = \frac{\sum_{n \leq N} a_n^{[k]}}{\sum_{n \leq N} a_n}. \tag{22}$$

The associated Dirichlet series

$$F(s) = \sum_{n > 1} \frac{a_n}{n^s}, \qquad G(s) = \sum_{n > 1} \frac{b_n}{n^s}, \qquad G_k(s) = \sum_{n > 1} \frac{a_n^{[k]}}{n^s}. \tag{23}$$

can be easily expressed in terms of $S(s,w)$,

$$F(s) = S(s, 1), \qquad G(s) = \frac{d}{dw} S(s,w)|_{w=1}, \tag{24}$$

and more generally, $k!\, G_k(s)$ is closely related to the k-th derivative of $S(s,w)$ with respect to w, taken at $w = 1$. Then, the evaluation of the average $E_N[X]$ and of the higher moments $E_N[X^k]$ of order k will involve partial sums of the coefficients of Dirichlet series $F(s), G(s), G_k(s)$.

4.2. Tauberian Theorems. We have shown that the average number of steps $E_N[X]$ or the higher moments of the cost $E_N[X^k]$ of the ten algorithms on Ω_N is a ratio where the numerators and the denominators involve the partial sums of the Dirichlet series $F(s), G(s), G_k(s)$. Thus, the asymptotic evaluation of $E_N[X], E_N[X^k]$ (for $N \to \infty$) is possible if we can apply the following Tauberian theorem [9], [44] to the Dirichlet series $F(s), G(s), G_k(s)$.

Tauberian Theorem. [Delange] *Let $F(s)$ be a Dirichlet series with non negative coefficients such that $F(s)$ converges for $\Re(s) > \sigma > 0$. Assume that*
(i) $F(s)$ is analytic on $\Re(s) = \sigma, s \neq \sigma$, and
(ii) for some $\gamma \geq 0$, one has $F(s) = A(s)(s - \sigma)^{-\gamma - 1} + C(s)$, where A, C are analytic at σ, with $A(\sigma) \neq 0$.
Then, as $N \to \infty$,

$$\sum_{n \leq N} a_n = \frac{A(\sigma)}{\sigma \Gamma(\gamma + 1)} \ N^\sigma \log^\gamma N \quad [1 + \varepsilon(N)], \quad \varepsilon(N) \to 0.$$

In the remainder of the paper, we shall show that the Tauberian Theorem applies to the Dirichlet series defined in (23) with $\sigma = 2$. For $G_0(s) := F(s)$, it applies with $\gamma = 0$. For $G_1(s) := G(s)$, it applies with $\gamma = 1$ or $\gamma = 2$. For the slow algorithms, γ equals 2, and the average number of steps will be of order $\log^2 N$. For the fast algorithms, γ equals 1, and this will prove the logarithmic behaviour of the average number of steps.

We first examine the case of function $F(s)$ defined in (23). The Tauberian Theorem applies to $F(s)$ and with $\sigma = 2$ and $\gamma = 0$. First, the Dirichlet series $F(s)$ is closely linked to the $\bar{\zeta}$ function defined in (20), in both cases (valid= all or valid= odd). In the first case, $F(s)$ is related to the classical Zeta function $\zeta(s)$ itself,

$$F(s) = \frac{\rho}{\zeta(s)} \sum_{v \geq 1} \frac{v-1}{v^s} = \rho[\frac{\zeta(s-1)}{\zeta(s)} - 1].$$

In the "odd" case,

$$F(s) = \frac{\rho}{2\bar{\zeta}(s)} \sum_{v \text{ odd}} \frac{v-1}{v^s} = \frac{\rho}{2}[\frac{\bar{\zeta}(s-1)}{\bar{\zeta}(s)} - 1] \qquad \text{with} \qquad \bar{\zeta}(s) = (1 - \frac{1}{2^s})\, \zeta(s).$$

Thus, from classical properties of the Riemann Zeta function, it is clear that the Tauberian Theorem applies to $F(s)$, with $\sigma = 2$ and $\gamma = 0$. More precisely, at $s = 2$, one has:

$$F(s) \simeq \frac{2\rho}{\pi^2}\frac{1}{s-2} \quad \text{if valid= odd} \quad \text{and} \quad F(s) \simeq \frac{6\rho}{\pi^2}\frac{1}{s-2} \quad \text{if valid= all.} \tag{25}$$

It is not a priori clear how to directly apply Tauberian Theorems to $G(s)$ and more generally to $G_k(s)$ with $k \geq 2$. In the following, we obtain alternative expressions for $S(s, w)$ from which the location and the nature of the singularities of $G(s), G_k(s)$ will become apparent. Our analysis involves the Ruelle operators that have been introduced in Section 2. 6. They play here the rôle of generating operators, since they generate themselves the generating function $S(s, w)$.

4.3. Ruelle operators and generating functions. We recall that the first seven algorithms are generic, since they use the same set $\mathcal{H}$ at each generic step. In this case, the ℓ-th iterate of the Ruelle operator $\mathbf{H}_s$ generates all the LFT's used in ℓ (generic) steps of the algorithm. The last three algorithms are Markovian. Here, each step has two possible states, 0 and 1 . Then, the ℓ-th iterate of the Ruelle matrix operator $\mathbf{U}_s$ generates all the elements of $\mathcal{U}^{<\ell>}$, i.e., all the possible LFT's of depth ℓ. More precisely, the coefficient of index (i, j) of the matrix $\mathbf{U}_s^\ell$ is the Ruelle operator relative to the set $\mathcal{U}_{i|j}^{<\ell>}$ that brings rationals from state j to state i in ℓ steps.

In both cases, the Ruelle operator is then a "generating" operator, and generating functions themselves can be easily expressed with the Ruelle operator:

Proposition 1. *The double generating function $S(s, w)$ of the sequence $(\nu_n^{[\ell]})$ can be expressed as a function of the Ruelle operators associated to the algorithm.*
In the generic case,

$$S(s, w) = w\mathbf{K}_s[1](a) + w^2\, \mathbf{F}_s \circ (I - w\mathbf{H}_s)^{-1} \circ \mathbf{J}_s[1](a).$$

Here, the Ruelle operator $\mathbf{H}_s$ is related to the set $\mathcal{H}$ used at each generic step, whereas $\mathbf{F}_s$ is related to the set $\mathcal{F}$ used at the final step, and $\mathbf{J}_s$ is related to the set $\mathcal{J}$ used at the initial step; The operator $\mathbf{K}_s$ is related to the set $\mathcal{K} := \mathcal{J} \cap \mathcal{F}$ and solely intervenes when the algorithm performs only one step; the value a is the final value of the rational u/v.

In the Markovian case, the Ruelle operator $\mathbf{U}_s$ *is a matrix operator, and*

$$S(s,w) = (\,0 \quad 1\,)\; w\mathbf{U}_s\,(I - w\mathbf{U}_s)^{-1}\,\begin{pmatrix}1\\0\end{pmatrix}(a).$$

Proof. *Generic case.* Consider an element (u,v) of $\Omega^{[\ell+2]}$ $(\ell \geq 0)$; since the algorithm performs $\ell+2$ divisions on this input, and finishes with the value a, there exists exactly one linear fractional transformation h in the set $\mathcal{J}\mathcal{H}^\ell\mathcal{F}$ such that $u/v = h(a)$. Then, from (10) and for any $\ell \geq 0$,

$$\mathbf{F}_s \circ \mathbf{H}_s^\ell \circ \mathbf{I}_s[f](a) = \sum_{h\in\mathcal{J}\mathcal{H}^\ell\mathcal{F}} \frac{1}{D[h](a)^s}\, f\circ h(a) = \sum_{n>1}\ \sum_{(u,n)\in\Omega^{[\ell+2]}} \frac{1}{n^s} f(\frac{u}{n}). \tag{26}$$

Markovian case. The algorithm begins in the 0 state and ends in the 1 state. When it performs ℓ steps on the input (u,v), it uses the set $\mathcal{U}_{1|0}^{<\ell>}$. Moreover, the last value of the rational u/v equals a. Then

$$(\,0 \quad 1\,)\,\mathbf{U}_s^\ell \begin{pmatrix}f\\0\end{pmatrix}(a) = \sum_{h\in\mathcal{U}_{1|0}^{<\ell>}} \frac{1}{D[h](a)^s}\, f\circ h(a) = \sum_{n>1}\ \sum_{(u,n)\in\Omega^{[\ell]}} \frac{1}{n^s} f(\frac{u}{n}). \tag{27}$$

In both cases, the result is obtained when choosing $f = 1$ or $f = (1,0)$, and summing expressions over all the possible values of ℓ. ∎

4.4. Induced Ruelle operators and generating functions. We consider here an algorithm (H) whose set $\mathcal{H}$ of inverse branches contains a "bad" LFT p with an indifferent point a (i.e., a fixed point where the absolute value of the derivative equals 1) which is also the last value of the algorithm. Then, since the LFT p does not belong to the final set $\mathcal{F}$, this final set $\mathcal{F}$ is a subset of $\mathcal{Q} := \mathcal{H}\setminus\{p\}$. We then use here the Ruelle operator $\widetilde{\mathbf{H}}_s$ relative to the induced set $\widetilde{\mathcal{H}} := p^\star\mathcal{Q}$, together with the Ruelle operator $\widetilde{\mathbf{F}}_s$ relative to the last induced set $\widetilde{\mathcal{F}} := p^\star\mathcal{F}$. They involve the Ruelle operators $\mathbf{Q}_s, \mathbf{P}_s, \mathbf{F}_s$ relative to the sets $\{p\}, \mathcal{Q}, \mathcal{F}$, under the form

$$\widetilde{\mathbf{H}}_s = \sum_{k\geq 0} \mathbf{Q}_s\ \mathbf{P}_s^k = \mathbf{Q}_s\ (I-\mathbf{P}_s)^{-1}, \qquad \widetilde{\mathbf{F}}_s = \sum_{k\geq 0} \mathbf{F}_s\ \mathbf{P}_s^k = \mathbf{F}_s\ (I-\mathbf{P}_s)^{-1}. \tag{28}$$

As Prellberg and Slawny suggest it in another context [34], we introduce two Ruelle operators that depend also on the variable w, as

$$\widetilde{\mathbf{H}}_{s,w} = \sum_{k\geq 0} w^{k+1}\ \mathbf{Q}_s\ \mathbf{P}_s^k, \qquad \widetilde{\mathbf{F}}_{s,w} = \sum_{k\geq 0} w^{k+1}\ \mathbf{F}_s\ \mathbf{P}_s^k, \tag{29}$$

In the Markovian case, the operators $\mathbf{Q}_s, \mathbf{P}_s$ are now matrix operators, the operator $\mathbf{F}_s$ relative to the final set coincides with $\mathbf{Q}_s$, and the matrix operators $\widetilde{\mathbf{U}}_s, \widetilde{\mathbf{U}}_{s,w}$ are defined in the same way as $\widetilde{\mathbf{H}}_s, \widetilde{\mathbf{H}}_{s,w}$.
in both cases (generic case or Markovian case), these (induced) Ruelle operators can be used to express the generating function $S(s,w)$.

Proposition 2. *The double generating function $S(s,w)$ relative to the parameter "number of steps" and the double generating function $\widetilde{S}(s,w)$ relative to the parameter "number of good steps" can be expressed as a function of the induced Ruelle operators associated to the algorithm defined in (28), (29).*
In the generic case,

$$S(s,w) = \widetilde{\mathbf{F}}_{s,w}\ (I-\widetilde{\mathbf{H}}_{s,w})^{-1}[1](a), \qquad \widetilde{S}(s,w) = w\widetilde{\mathbf{F}}_s\ (I-w\widetilde{\mathbf{H}}_s)^{-1}[1](a), \tag{30}$$

Here, the Ruelle operator $\widetilde{\mathbf{H}}_s$ is related to the set $\widetilde{\mathcal{H}}$ used at each generic (induced) step, whereas $\widetilde{\mathbf{F}}_s$ is related to the final (induced) set $\widetilde{\mathcal{F}}$.

In the Markovian case,

$$S(s,w) = (\,0 \quad 1\,)\ \widetilde{\mathbf{U}}_{s,w}\,(I-\widetilde{\mathbf{U}}_{s,w})^{-1}\,\begin{pmatrix}1\\0\end{pmatrix}(a), \qquad \widetilde{S}(s,w) = (\,0 \quad 1\,)\ w\,\widetilde{\mathbf{U}}_s\,(I-w\,\widetilde{\mathbf{U}}_s)^{-1}\,\begin{pmatrix}1\\0\end{pmatrix}(a).$$

In both cases, the value a is the final value of the rational u/v.

4.5. The quasi inverse of the Ruelle operators. The various generating functions $S(s,w)$ always involve a quasi-inverse of some Ruelle operator –basic or induced–. There are two different cases: the first form involves the quasi-inverse $(I-\mathbf{H}_s)^{-1}$ or $(I-\widetilde{\mathbf{H}}_s)^{-1}$, where $\mathbf{H}_s, \widetilde{\mathbf{H}}_s$ do not depend on parameter w, whereas some mixing between the variables s and w intervenes in the second form $(I-\mathbf{H}_{s,w})^{-1}$. In the first case, the derivatives of order k of $S(s,w)$ with respect of w will involve powers of the quasi–inverse under the form $(I-w\mathbf{H}_s)^{-k-1}\circ\mathbf{H}_s^k$. In the second case, the derivatives of $(I-\widetilde{\mathbf{H}}_{s,w})^{-1}$ also involve the successive derivatives of $\widetilde{\mathbf{H}}_{s,w}$ with respect to w. More precisely, the k-th derivative of $S(s,w)$ contains one term of the form

$$(I-\widetilde{\mathbf{H}}_{s,w})^{-1}\circ\frac{d^k}{dw^k}\widetilde{\mathbf{H}}_{s,w}\circ(I-\widetilde{\mathbf{H}}_{s,w})^{-1},$$

that will be called the main term, and all the other terms are products of the quasi-inverse $(I-\widetilde{\mathbf{H}}_{s,w})^{-1}$ with lower derivatives (with respect of w) of $\widetilde{\mathbf{H}}_{s,w}$.

When $w=1$, we are then interested in the singularities of the quasi-inverses $(I-\mathbf{H}_s)^{-1}$, $(I-\widetilde{\mathbf{H}}_s)^{-1}$ of the Ruelle operators –induced or basic–. When the induced operator is useful, we are also interested in the singularities of the operator

$$\mathbf{M}_s^{[k]} := \frac{d^k}{dw^k}\widetilde{\mathbf{H}}_{s,w|_{w=1}}.$$

5 Functional Analysis.

Here, we consider a dynamical system relative to some piecewise analytic map of the interval $\mathcal{I}$ and the following conditions $\mathcal{C}(\mathcal{H})$ on its set $\mathcal{H}$ of inverse branches. While the first condition (C_1) is only technical, the second one (C_2) expresses a very important "contracting property". These two conditions will entail that the quasi-inverse of the Ruelle operator fulfills all the properties that we need for applying the Tauberian Theorem.

Conditions $\mathcal{C}(\mathcal{H})$.

(C_1) *The set $\mathcal{H}$ is a set of LFTs with integer coefficients which contains, for some integer $A>0$ a subset*

$$\mathcal{D} := \{h \mid h(x) = A/(c+x) \quad \text{with integers } c\to\infty\}.$$

(C_2) *There exist an open disk $\mathcal{V}$ that contains $\mathcal{I}$, and a real $\alpha<2$ such that*

(i) every LFT $h\in\mathcal{H}$ has an analytic continuation on $\mathcal{V}$, and maps the closure $\bar{\mathcal{V}}$ of disk $\mathcal{V}$ inside $\mathcal{V}$;

(ii) For each $h\in\mathcal{H}$, there exists $\delta(h)<1$ for which the analytic continuation of the function $|h'|$, denoted by $\widetilde{h}$, satisfies $0<|\widetilde{h}(z)|\le\delta(h)$ for all $z\in\mathcal{V}$

(iii) the series $\displaystyle\sum_{h\in\mathcal{H}}\left|\frac{\delta(h)}{\det(h)}\right|^{s/2}$ *converges on the plane $\Re(s)>\alpha$.*

In the sequel, we consider two kinds of dynamical systems, the generic ones or the Markovian ones. By definition, a generic dynamical system satisfies condition $\mathcal{C}$ if its set $\mathcal{H}$ of inverse branches fulfills conditions $\mathcal{C}(\mathcal{H})$. In the same vein, a Markovian dynamical system with two states 0 and 1 satisfies condition $\mathcal{C}$ if each of its four sets $\mathcal{U}_{i|j}$ of inverse branches fulfills conditions $\mathcal{C}(\mathcal{U}_{i|j})$. In both cases, the dynamical system is said to be a $\mathcal{C}$–dynamical system. The sequel of the section is devoted to the proof of the following Theorem which relates analytical properties of the set of LFTs (in the complex plane) and analytical properties of the quasi-inverse of the Ruelle operators (as a function of parameter s).

Theorem 1. *Consider a $\mathcal{C}$–generic dynamical system, defined from a piecewise analytic map of the interval $\mathcal{I}$ whose set of inverse branches is $\mathcal{H}$. Let $\mathbf{H}_s$ be the Ruelle operator relative to this dynamical system. Then, the powers $(I-\mathbf{H}_s)^{-p}$ of the quasi-inverse are analytic on the punctured plane $\{\Re(s)=2, s\neq 2\}$ and have a pole of order p at $s=2$. Near $s=2$, one has, for any function f positive on $\mathcal{V}\cap\mathbf{R}$, and any $x\in\mathcal{V}\cap\mathbf{R}$,*

$$(I-\mathbf{H}_s)^{-p}[f](x) \sim \frac{1}{(s-2)^p}\,(\frac{2}{h(\mathcal{H})})^p\,\psi(x)\int_0^1 f(x)dx, \tag{31}$$

where $h(\mathcal{H})$ is the entropy of the dynamical system and ψ is the invariant function of the Perron-Frobenius operator $\mathbf{H}$ defined by the normalization condition $\int_0^1\psi(x)dx=1$.
Consider now a $\mathcal{C}$–Markovian dynamical system with two states O and 1, defined from two piecewise analytic maps of the interval $\mathcal{I}$ whose sets of inverse branches are $\mathcal{U}_{i|j},(i,j=0,1)$. Let $\mathbf{U}_s$ be the Ruelle operator relative to this dynamical system. Then, the powers $(I-\mathbf{U}_s)^{-p}$ of the quasi-inverse are analytic on the punctured plane $\{\Re(s)=2, s\neq 2\}$ and have a pole of order p at $s=2$. Near $s=2$, one has, for any pair $f=(f^{[0]},f^{[1]})$ of functions positive on $\mathcal{V}\cap\mathbf{R}$, and any $x\in\mathcal{V}\cap\mathbf{R}$,

$$(I-\mathbf{U}_s)^{-p}[f](x) \sim \frac{1}{(s-2)^p}\,(\frac{2}{h(\mathcal{H})})^p\,\psi(x)\int_0^1 [f^{[0]}(x)+f^{[1]}(x)]dx,$$

where where $h(\mathcal{H})$ is the entropy of the dynamical system and ψ is the invariant function of the Perron-Frobenius operator $\mathbf{U}$ defined by the normalization condition $\int_0^1[\psi^{[0]}(x)+\psi^{[1]}(x)]dx=1$.

We will prove now the following facts: on a convenient functional space, and for $s>\alpha$, the Ruelle operators are compact, and possess, for real values of parameter s, dominant spectral properties. When s is near the real axis, these operators have thus a spectral gap. Moreover, the spectral radius of the operators satisfies maximum properties along horizontal and vertical lines. At $s=2$, the Ruelle operators are density transformers, and their dominant spectral objects have special values.

5.1. Compacity. We consider the space $A_\infty(\mathcal{V})$ of all functions f that are holomorphic in the domain $\mathcal{V}$ and are continuous on the closure $\bar{\mathcal{V}}$. Endowed with the sup-norm,

$$||f|| = \sup\ \{|f(u)|;\ u\in\mathcal{V}\},$$

$A_\infty(\mathcal{V})$ is a Banach space. Under the contracting condition (C_2), the Ruelle operator $\mathbf{H}_s$ acts on $\mathcal{A}_\infty(\mathcal{V})$ for $\Re(s)>\sigma$ and the operator $\mathbf{U}_s$ acts on $\mathcal{A}_\infty(\mathcal{V})^2$ for $\Re(s)>\alpha$. Such operators are studied in an extensive way by several authors, in particular Shapiro [41], [42] who proves that the operators are compact. They have even stronger properties, since they are nuclear of order 0 (in the sense of Grothendieck [17], [18]). This means that most of the matrix calculus can be adapted to this infinite–dimensional case. In particular, the trace of such an operator is well-defined.

5.2. Positivity and dominant spectral properties. Furthermore, for real values of parameter s, Ruelle operators have dominant spectral properties:

Under conditions $\mathcal{C}$, and for real values of parameter $s>\alpha$, the Ruelle operators $\mathbf{H}_s,\mathbf{U}_s$ admit a unique dominant eigenvalue $\lambda(s)$ positive, analytic for $s>\alpha$, and a unique (under normalization) dominant eigenfunction denoted by ψ_s.

Proof. We follow the lines of Mayer's work [33] that we adapt in our context. Mayer himself uses a result due to Krasnoselsky [27].

A subset K of a real Banach space B is called a proper cone if (i) $\rho K \subset K$ for $\rho > 0$ and (ii) $K \cap -K = \{0\}$. A proper cone is called reproducing if $B = K - K$, i.e., every element g of B is a difference of two elements of K. A linear operator $\mathcal{L} : B \to B$ is positive with respect to K if $\mathcal{L}K \subset K$. A positive operator $\mathcal{L} : B \to B$ is u_0-positive, for some some u_0 in the interior K^* of K, if there exist, for every non-zero $f \in K$, an integer p and strictly positive reals α, β for which

$$\alpha\ u_0 \le \mathcal{L}^p[f] \le \beta\ u_0, \tag{32}$$

where the order is defined with respect to K. Here is the result that we shall use.

Positivity Theorem. *[Krasnoselsky] Any compact u_0-positive operator $\mathcal{L} : B \to B$ satisfies a Perron-Frobenius property: it has a unique eigenvector in K^* and the relative eigenvalue is simple, positive, and in absolute value strictly larger than the other eigenvalues of $\mathcal{L}$.*

We show how to apply this result, first in the generic case. For real s, $\mathbf{H}_s$ acts on the real Banach space $A_{\infty\mathbf{R}}(\mathcal{V})$ formed with elements f of $A_\infty(\mathcal{V})$ which are real on the real segment $\mathcal{J}$. We denote by A_+ the subset of $A_{\infty\mathbf{R}}(\mathcal{V})$ formed with elements f which are positive on the real segment $\mathcal{J}$. For real s, $\mathbf{H}_s$ acts on A_+, and A_+ is a cone, proper and reproducing. The interior of the cone, denoted by A_+^*, is formed with elements f of $A_\infty(\mathcal{V})$ which are strictly positive on the real segment $\mathcal{J}$. We define the function u_0 to be equal to the constant function 1, and we show now that the operator $\mathbf{H}_s$ is u_0–positive with respect to the cone A_+: the upper bound of (32) is clear. For the lower bound, suppose that there exist $f \in A_+$ and x in $\mathcal{J}$ for which $\mathbf{H}_s[f](x) = 0$. Then, f is zero at each point $h(x)$ associated to an element h of $\mathcal{H}$. This is true in particular for points $h(x)$ associated to set $\mathcal{D}$ of condition (C_3) which form a denumerable set of distincts points. Since f is analytic, then f is zero.

Then, we apply Krasnoselsky's theorem: since $\mathbf{H}_s : A_{\infty\mathbf{R}}(\mathcal{V}) \to A_{\infty\mathbf{R}}(\mathcal{V})$ is a compact u_0–positive operator with respect to the proper and reproducing cone A_+, the restriction of $\mathbf{H}_s$ to the real Banach space $A_{\infty\mathbf{R}}(\mathcal{V})$ has a unique positive dominant eigenvalue $\lambda(s)$ strictly positive. One can choose the dominant eigenvector ψ_s in the cone A_+^*, which means that ψ_s is strictly positive on $\mathcal{J}$. Moreover, a direct calculus using the nuclearity (and the trace formula) shows that the spectra of the two operators, the operator $\mathbf{H}_s : A_\infty(\mathcal{V}) \to A_\infty(\mathcal{V})$ and its restriction to $A_{\infty\mathbf{R}}(\mathcal{V})$ are the same. Finally, the operator $\mathbf{H}_s : A_\infty(\mathcal{V}) \to A_\infty(\mathcal{V})$ has itself dominant spectral properties.

This ends the proof for the operator $\mathbf{H}_s$ in the generic case. This proof can be easily generalized to the Markov case. The real Banach space is then $A_{\infty\mathbf{R}}(\mathcal{V})^2$. The associated cone is $(A_+)^2$. For the lower bound of (32), suppose that there exist $f \in (A_+)^2$ and x in $\mathcal{J}$ for which $\mathbf{U}_s[f](x) = 0$. Here f has two components $f = (f^{[0]}, f^{[1]})$, and $f^{[i]}$ is zero at each point $h(x)$ associated to an element h of $\mathcal{U}_i$. This is true in particular for points $h(x)$ associated to set $\mathcal{D}_i$ which form a denumerable set of distincts points. Since $f^{[i]}$ is analytic, then $f^{[i]}$ is zero.

5.3. Spectral gap. The previous two properties –compacity, dominant spectral properties– entail the existence of a spectral gap that separates the dominant eigenvalue from the remainder of the spectrum.

Under conditions $\mathcal{C}$, and for a real number $\rho > \alpha$, there exists a neighborhood $\mathcal{R}$ of ρ, such that for any $s \in \mathcal{R}$, the operator $\mathbf{H}_s : \mathcal{A}_\infty(\mathcal{V}) \to \mathcal{A}_\infty(\mathcal{V})$ has a spectral gap. The same is true for the operator $\mathbf{U}_s : \mathcal{A}_\infty(\mathcal{V})^2 \to \mathcal{A}_\infty(\mathcal{V})^2$.

Proof. Since the operator $\mathbf{H}_s$ is compact, its spectrum is discrete with only an accumulation point at 0. Furthermore, for real $s = \rho$, it admits a unique dominant eigenvalue $\lambda(\rho)$. Then, this dominant eigenvalue is separated from the remainder of the spectrum by a gap, i.e., the supremum

$$\mu(\rho) := \ \mathrm{Sup}\{|\nu| \mid \nu \in \mathrm{Sp}\, \mathbf{H}_\rho,\ \nu \ne \lambda(\rho)\}$$

is strictly less than $\lambda(\rho)$. Thus, the operator $\mathbf{H}_\rho$ admits a spectral decomposition of the form $\lambda(\rho)\mathbf{R}_\rho + \mathbf{N}_\rho$, where $\mathbf{R}_\rho$ is the projection on the dominant eigensubspace, and $\mathbf{N}_\rho$ has a spectral radius equal to $\mu(\rho)$. More precisely, the operator $\mathbf{R}_s$ can be written as $\mathbf{R}_\rho[f](z) = e_\rho[f]\psi_\rho(z)$, where ψ_ρ is the dominant eigenvector of

$\mathbf{H}_\rho$ and e_ρ the projector on the dominant eigensubspace, with the normalization condition $e_\rho[\psi_\rho] = 1$. More generally, for any $\ell \geq 1$, and for $z \in \mathcal{V}$

$$\mathbf{H}_\rho^\ell[f](z) = \lambda(\rho)^\ell \mathbf{R}_\rho[f](z) + \mathbf{N}_\rho^\ell[f](z). \tag{33}$$

By Perturbation theory [23], the decomposition (33) extends for $\mathbf{H}_s$ to a (complex) neighborhood of $s = \rho$, the dominant eigenvalue $\lambda(s)$ is analytic there, and $\mathbf{N}_s$ has a spectral radius less than $\beta\lambda(s)$ (with $\beta < 1$). Then, one has, for all $\ell \geq 1$, and for $z \in \mathcal{V}$, the following holds

$$\mathbf{H}_s^\ell[f](z) = \lambda(s)^\ell \mathbf{R}_s[f] + \mathbf{N}_s^\ell[f](z). \tag{34}$$

and leads to

$$(I - \mathbf{H}_s)^{-1}[f](z) = \frac{\lambda(s)}{1-\lambda(s)} \psi_s(z)\ e_s[f] + \ (I - \mathbf{N}_s)^{-1}[f](z). \tag{35}$$

The same is true for the operator $\mathbf{U}_s$ in the Markov case, so that this part of the proof is common to both cases (generic and Markov).

5.4. Maximum properties on horizontal lines. *The dominant eigenvalue function $s \to \lambda(s)$ is strictly decreasing along the real line for $s > \alpha$.*

Proof. We consider both cases separately, first the generic case: When f is stricly positive on $\mathcal{J}$, the quantity $e_\rho[f]$ is strictly positive, and ψ_ρ is strictly positive on $\mathcal{J}$. Then, when taking $f = 1$ and $z = 0$ in (33), we obtain

$$\lambda(\rho) = \lim_{\ell\to\infty} \Big(\sum_{h \in \mathcal{H}^\ell} \frac{1}{D[h](0)^\rho} \Big)^{1/\ell}. \tag{36}$$

From contracting conditions (C_2), there exists $\gamma < 1$ such that

$$\text{Sup}\,\{\frac{1}{D[h](0)} \mid h \in \mathcal{H}\} = \gamma,$$

and then, by multiplicative properties (10) of D, and for any ℓ

$$\text{Sup}\,\{\frac{1}{D[h](0)} \mid h \in \mathcal{H}^\ell\} \leq \gamma^\ell.$$

Now, for $\alpha > 0$

$$\lambda(\rho+\alpha) = \lim_{\ell\to\infty} \Big(\sum_{h \in \mathcal{H}^\ell} \frac{1}{D[h](0)^{\rho+\alpha}} \Big)^{1/\ell} \leq \lim_{\ell\to\infty} \Big(\sum_{h \in \mathcal{H}^\ell} \gamma^{\alpha\ell}\ \frac{1}{D[h](0)^\rho} \Big)^{1/\ell} \leq \gamma^\alpha \lambda(\rho) < \lambda(\rho).$$

We easily adapt the previous proof to the Markov case: The relation

$$(1 \quad 1)\, \mathbf{U}_\rho^\ell \begin{pmatrix} 1 \\ 1 \end{pmatrix} (0) = \sum_{h \in \mathcal{U}^{<\ell>}} \frac{1}{D[h](0)^\rho},$$

that uses the set $\mathcal{U}^{<\ell>}$ of all possible LFTs of depth ℓ, proves that

$$\lambda(\rho) = \lim_{\ell\to\infty} \Big(\sum_{h \in \mathcal{U}^{<\ell>}} \frac{1}{D[h](0)^\rho} \Big)^{1/\ell}.$$

We now finish the proof in the same way as previously. ■

5.5. Maximum Properties along vertical lines. We describe now the behaviour of the spectrum of the Ruelle operator when parameter s varies on a vertical line:

Under conditions $\mathcal{C}$, for any real $\rho > \alpha$, for any s on the vertical line $\Re(s) = \rho, s \neq \rho$, the spectral radius of the operators $\mathbf{H}_s, \mathbf{U}_s$ is strictly less than $\lambda(\rho)$.

Proof. We adapt a proof due to Faivre [12]. **Generic case.** Let λ be an eigenvalue of $\mathbf{H}_s$ and f denote an eigenvector relative to λ. In the same way, the vector f_ρ denotes a dominant eigenvector relative to $\lambda(\rho)$. This function is strictly positive on the segment $\mathcal{J}$, non–zero on $\mathcal{V}$ and normalized by the condition $f_\rho(0) = 1$; Moreover, one can suppose that the function μ

$$\mu(x) := \frac{f(x)}{f_\rho(x)} \tag{37}$$

is of modulus at most 1 on $[0,1]$ and attains modulus 1 at point x_0. One always has

$$|\lambda\, f(x_0)| = |\mathbf{H}_s[f](x_0)| = |\sum_{h\in\mathcal{H}} D[h](x_0)^{-s}\ f\circ h(x_0)| \leq \sum_{h\in\mathcal{H}} D[h](x_0)^{-\rho}\ |f\circ h(x_0)| \tag{38}$$

$$\leq \sum_{h\in\mathcal{H}} D[h](x_0)^{-\rho}\ f_\rho\circ h(x_0) = \lambda(\rho) f_\rho(x_0), \tag{39}$$

and the definition of x_0 proves the inequality $|\lambda| \leq \lambda(\rho)$.
We suppose that, for $s = \rho + it, t \neq 0$, there exists $\lambda \in \mathrm{Sp}\,\mathbf{H}_s$ that satisfies $|\lambda| = \lambda(\rho)$. Then the sequence of inequalities (38), (39) becomes a sequence of equalities. For any $h \in \mathcal{H}$, the equality

$$|f\circ h(x_0)| \ = \ f_\rho\circ h(x_0) \tag{40}$$

holds. On the other side, the sequence $a_h := D[h](x_0)^{-\rho-it} f\circ h(x_0)$ satisfies the equality $|\sum a_h| = \sum |a_h|$. Then, there exists θ (of modulus 1) such that $a_h = \theta|a_h|$ for any h, and then the relation

$$f\circ h(x_0)\ D[h](x_0)^{-it} \ = \ \theta\ |f\circ h(x_0)| \tag{41}$$

holds. Both relations (40), (41) are in particular valid for the subset $\mathcal{D}$ of $\mathcal{H}$. Then, for $c \to \infty$, the sequence $h(x_0)$ tends to 0, and, equality (40) proves that $|f(0)| \ = \ f_\rho(0) \neq 0$. Now, for $c \to \infty$, Relation (41) shows that the sequence

$$D[h](x_0)^{-it} = (\frac{1}{c+x_0})^{it}$$

has a limit equal to θ, which can be only true for $t = 0$.
For the operator $\mathbf{U}_s$ in the Markov case, we consider the same objects: λ is an eigenvalue of $\mathbf{U}_s$ and $f = (f^{[0]}, f^{[1]})$ denotes an eigenvector relative to λ. In the same way, the vector $f_\rho = (f_\rho^{[0]}, f_\rho^{[1]})$ denotes a dominant eigenvector relative to $\lambda(\rho)$. This function has all its components strictly positive on the segment $\mathcal{J}$, non–zero on $\mathcal{V}$; moreover, one can suppose that the two functions μ_i

$$\mu_i(x) := \frac{f^{[i]}(x)}{f_\rho^{[i]}(x)} \tag{42}$$

are of modulus at most 1 on $[0,1]$, and the function μ_ℓ (for some $\ell \in \{0,1\}$) attains modulus 1 at point x_0. One always has

$$|\lambda\, f^{[\ell]}(x_0)| = \mathbf{U}_{s,\ell|0}[f^{[0]}](x_0) + \mathbf{U}_{s,\ell|1}[f^{[1]}](x_0) \leq \mathbf{U}_{\rho,\ell|0}[|f^{[0]}|](x_0) + \mathbf{U}_{\rho,\ell|1}[|f^{[1]}|](x_0) \tag{43}$$

$$\leq \mathbf{U}_{\rho,\ell|0}[f_\rho^{[0]}](x_0) + \mathbf{U}_{\rho,\ell|1}[f_\rho^{[1]}](x_0) = \lambda(\rho) f_\rho^{[\ell]}(x_0), \tag{44}$$

and the definition of point x_0 and index ℓ proves the inequality $|\lambda| \leq \lambda(\rho)$. Now, we suppose that, for $s = \rho + it, t \neq 0$, there exists $\lambda \in \operatorname{Sp} \mathbf{U}_s$ that satisfies $|\lambda| = \lambda(\rho)$. Then, the sequence of inequalities (43), (44) becomes a sequence of equalities. In particular, for any symbol $j = 0, 1$, we deduce the equalities

$$|f^{[j]} \circ h(x_0)| = f_\rho^{[j]} \circ h(x_0), \qquad \text{for } h \in \mathcal{U}_{\ell|j} \tag{45}$$

$$f^{[j]} \circ h(x_0)\ D[h](x_0)^{-it} \ = \ \theta\ |f^{[j]} \circ h(x_0)|, \qquad \text{for } h \in \mathcal{U}_{\ell|j} \tag{46}$$

that are in particular valid for the subset $\mathcal{D}_j$ of $\mathcal{U}_{\ell|j}$. Then, for $c \to \infty$, the sequence $h(x_0)$ tends to 0, and, equality (45) proves that $|f^{[j]}(0)| \ = \ f_\rho^{[j]}(0) \neq 0$. Now, for $c \to \infty$, Relation (46) shows that the sequence

$$D[h](x_0)^{-it} = (\frac{1}{c + x_0})^{it}$$

has a limit equal to θ, which can be only true for $t = 0$. ■

5.6. Some explicit values values at $s = 2$. *Under conditions $\mathcal{C}$, the Ruelle operators $\mathbf{H}_s, \mathbf{U}_s$ are at $s = 2$ density transformers; their dominant eigenvalue $\lambda(2)$ equals 1. Their dominant projector satisfies*

$$e_2[f] = \int_{\mathcal{I}} f(x)dx, \qquad \textit{or} \qquad e_2[f] = \int_{\mathcal{I}} [f^{[0]}(x) + f^{[1]}(x)]dx,$$

respectively in the generic and in the Markovian case. The entropy $h(\mathcal{H})$ of the dynamical system is well defined and is closely related with the derivative of the dominant eigenvalue function $s \to \lambda(s)$ at $s = 2$ via the equality $-2\lambda'(2) = h(\mathcal{H})$.

Proof. The alternative expression of Rohlin's formula for the entropy (16) involves the derivative $\Delta\mathbf{H}_s, \Delta\mathbf{U}_s$, of the Ruelle operator with respect to s, under the forms

$$h(\mathcal{H}) = -2 \int_{\mathcal{I}} \Delta\mathbf{H}[\psi](t)dt \qquad \text{or} \qquad h(\mathcal{H}) = -2 \int_{\mathcal{I}} (1\ \ 1)\ \Delta\mathbf{U}\ [\psi](t)dt \tag{47}$$

Taking the derivative (with respect to s) of relation $\mathbf{H}_s[\psi_s] = \lambda(s)\psi_s$ leads to

$$\Delta\mathbf{H}_s[\psi_s] + \mathbf{H}_s[\Delta\psi_s] = \lambda'(s)\psi_s + \lambda(s)\Delta\psi_s.$$

When choosing $s = 2$, taking the integrals on $\mathcal{I}$, and using the fact that $\mathbf{H}$ is a density transformer, Relation (47) leads to equality $-2\lambda'(2) = h(\mathcal{H})$. We obtain the same formula in the Markovian case.

5.7. End of the proof. The equality $\lambda(2) = 1$ of paragraph 5.6 together with maximum properties along vertical and horizontal lines of paragraphs 5.4 and 5.5 entail that the quasi-inverses $(I - \mathbf{H}_s)^{-1}, (I - \mathbf{U}_s)^{-1}$ are analytical on the punctured plane $\{\Re(s) > 2, s \neq 2\}$. Furthermore, the existence of a spectral gap between the dominant eigenvalue $\lambda(s)$ and the remainder of the spectrum (proven in paragraph 5.3) splits the quasi-inverse $(I - \mathbf{H}_s)^{-1}$ into two parts [see equation (35)]: the "part" relative to the dominant eigensubspace and the "part" relative to the remainder of the spectrum. On a (complex) neighborhood of $s = 2$, the spectral radius of $\mathbf{N}_s$ is strictly less than 1, and $(I - \mathbf{N}_s)^{-1}$ is analytic there. Now, using the derivability of $s \to \lambda(s)$ at $s = 2$ and the equality $\lambda(2) = 1$, the residues at $s = 2$ are easily evaluated from special values at $s = 2$ (cf paragraph 5.6). This ends the proof of Theorem 1.

6 Average-case Analysis of Euclidean Algorithms.

We come back now to the analysis of the ten algorithms. It appears that there will be only two possibilities. The first case arises when the algorithm is relative to a $\mathcal{C}$–dynamical system. Then, the average number of steps is of logarithmic order, and the algorithm is said to be fast. The second case arises when the dynamical system associated to the algorithm is not a $\mathcal{C}$–dynamical system. But, in this case, it appears that the induced dynamical system is a $\mathcal{C}$–dynamical system. The dynamical system itself is said to be a $\widetilde{\mathcal{C}}$–dynamical system. Then, the average number of steps is of log-squared order, and the algorithm is said to be slow.

We thus exhibit a general criterion for logarithmic versus log-squared behaviour that separates the algorithms to be studied in two Classes, the Fast Class, where the average number of steps is of logarithmic order, and the Slow Class, where the average number of steps is of log-squared order. It appears that this classification coincides with the previous classification between the Good Class and the Bad Class.

Finally, we shall obtain the following equalities between classes:

the Fast Class = the Class of the $\mathcal{C}$–dynamical systems = the Good Class;

the Slow Class = the Class of the $\widetilde{\mathcal{C}}$–dynamical systems = the Bad Class.

At the end, we come back to the study of the pseudo-version $(\widehat{T})$ of the Subtractive algorithm which is well-known as the Binary algorithm (B).

6.1. The Fast Class. This class contains the Euclidean Algorithms that are associated to $\mathcal{C}$–dynamical systems. The Fast Class coincides with the Good Class.

Theorem 2. *A Euclidean Algorithm associated to a $\mathcal{C}$–dynamical system (generic or Markovian) performs an average number of steps on valid rationals of $\mathcal{I}$ with denominator less than N that is asymptotically logarithmic,*

$$H_N \sim \frac{2}{h(\mathcal{H})} \log N$$

For any integer ℓ, the moment of order ℓ of the number of steps function is asymptotic to the ℓ–th power of the mean,

$$H_N^{[\ell]} \sim (\frac{2}{h(\mathcal{H})})^\ell \log^\ell N.$$

In particular the standard deviation is $o(\log N)$. Consequently the random variable expressing the number of steps satisfies the concentration of distribution property.

Proof. From Proposition 1, the main terms of Dirichlet series $F(s), G(s), G_k(s)$ –i.e., the terms that involve the higher powers of the quasi inverse of the Ruelle operators– admit the following expressions; in generic case,

$$F(s) \asymp \mathbf{F}_s \circ (I - \mathbf{H}_s)^{-1} \circ \mathbf{J}_s[1](a), \qquad G(s) \asymp \mathbf{F}_s \circ (I - \mathbf{H}_s)^{-2} \circ \mathbf{H}_s \circ \mathbf{J}_s[1](a),$$
$$G_k(s) \asymp \mathbf{F}_s \circ (I - \mathbf{H}_s)^{-k-1} \circ \mathbf{H}_s^k \circ \mathbf{J}_s[1](a).$$

and, in Markovian case,

$$F(s) \asymp (0 \quad 1)\ \mathbf{U}_s (I - \mathbf{U}_s)^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (a), \qquad G(s) \asymp (0 \quad 1)\ \mathbf{U}_s\ (I - \mathbf{U}_s)^{-2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (a).$$

$$G_k(s) \asymp (0 \quad 1)\ \mathbf{U}_s^k\ (I - \mathbf{U}_s)^{-k-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (a).$$

On the other side, Theorem 1 proves that the k-th power of the quasi-inverse of the Ruelle operator relative to a $\mathcal{C}$-dynamical system fulfills all the hypotheses of the Tauberian Theorem with $\sigma = 2$ and $\gamma = k$. ■

It is easy to verify that the generic sets $\mathcal{G}, \mathcal{K}, \mathcal{O}, \widehat{\mathcal{O}}$ relative to the Classical Algorithm, the Centered Algorithm, the Odd Algorithm, the pseudo-Odd Algorithm or the Markovian sets $\widehat{\mathcal{G}}, \widehat{\mathcal{K}}$ relative to the pseudo-versions of the the Classical or the Centered Algorithm fulfill conditions $\mathcal{C}$. Then, we obtain our first main result:

Theorem 3. *Consider the following six algorithms: the Classical Algorithm (G), the Classical Centered (K), the Odd Algorithm (O), the pseudo-Classical Algorithm $(\widehat{G})$, the pseudo-Centered Algorithm $(\widehat{K})$,* *and the pseudo-Odd algorithm* $(\widehat{O})$. *The average numbers of division steps performed by each of these five algorithms, on the set of valid inputs of denominator less than* N *are of asymptotic logarithmic order. They all satisfy*

$$H_N \sim \frac{2}{h(\mathcal{H})} \log N \qquad \text{for} \qquad H \in \{G, K, O, \widehat{G}, \widehat{K}, \widehat{O}\}.$$

For any integer ℓ, *the moment of order* ℓ *of the number of steps function is asymptotic to the* ℓ*–th power of the mean,*

$$H_N^{[k]} \sim (\frac{2}{h(\mathcal{H})})^k \ \log^k N.$$

Here $h(\mathcal{H})$ *is the entropy of the dynamical system relative to the algorithm. For the three Euclidean algorithms, the entropies are explicit,*

$$h(\mathcal{G}) = \frac{\pi^2}{6\log 2}, \qquad h(\mathcal{K}) = \frac{\pi^2}{6\log\phi}, \qquad h(\mathcal{O}) = \frac{\pi^2}{9\log\phi},$$

so that the three constants relative to these algorithms satisfy

$$A_G \approx 0.8428 \qquad A_K \approx 0.5851 \qquad A_O \approx 0.8777.$$

The entropies relative to the three pseudo-Euclidean Algorithms are not explicit, but they are computable numbers. The three constants relative to these pseudo-Euclidean Algorithms satisfy

$$A_{\widehat{G}} \approx 0.535 \pm 0.005, \qquad A_{\widehat{K}} \approx 0.430 \pm 0.005 \qquad A_{\widehat{O}} \approx 0.435 \pm 0.005.$$

Via Rohlin's formula, each of the six entropies admits an alternative form that involves the dominant eigenfunction ψ. Even if the dominant eigenfunction does not seem to be explicit for the pseudo-Euclidean algorithms, it can be efficiently computed, by adapting methods developed in previous papers [8, 46]. What we have at the moment is values from simulations that already provide a consistent picture of the relative merits of the pseudo-versions of the Classical, Centered and Odd Algorithms. It is to be noted that the computer algebra system MAPLE makes use of the pseudo-Classical Algorithm, (perhaps on the basis that only unsigned integers need to be manipulated), although this algorithm appears to be from our analysis the pseudo-Euclidean algorithm that has the *worst* convergence rate.

6.2. The Slow Class. This class contains the Euclidean Algorithms that are associated to $\widetilde{\mathcal{C}}$–dynamical systems. The Slow Class coincides with the Bad Class.

This class contains the algorithms for which the set $\mathcal{H}$ of inverse branches is only almost well-behaved. When the induced set $\widetilde{\mathcal{H}}$ fulfills conditions $\mathcal{C}(\widetilde{\mathcal{H}})$, we can adapt the previous methods to the induced system and we prove the following:

Theorem 4. *A Euclidean Algorithm associated to a* $\widetilde{\mathcal{C}}$*–dynamical system performs an average number of steps on valid rationals of* $\mathcal{I}$ *with denominator less than* N *that is asymptotically of log–squared type,*

$$H_N \sim \frac{1}{\bar\zeta(2)} \log^2 N.$$

For any integer $\ell \geq 2$, *the* ℓ*–th moment of total number of steps is*

$$H_N^{[\ell]} \sim \frac{\bar\zeta(2)}{2}(\frac{\bar\zeta(\ell)}{\bar\zeta(\ell+1)} - 1)^2 \ N^{\ell-1}$$

where $\bar\zeta(s)$ *is the Zeta function relative to valid numbers. In particular the standard deviation is* $\Theta(\sqrt{N})$.

The average number $\widetilde{H}_N$ of good steps performed by the Euclidean Algorithm on valid rationals of $\mathcal{I}$ with denominator less than N satisfies

$$\widetilde{H}_N \sim \frac{2}{h(\widetilde{\mathcal{H}})} \log N,$$

where $h(\widetilde{\mathcal{H}})$ is the entropy of the induced dynamical system $(\widetilde{H})$. The ℓ–th moment of the number of good steps is asymptotic to the ℓ–th power of the mean.

Proof. We begin with Proposition 2. When differentiating $S(s,w)$ with respect to w, and/or evaluating at $w=1$, one obtains expressions for the Dirichlet series $F(s)$, $G(s)$, $G_k(s)$. More precisely, their main terms now involve the operators

$$\widetilde{\mathbf{H}}_s := \widetilde{\mathbf{H}}_{s,1} \qquad \widetilde{\mathbf{F}}_s := \widetilde{\mathbf{F}}_{s,1} \qquad \mathbf{M}_s := \frac{d}{dw}\widetilde{\mathbf{H}}_{s,w|_{w=1}}, \qquad \mathbf{M}_s^{[k]} := \frac{d^k}{dw^k}\widetilde{\mathbf{H}}_{s,w|_{w=1}},$$

under the form

$$F(s) \asymp \widetilde{\mathbf{F}}_s\ (I-\widetilde{\mathbf{H}}_s)^{-1}[1](a), \qquad G(s) \asymp \widetilde{\mathbf{F}}_s \circ (I-\widetilde{\mathbf{H}}_s)^{-1} \circ \mathbf{M}_s \circ (I-\widetilde{\mathbf{H}}_s)^{-1}[1](a),$$

$$G_k(s) \asymp \widetilde{\mathbf{F}}_s \circ (I-\widetilde{\mathbf{H}}_s)^{-1} \circ \mathbf{M}_s^{[k]} \circ (I-\widetilde{\mathbf{H}}_s)^{-1}[1](a).$$

Now, since the induced set $\widetilde{\mathcal{H}}$ fulfills conditions $\mathcal{C}$, we can apply Theorem 1 to the induced dynamical system: The quasi-inverse of the induced Ruelle operator $\widetilde{\mathbf{H}}_s$ satisfies all the hypotheses of Tauberian Theorem. On the other side, the operator $\mathbf{M}_s$ has a simple pole at $s=2$. Then, properties of the powers of the quasi–inverse $(I-\widetilde{\mathbf{H}}_s)^{-1}$ can be transferred to $F(s)$ and $G(s)$, that fulfill the hypotheses of the Tauberian Theorem. They decompose as

$$F(s) = \frac{A}{s-2} + C(s) \qquad G(s) = \frac{B}{(s-2)^3} + D(s),$$

where $C(s)$ and $D(s)$ are analytic at $s=2$. A and B involve the spectral dominant objects of $\widetilde{\mathbf{H}}_s$, and the comparison between the two expressions of the residue of F respectively obtained in (25) and (31) gives an alternative expression of $B/(2A)$ in terms of $\bar{\zeta}(2)$. For the higher moments of order $\ell \geq 2$, the situation appears to be quite different, since $G_\ell(s)$ has now a dominant pôle at $s=\ell+1$. This pôle is brought by the operator $\mathbf{M}_s^{[\ell]}$, while the other operators that intervene in the expressions of are regular at $s=\ell+1$.

We can also study another important parameter which is the number of steps that use LFTs of $\mathcal{Q}$. We begin with Proposition 2. When differentiating $\widetilde{S}(s,w)$ with respect to w, and/or evaluating at $w=1$, one obtains expressions for the Dirichlet series $\widetilde{F}(s)$, $\widetilde{G}(s)$, $\widetilde{G}_k(s)$. More precisely, their main terms now only involve the operators $\widetilde{\mathbf{H}}_s, \widetilde{\mathbf{F}}_s$ under the form

$$\widetilde{F}(s) \asymp \widetilde{\mathbf{F}}_s\ (I-\widetilde{\mathbf{H}}_s)^{-1}[1](a), \qquad \widetilde{G}(s) \asymp \widetilde{\mathbf{F}}_s \circ (I-\widetilde{\mathbf{H}}_s)^{-2} \circ \widetilde{\mathbf{H}}_s[1](a),$$

$$G_k(s) \asymp \widetilde{\mathbf{F}}_s \circ (I-\widetilde{\mathbf{H}}_s)^{-k-1} \circ \widetilde{\mathbf{H}}_s^k\,[1](a).$$

Now, since the induced set $\widetilde{\mathcal{H}}$ fulfills conditions $\mathcal{C}$, we can apply Theorem 1 to the induced dynamical system. It entails that the quasi-inverse of the induced Ruelle operator $\widetilde{\mathbf{H}}_s$ satisfies all the hypotheses of Tauberian Theorem.∎

It is clear that the induced sets $\widetilde{\mathcal{E}}, \widetilde{\mathcal{L}}, \widetilde{\mathcal{T}}, \widetilde{\widehat{\mathcal{L}}}$ satisfy conditions $\mathcal{C}$. Then, we obtain our second main result:

Theorem 5. *Consider the four algorithms, the By-Excess Algorithm (L), the Subtractive Algorithm (T), the Even Algorithm (E), the pseudo-By-Excess algorithm $(\widehat{L})$. The average numbers of steps performed by each of the four algorithms, on the set of valid inputs of denominator less than N are of asymptotic log–squared order. They satisfy*

$$L_N \sim \frac{3}{\pi^2}\log^2 N, \qquad T_N \sim \frac{6}{\pi^2}\log^2 N, \qquad E_N \sim \frac{2}{\pi^2}\log^2 N.$$

The average numbers of good steps performed by the algorithms on the set of valid inputs of denominator less than N satisfy

$$\widetilde{L}_N \sim \frac{6 \log 2}{\pi^2} \log N \qquad \widetilde{T}_N \sim \frac{12 \log 2}{\pi^2} \log N, \qquad \widetilde{E}_N \sim \frac{4 \log 3}{\pi^2} \log N.$$

6.3. The Binary Algorithm. The Binary Algorithm (B) uses only subtractions and right shifts, since it performs a sequence of operations of the form $v := (v-u)/2^b$, where b is the dyadic valuation of $v-u$, denoted by $b := \mathrm{Val}_2\,(v-u)$, and defined as the largest exponent b such that 2^b divides $v-u$.

This algorithm has two nested loops: The external loop corresponds to an exchange. Between two exchanges, there is a sequence of iterations that constitutes the internal loop. It operates on odd-integer pairs.

Binary Euclidean Algorithm (u, v)

While $u \neq v$ **do**
  While $u < v$ **do**
    $b := \mathrm{Val}_2\,(v-u)$;
    $v := (v-u)/2^b$;
  Exchange u and v;

Output: u (or v).

Each internal step consists in subtractions and shifts and a sequence of internal steps can be written as

$$v = u + 2^{b_1} v_1, \qquad v_1 = u + 2^{b_2} v_2, \qquad v_2 = u + 2^{b_3} v_3, \qquad \ldots \qquad v_{\ell-1} = u + 2^{b_\ell} v_\ell.$$

Here v_ℓ is strictly less than u, and plays the rôle of a remainder r, so that a sequence of internal steps, followed by an exchange results in a decomposition of the form $v = mu + 2^k s$, with m odd, $m < 2^k$ and $s < u$, and the number of steps in the sequence equals $b(m)$, where $b(x)$ denotes the number of ones in the binary expansion of x.

It is clear that the algorithm (when it is viewed as a succession of internal steps) is exactly the pseudo-version $(\widehat{T})$ of the Subtractive Algorithm (T). As previously, the relative dynamical system is random, since the pseudo-division is related to dyadic valuation: we define random binary continued fraction for real numbers when choosing at random the dyadic valuation k of a real number in the same way as in 3.1. There is only one state, the state 1. The set $\widehat{\mathcal{T}}$ of LFTs used in each internal step is then the union of sets $\mathcal{T}_{[k]}$ for $k \geq 1$, and the set $\mathcal{T}_{[k]}$ that is chosen with probability 2^{-k} is formed with two LFTs of determinant 2^k,

$$\mathcal{T}_{[k]} = \{p_k(x) := \frac{x}{x + 2^k}, \qquad q_k(x) := \frac{1}{1 + 2^k x}\}.$$

We denote by $\mathcal{P}$ the set of all the p_k's and by $\mathcal{Q}$ the set of all the q_k's. Since the set $\widehat{\mathcal{T}}$ is not well-behaved, it is thus easier to use again the method of inducing. Then the global result of a sequence of internal steps followed by an exchange uses a LFT that belongs to the induced set

$$\mathcal{B} := \mathcal{P}^\star \mathcal{Q} = \{h(x) = \frac{1}{m + 2^s x}, \quad m \text{ odd}, \; m < 2^s\}.$$

Then, the Binary algorithm (B), when it is considered as a sequence of external loops, can be viewed exactly as the induced pseudo-version of the Subtractive algorithm, i.e.,

$$(B) = (\widetilde{\widehat{T}}).$$

Even if the induced set $\mathcal{B}$ has now a better behaviour than the set $\widehat{\mathcal{T}}$, the algorithm (B) is more difficult to analyze [7, 47] than the algorithms of this paper, because it is not possible to find an open disk whose diameter contains the basic interval $\mathcal{I} := [0,1]$ and on which all the LFT's are analytic. The reason is that the sequence of poles of LFT's is of the form $x = -m/2^s$ and has an accumulation point at $x = 0$. As in [48], we choose for $\mathcal{V}$ an open disk of diameter $]0, \beta[$ with $1 < \beta < 2$, and a convenient functional space is then the Hardy space of order two relative to $\mathcal{V}$. It is denoted by $\mathcal{H}^2(\mathcal{V})$ and is formed with all functions f analytic inside $\mathcal{V}$ and such that $|f|^2$ is integrable along the frontier of $\mathcal{V}$. The Ruelle operator $\mathbf{B}_s$ relative to set $\mathcal{B}$ acts on this functional space and is compact provided that $\Re(s) > (3/2)$. On this space, the main properties that we need for $\mathbf{B}_s$ are fulfilled, and this proves the logarithmic behaviour for the average number of steps.

Acknowledgements. I wish to thank Pierre Ducos for earlier discussions of 1992, Charlie Lemée for his master's thesis closely related to this work, Thomas Prellberg for his explanations of the inducing method, Dieter Mayer for his introduction to random dynamical systems, Ilan Vardi for Hickerson's formula, Philippe Flajolet for the experimentations and some nice figures, and Jérémie Bourdon for other nice figures and his help for the computation of higher moments.

References

[1] Akhavi, A., Vallée, B. Average bit–complexity of Euclidean Algorithms, 12p, submitted to ICALP'00.

[2] Babenko, K. I. On a problem of Gauss. *Soviet Mathematical Doklady 19*, 1 (1978), 136–140.

[3] Bauer, M. and Lopes, A. A billiard in the hyperbolic plane with decay of correlations of type n^{-2}, *Discrete and continuous dynamical systems* 3 (1997) pp 107-116.

[4] Bedford, T., Keane, M., and Series, C., Eds. *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, Oxford University Press, 1991.

[5] Blum, L., Blum, M. and Shub, M. A simple unpredictable random number generator, *SIAM J. on Comp.* 15 (1986) 364–363

[6] Bowen, R. Invariant measures for Markov maps of the interval, *Commun. Math. Phys.* 69 (1979) 1–17.

[7] Brent, R.P. Analysis of the binary Euclidean algorithm, Algorithms and Complexity, New directions and recent results, ed. by J.F. Traub, Academic Press 1976, pp 321–355

[8] Daudé, H., Flajolet, P., and Vallée, B. An average-case analysis of the Gaussian algorithm for lattice reduction, *Combinatorics, Probability and Computing* (1997) 6 397–433

[9] Delange, H. Généralisation du Théorème d'Ikehara, *Ann. Sc. ENS, (1954) 71, pp 213–242.*

[10] Dixon, J. D. The number of steps in the Euclidean algorithm, *Journal of Number Theory 2* (1970), 414–422.

[11] Eisenstein, G. Einfacher Algorithmus zur Bestimmung der Werthes von $(\frac{a}{b})$, *J. für die Reine und Angew. Math.* 27 (1944) 317-318.

[12] Faivre, C. Distribution of Lévy's constants for quadratic numbers, *Acta Arithmetica* 61, 1 (1992), 13–34.

[13] Flajolet, P. Analytic analysis of algorithms, In Proceedings of the 19th International Colloquium "Automata, Languages and Programming", Vienna, July 1992, W. Kuich, editor, Lecture Notes in Computer Science 623, pp 186–210

[14] Flajolet, P. and Sedgewick, R. Analytic Combinatorics, Book in preparation (1999), see also INRIA Research Reports 1888, 2026, 2376, 2956.

[15] Flajolet, P., and Vallée, B. Continued fraction Algorithms, Functional operators and Structure constants, *Theoretical Computer Science* 194 (1998), 1–34.

[16] GOLDWASSER, S. AND MICALI, S. Probabilistic encryption, *Journal of Computer and Systems Science* 28 (1984) 270–299.

[17] GROTHENDIECK, A. Produits tensoriels topologiques et espaces nucléaires, *Mem. Am. Math. Soc.* 16 (1955)

[18] GROTHENDIECK, A. La théorie de Fredholm, *Bull. Soc. Math. France* 84 pp 319-384.

[19] HEILBRONN, H. On the average length of a class of continued fractions, Number Theory and Analysis, ed. by P. Turan, New-York, Plenum, 1969, pp 87-96.

[20] HENSLEY, D. The number of steps in the Euclidean algorithm, *Journal of Number Theory* 49, 2 (1994), 142–182.

[21] HICKERSON, D. Continued fractions and density results for Dedekind sums, *J. Reine Angew. Math.* 290 (1977) 113-116.

[22] JACOBI, C.G.J. Uber die Kreistheilung und ihre Anwendung auf die Zalhentheorie, *J. für die Reine und Angew. Math.* 30 (1846) 166–182.

[23] KATO, T. *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[24] KHINCHIN, A. I. *Continued Fractions.* University of Chicago Press, Chicago, 1964. A translation of the Russian original published in 1935.

[25] KRAAIKAMP, C. AND LOPES, A. The Theta group and the continued fraction expansion with even partial quotients. preprint, 1995

[26] KNUTH, D.E. The art of Computer programming, Volume 2, 3rd edition, Addison Wesley, Reading, Massachussets, 1998.

[27] KRASNOSELSKY, M. *Positive solutions of operator equations,* P. Noordhoff, Groningen, 1964.

[28] KUZMIN, R. O. Sur un problème de Gauss, Atti del Congresso Internationale dei Matematici 6 (Bologna, 1928) pp 83-89.

[29] LEBESGUE V. A. Sur le symbole (a/b) et quelques unes de ses applications, *J. Math. Pures Appl.* 12 pp 497–517

[30] LÉVY, P. Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue. *Bull. Soc. Math. France* 57 (1929) pp 178-194

[31] LORCH, E. R. *Spectral Theory,* Oxford University Press, New York, 1962.

[32] MAYER, D. H. Continued fractions and related transformations, In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces*, T. Bedford, M. Keane, and C. Series, Eds. Oxford University Press, 1991, pp. 175–222.

[33] MAYER, D. H. Spectral properties of certain composition operators arising in statistical mechanics, *Commun. Math. Phys.* 68, 1-8 (1979)

[34] PRELLBERG, T. AND SLAWNY, J. Maps of intervals with Indifferent fixed points: Thermodynamic formalism and Phase transitions. *Journal of Statistical Physics* 66 (1992) 503-514

[35] RIEGER, G. J. Uber die mittlere Schrittazahl bei Divisionalgorithmen, *Math. Nachr.* (1978) pp 157-180

[36] RUELLE, D. *Thermodynamic formalism,* Addison Wesley (1978)

[37] RUELLE, D. *Dynamical Zeta Functions for Piecewise Monotone Maps of the Interval,* vol. 4 of *CRM Monograph Series*, American Mathematical Society, Providence, 1994.

[38] SHALLIT, J. On the worst–case of the three algorithmss for computing the Jacobi symbol, *Journal of Symbolic Computation* 10 (1990) 593–610.

[39] SHALLIT, J. Origins of the analysis of the Euclidean Algorithm, *Historia Mathematica* 21 (1994) pp 401-419

[40] SCHWEIGER, F. Continued fractions with odd and even partial quotients. Mathematisches Institut der Universitat Salzburg, Arbeitsbericht 2/1982

[41] SHAPIRO, J. *Composition operators and classical function theory*, Universitext: Tracts in Mathematics, Springer-Verlag, 1993.

[42] SHAPIRO, J Compact composition operators on spaces of boundary regular holomorphic functions, Proceedings of the AMS, 100, (1997), pp 49-57

[43] SOLOVAY, R. AND STRASSEN, V. A fast Monte-Carlo test for primality, *SIAM J. Comp.* 6 (1977) 84-85; erratum 7 (1978) 118.

[44] TENENBAUM, G. *Introduction à la théorie analytique des nombres*, vol. 13. Institut Élie Cartan, Nancy, France, 1990.

[45] VALLÉE, B. Opérateurs de Ruelle-Mayer généralisés et analyse des algorithmes d'Euclide et de Gauss, *Acta Arithmetica* 81.2 (1997) 101–144.

[46] VALLÉE, B. Fractions continues à contraintes périodiques, *Journal of Number Theory* 72 (1998) pp 183–235.

[47] VALLÉE, B. Dynamics of the Binary Euclidean Algorithm: Functional Analysis and Operators., *Algorithmica* (1998) vol 22 (4) pp 660–685.

[48] VALLÉE, B. A Unifying Framework for the analysis of a class of Euclidean Algorithms., To appear in the proceedings of LATIN'2000, LNCS

[49] VARDI, I. Continued fractions, Preprint, chapter of a book in preparation.

[50] WIRSING, E. On the theorem of Gauss–Kusmin–Lévy and a Frobenius–type theorem for function spaces. *Acta Arithmetica 24* (1974), 507–528.

[51] YAO, A.C., AND KNUTH, D.E. Analysis of the subtractive algorithm for greatest common divisors. *Proc. Nat. Acad. Sc. USA* 72 (1975) pp 4720-4722.