Average–case Analyses of three algorithms for computing the Jacobi Symbol.

Brigitte VALLÉE, Charly LEMÉE
GREYC, Université de Caen, F-14032 Caen (France)
[Brigitte.Vallee@info.unicaen.fr, clemee@etu.info.unicaen.fr]

October 1998

Abstract

We provide here a complete average–case analysis of the three algorithms for computing the Jacobi symbol, for positive odd integers less than N. We analyse the average number of steps used for each of the algorithms. The average values are shown to be asymptotic to $A_1 \log N$ or $A_2 \log N$ for two of them, whereas it is asymptotic to $A_3 \log^2 N$ for the third algorithm. The three constants A_i are related to the invariant measure of the Perron-Frobenius operator linked to the dynamical system. More precisely, they can be expressed with the entropy of the system.

1 Introduction.

The Jacobi symbol, introduced in [24], is a very important tool in algebra, since it is related to quadratic characteristics of modular arithmetics. Interest in its efficient computation is now reawakened with its utilisation in primality tests [40] 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 [4] or probabilistic encryption proposed by Goldswasser and Micali [15].

All the efficient algorithms that compute the Jacobi symbol have a similar structure to the GCD's algorithms: they perform a sequence of iterations, and each iteration uses a division and an exchange. For computing the Jacobi Symbol, instead of a classical euclidean division, one performs a slightly modified division, that is called a pseudo–euclidean division, whose cost is very close to the classical division. There are (at least) three kinds of pseudo–division that can be defined, and each of them creates an algorithm for computing the Jacobi Symbol. The analysis of each algorithm is thus devoted to evaluate the number of iterations that are performed during its execution.

Shallit [36] provided a complete worst-case analysis of all these algorithms, when the input is a pair of odd integers less than N. He proved that two of these algorithms –the Odd and the Ordinary Algorithms– perform, in the worst–case, $\Theta(\log N)$ iterations, whereas the last one (the Even Algorithm) performs $\Theta(N)$ iterations. Shallit precisely describes the worst–case configurations. At the end of his paper, Shallit wrote: "Bach has also suggested that one could investigate the average number of division steps in computing the Jacobi symbol .... This analysis is probably feasible to carry out for the Even Algorithm, and it seems likely that the average number of division steps is $O(\log^2 N)$. However, determining the average behaviour for the two other algorithms seems quite hard."

Here, we follow Bach's suggestion and we are interested in the average–case analysis of the algorithms, and, in particular, we wish to compare it to the average analysis of GCD's algorithms. The behaviour of the GCD algorithm is now well–understood. Heilbronn [18] and Dixon [9] have independently shown that the average number G_N of Euclidean divisions on a random pair of integers less than N is asymptotically logarithmic,

$$G_N \sim \frac{12 \log 2}{\pi^2} \log N.$$

Recently, the first author [46] analysed the Binary GCD algorithm, and she proved that the average number of iterations B_N of the Binary GCD algorithm on a random pair of odd integers less than N is also asymptotically logarithmic,

$$B_N \sim A_B \log N,$$

where A_B is a constant related to spectral properties of a functional operator linked to the dynamics of the algorithm. Note that this algorithm can be used for computing the Jacobi symbol.

We prove here that two of the algorithms for computing the Jacobi Symbol (the Odd Algorithm and the Ordinary Algorithm) have a similar behaviour to these algorithms: the average number of iterations O_N and U_N of the Odd algorithm and of Ordinary Algorithms on a random pair of odd integers less than N are also asymptotically logarithmic,

$$O_N \sim A_O \log N, \qquad U_N \sim A_U \log N.$$

The Even Algorithm behaves in a quite different way, since the average number of iterations E_N of the Even Algorithm on a random pair of odd integers less than N is asymptotically of order $\log^2 N$,

$$E_N \sim A_E \log^2 N,$$

so that the behaviour of the Even Algorithm is similar to the GCD substractive algorithm. We also compare the Even Algorithm to another GCD algorithm introduced by Ilan Vardi [47], so–called the non–Euclidean GCD algorithm. We show that these three algorithms (Even, Subtraction and Non-Euclidean) have almost the same structure, and we also explain why the Even algorithm is much slower than the Odd Algorithm or the Ordinary Algorithm.

The methods developed here use quite varied tools: generating functions, Ruelle operators, Tauberian methods, functional analysis. First, we use classical tools in the average–case analysis of algorithms [12], [13]: we introduce the generating functions related to the parameters to be analyzed; as is usual in the context of computational number theory, these generating functions are Dirichlet series. Second, we prove that these generating functions are closely linked to the some operators associated to the algorithms. These operators contain all the information on the dynamics of the algorithm. In the context of dynamical systems [34], they are called the Ruelle operators relative to the system. More precisely, the generating functions involve the quasi–inverse operator $(I - \mathbf{H}_s)^{-1}$, and the expectations to be studied are partial sums of coefficients of these Dirichlet series, so the main results of the paper will come from the application of Tauberian Theorems [8], provided that they *can* be applied. This will be the case as soon as the operator $\mathbf{H}_s$ when acting in a suitable Banach space has a "spectral gap", i.e. a unique dominant eigenvalue separated from the remainder of the spectrum by a gap. When acting on suitable spaces of holomorphic functions, the operator $\mathbf{H}_s$ is proven to be compact and positive (in the sense of Krasnoselsky) [28] for real values of parameter s, and then it has a spectral gap. Since Tauberian theorems link the asymptotics of coefficients to the dominant singularity of the function, the constants A_i involve the dominant singularity of the quasi–inverse $(I - \mathbf{H}_s)^{-1}$, i.e. the dominant spectral objects of the Ruelle operator $\mathbf{H}_s$. These dominant spectral objects are all explicit in the case of the Even Algorithm, and we show the equality $A_E = 1/\pi^2$. However, they do not seem be related to other classical constants in the case of the Odd and Ordinary Algorithm. This is already the case for the constant A_B that intervenes in the analysis of the Binary GCD algorithm.

These "dynamical" methods have been already used in other domains of algorithmic number theory, mainly in the study of the classical Euclidean algorithm, or related algorithms, for instance [32] [19] [20] [21] [22] [23] [42] [43] [44] [45] [46] [7]. In particular, the first author [14] uses them to obtain an alternative proof of the results of Heilbronn and Dixon about the logarithmic behaviour of the Classical Euclidean algorithm. Recently, she also provided a complete analysis of the Binary GCD algorithm with similar methods [46], and she has extended them in the general context of information theory [45], where she describes algorithmic properties of dynamical sources.

In this paper, our analyses confirm Bach's intuition, since we prove that the average number of divisions performed by the Even Algorithm is of order $O(\log^2 N)$. But, it is not true that the analyses of the Odd Algorithm or the Ordinary algorithm were harder than the analysis of the Even Algorithm. On the contrary, these two last analyses use a standard application of the methods that are previously described, whereas the analysis of the Even Algorithm is more involved, because the Ruelle operator associated to the Even Algorithm is not well–behaved. We must change our operator and use the method of inducing, classical in the theory of dynamical systems. It was introduced by Bowen [5] and well explained in a paper of Prellberg and Slawny [33].

Plan of the paper. In Section 2, we describe precisely the three algorithms, and the set of transformations that they use. Then, we introduce the generating functions and the Ruelle operators, and we exhibite their relation. We show how the Tauberian Theorems intervene, and explain how their application depend on spectral properties of the Ruelle operator. Section 4 is devoted to general results of functional analysis: we prove that, under some natural "contracting" conditions satisfied by the set of transformations, the Ruelle operator fullfills the hypotheses of Tauberian Theorems. Section 5 describes the applications of the general methods of Section 3 and 4 to each algorithm, first the Odd and the Ordinary Algorithm, and finally the Even Algorithm whose analysis is more involved.

2 The main algorithms.

Let u and v be two integers, with v an odd prime. Legendre's symbol is defined as follows:

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

With Legendre's symbol, one defines the Jacobi symbol which generalizes the Legendre symbol to the case when v is any odd integer. If v admits the decomposition

$$v = \prod_{i \in I} v_i^{e_i} \qquad \text{with odd primes } v_i,$$

then the Jacobi symbol $J(u, v)$ is defined as

$$J(u, v) := \prod_{i \in I} (\frac{u}{v_i})^{e_i}.$$

2.1. Main properties. The following properties play a central rôle in all the algorithms, since they make possible to compute the Jacobi symbol $J(u, v)$ even if the factorisation of v is unknown.

$$\text{Quadratic Reciprocity law:} \qquad J(u, v) = (-1)^{(u-1)(v-1)/4}\, J(v, u), \qquad \text{for } u, v \text{ odd positive integers;}$$

$$\text{Modulo law:} \qquad J(v, u) = J(v - bu, u);$$

$$\text{Multiplicativity law:} \qquad J(vw, u) = J(v, u)\, J(w, u);$$

$$\text{Special values:} \qquad J(2, v) = (-1)^{(v^2-1)/8}; \qquad J(\varepsilon, u) = \varepsilon^{(u-1)/2} \text{ for } \varepsilon = \pm 1;$$

All these algorithms are very similar to the Euclidean Algorithm, since they perform a sequence of successive euclidean divisions. But the Quadratic Reciprocity law is only true for two odd integers. So, one transforms the euclidean division into a pseudo–euclidean division where remainders are transformed into pseudo–remainders that will always be odd.

2.2. General principles. We begin with a pair (u, v) where u and v are positive and odd, $u < v$. We use a pseudo-euclidean division of the form

$$v = bu + \varepsilon 2^k s \qquad \text{with } \varepsilon = \pm 1,\ s \text{ odd and strictly less than } u$$

which gives another pair (s, u) for the following step. Then the Jacobi Symbol $J(u, v)$ is easily computed from the Jacobi symbol $J(s, u)$, when using successively the Quadratic Reciprocity law, the Modulo law, and the Multiplicativity law:

$$J(u, v) = (-1)^{(u-1)(v-1)/4}\, J(v, u)$$

$$J(v, u) = J(\varepsilon\, 2^k\, s, u) = J(\varepsilon, u)\ J(2, u)^k\, J(s, u).$$

Finally, with the special cases, one obtains:

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

(here, $\delta(i, j)$ denotes the Kronecker symbol).

There are three main different kinds of pseudo–euclidean divisions,
- divisions where all the pseudo–quotients are even,
- divisions where all the pseudo–quotients are odd,
- divisions where the pseudo–quotients are equal to the ordinary quotients: they may be even or odd.

There are thus three different algorithms according to the kind of pseudo–euclidean division that they perform and they are respectively called Even, Odd, Ordinary. We describe now each of these algorithms, that also be used to compute the GCD of the pair (u, v) that is denoted by $G(u, v)$.

We must note that the binary division, used in the Binary GCD algorithm, can also be used for computing the Jacobi symbol. However, it is of another kind since the pseudo–division that it uses is not a modification of the classical euclidean division. Anyway, the algorithm is already analysed. There are also other pseudo–euclidean divisions based on modifications of the centered euclidean division, in particular the Centered Ordinary division, that we do not analyse here. It is not more difficult, and our methods can be applied there.

2.3. The Even Algorithm. This algorithm is based on a pseudo–euclidean division where all the pseudo–quotients are even. This division is defined as follows: one first considers the ordinary euclidean division between two odd integers u, v with a non zero–remainder

$$v = au + r \qquad \text{with } 0 < r < u,$$

and one transforms it into a pseudo–euclidean division

$$v = bu + \varepsilon s \qquad \text{with } \varepsilon = \pm 1,\ s \text{ odd},\ 0 < s < u.$$

If the quotient a is even, then one lets

$$b := a, s := r, \varepsilon := 1,$$

else one lets

$$b := a + 1, s := u - r, \varepsilon = -1.$$

Since b is always even, $b = 2c$, then the pseudo–remainder s is always odd. As in the classical euclidean algorithm, the last division is non-trivial exact one, i.e. $v = au$ with $a > 1$ odd, that we transform into $v = (a-1)u + (+1)u$.

Even Algorithm.

Input: a pair (u, v) satisfying $0 < u < v$, u, v odd.

Output: $G(u, v)$ and $J(u, v)$

begin

$e := 0$;

While $u < v$ **do**

  begin

  Perform the even–euclidean division between v and u: $v = 2cu + \varepsilon s$;

  $e := e + \frac{1}{4}(u-1)(v-1) + \frac{1}{2}\delta(\varepsilon, 1)\,(u-1)$;

  $v := u; u := s$;

  end;

$G(u, v) := u$;

if $u = 1$ **then** $J(u, v) := (-1)^e$, **else** $J(u, v) := 0$.

end.

Instead of the integer pair (u, v), one considers the rational $x = u/v$. The rational at the following iteration is $y = s/u$. Now, the even-euclidean division can be written as

$$x = \frac{1}{2c + \varepsilon\, y}.$$

More generally, the whole process finishes with the value 1, and is summarized by a continued fraction

$$\frac{u}{v} = \cfrac{1}{2c_1 + \cfrac{\varepsilon_1}{2c_2 + \cfrac{\varepsilon_2}{\ddots \quad \ddots + \cfrac{\varepsilon_{\ell-1}}{2c_\ell + \varepsilon_\ell}}}}.$$

which uses at each step a set of linear fractional transformations (LFT's)

$$\mathcal{E} := \{h \mid h(x) = \frac{1}{2c + \varepsilon x} \qquad \text{with } c \geq 1,\ \varepsilon = \pm 1\ \}. \tag{1}$$

Remark that the last step is particular, since it only uses the LFT's associated to pairs (c, ε) with $\varepsilon = 1$. The LFT's associated are of the form

$$\mathcal{B}^+ := \{h \mid h(x) = \frac{1}{2c + x} \qquad \text{with } c \geq 1\ \}. \tag{2}$$

Finally, when the Even Algorithm runs on the integer pair (u, v), it finishes with the value 1, and it decomposes the rational $x = u/v$ in a unique way as

$$x = h_1 \circ h_2 \circ \ldots \circ h_\ell(1) \qquad \text{with } h_i \in \mathcal{E} \text{ for } 1 \leq i \leq \ell - 1 \text{ and } h_\ell \in \mathcal{B}^+. \tag{3}$$

The number ℓ which is the height of the continued fraction is exactly equal to the number of iterations of the Even Algorithm. Shallit [36] proves that the worst-case configuration is related to the choice

$$h_i(x) = \frac{1}{2 - x} \quad \text{for} \quad 1 \leq i \leq \ell - 1 \quad \text{and} \quad h_\ell(x) = \frac{1}{2 + x}.$$

The worst–case input is thus formed with two odd successive numbers, on the form $(2\ell - 1, 2\ell + 1)$ on which the Even Algorithm performs exactly ℓ steps. Then the Even Algorithm has *not* a polynomial–time complexity. Furthermore, the LFT $a(x) := 1/(2 - x)$ plays an important rôle in the worst-case configuration. We shall see that its rôle is also predominant in the average-case analysis.

2.4. The Odd Algorithm. This algorithm is based on a pseudo–euclidean division where all the pseudo–quotients are odd. Then, the remainders are even, and one removes the powers of 2. This division is defined as follows: We first consider the ordinary euclidean division between two odd integers u, v

$$v = au + r \qquad \text{with } 0 \le r < u,$$

that we modify as follows:

$$v = bu + \varepsilon\ r' \qquad \text{with } \varepsilon = \pm 1,\ r' \text{ even },\ 0 \le r' < u.$$

If the quotient a is odd, then one lets $b := a, r' := r, \varepsilon := 1$, else one lets $b := a+1, r' := u - r, \varepsilon = -1$. Since b is always odd, then r' is always even, and one lets $r' := 2^k s$ with s odd, and $k \ge 1$. Finally, one obtains the pseudo–euclidean division

$$v = bu + \varepsilon\, 2^k\, s \qquad \text{with } \varepsilon = \pm 1,\ s \text{ odd},\ 0 \le s < \tfrac{u}{2},\ k \ge 1.$$

Odd Algorithm.

Input: a pair (u, v) satisfying $0 < u < v$, u, v odd.
Output: $G(u, v)$ and $J(u, v)$.

begin
$e := 0$;
While $u \neq 0$ **do**
  begin
  Perform the odd–euclidean division between v and u: $v = bu + \varepsilon\, 2^k\, s$;
  $e := e + \frac{1}{4}(u-1)(v-1) + \frac{1}{8}k\,(u^2-1) + \frac{1}{2}\delta(\varepsilon, 1)\,(u-1)$;
  $v := u; u := s$;
  end;
$G(u, v) := v$;
if $v = 1$ **then** $J(u, v) := (-1)^e$, **else** $J(u, v) := 0$.
end.

We now consider two successive odd-euclidean divisions $v = bu + \varepsilon\, 2^k\, s$, $u = cs + \varepsilon'\, 2^{k'}\, s'$. Then, the values of k and c are related, since $2^k\, s < u < (c+1)s$ so that $2 \le 2^k < c+1$, and, since c is odd, we obtain the final condition $c \ge 2^k + 1$. Instead of the integer pair (u, v), one considers the rational $x = u/v$. The rationals of the following iterations are $y = s/u$ and $z = s'/s$. Now, the sequence of the two odd-euclidean divisions can be written as

$$x = \cfrac{1}{b + \cfrac{\varepsilon\, 2^k}{c + \varepsilon'\, 2^{k'} z}}.$$

More generally, the whole process finishes with the value 0, and is summarized by a continued fraction

$$\frac{u}{v} = \cfrac{1}{c_1 + \varepsilon_1 \cfrac{2^{k_2}}{c_2 + \varepsilon_2 \cfrac{2^{k_3}}{\ddots \quad \ddots + \varepsilon_{\ell-1} \cfrac{2^{k_\ell}}{c_\ell}}}}.$$

which uses at each step the set of LFT's

$$\mathcal{O} := \{h \mid h(x) = \frac{2^k}{c + \varepsilon x} \qquad \text{with } c \ge 3 \text{ odd},\ k \ge 1, \text{ and } c \ge 2^k + 1\}. \tag{4}$$

The first step is particular, since it uses the set

$$\mathcal{O}_0 := \{h \mid h(x) = \frac{1}{c+\varepsilon x} \qquad \text{with } c \text{ odd, } c \geq 1,\ \varepsilon = 1 \text{ for } c = 1\}. \tag{5}$$

Finally, when the Odd Algorithm runs on the integer pair (u, v), it finishes with the value 0, and it decomposes the rational $x = u/v$ as

$$x = h_1 \circ h_2 \circ \ldots \circ h_\ell(0) \qquad \text{with } h_1 \in \mathcal{O}_0 \text{ and } h_i \in \mathcal{O} \text{ for } 2 \leq i \leq \ell. \tag{6}$$

The number ℓ which is the height of the continued fraction is exactly equal to the number of iterations of the Odd Algorithm. Shallit does not analyse this algorithm, but it is clear that the worst-case configuration is related to the choice

$$h_1(x) = \frac{1}{1+x} \quad \text{and} \quad h_i(x) = \frac{2}{3-x} \quad \text{for} \quad 2 \leq i \leq \ell.$$

The worst–case input is thus formed with two odd numbers, on the form $(2^\ell - 1, 2^{\ell+1} - 3)$ on which the Odd Algorithm performs exactly ℓ steps. Then the Odd Algorithm has a polynomial–time complexity, and its maximum number of iterations on odd integers less than N is asymptotically equal to $\log_2 N$.

2.5. The Ordinary Algorithm. This algorithm is based on a pseudo–euclidean division where the pseudo–quotients are equal to the "ordinary" quotients: they may be odd or even. When the quotients are odd, the remainders are even, and one removes the powers of 2. When the quotients are even, the remainders are odd, and they are kept as pseudo-remainders. The ordinary euclidean division between two odd integers u, v

$$v = au + r \qquad \text{with } 0 \leq r < u,$$

is modified as follows:

$$v = bu + 2^k s \qquad \text{with } b := a,\ s = 0 \text{ or } s \text{ odd, } s < u,\ k \geq 0.$$

Ordinary Algorithm.

Input: a pair (u, v) satisfying $0 < u < v$, u, v odd.
Output: $G(u, v)$ and $J(u, v)$

begin
$e := 0$;
While $u \neq 0$ **do**
  begin
  Perform the ordinary–euclidean division between v and u: $v = bu + 2^k s$
  $e := e + \frac{1}{4}(u-1)(v-1) + \frac{1}{8}k\,(u^2-1)$
  $v := u; u := s;$
  end;
$G(u, v) := v$;
if $v = 1$ **then** $J(u, v) := (-1)^e$, **else** $J(u, v) := 0$.
end.

We now consider two successive ordinary–euclidean divisions

$$v = bu + 2^k s \qquad u = cs + 2^{k'} s'.$$

We have already remarked that the sizes of k and c are related, since $2^k\, s < u < (c+1)s$ so that $2^k < c+1$. Then for odd c, one obtains $c \geq 2^k+1$, and for even c, one obtains $c \geq 2^k$. But, there exists another link which involves parity inside the same ordinary euclidean division: 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 link is of markovian type, and we consider two states: the 0 state, which means "the remainder of (u, v) is odd", i.e. $k = 0$, and the 1 state, which means "the remainder of (u, v) is even ", i.e. $k \geq 1$. Instead of the integer pair (u, v), one considers the rational $x = u/v$, and we denote by $\mathcal{U}_{j|i}$ the set of LFT's which brings rationals from state i to state j. We obtain four different sets

$$\mathcal{U}_{0|0} = \{h \mid h(x) = \frac{1}{c+x} \qquad \text{with } c \text{ even, } c \geq 2\ \}, \tag{7}$$

$$\mathcal{U}_{1|0} = \{h \mid h(x) = \frac{1}{c+x} \qquad \text{with } c \text{ odd, } c \geq 1\ \}, \tag{8}$$

$$\mathcal{U}_{0|1} = \{h \mid h(x) = \frac{2^k}{c+x} \qquad \text{with } c \geq 2 \text{ even, } k \geq 1,\ c \geq 2^k\ \}, \tag{9}$$

$$\mathcal{U}_{1|1} = \{h \mid h(x) = \frac{2^k}{c+x} \qquad \text{with } c \geq 3 \text{ odd, } k \geq 1,\ c \geq 2^k+1\ \}. \tag{10}$$

The initial state is the 0 state, and the process finishes at the 1 state. Finally, the whole process decomposes the rational $x = u/v$ as

$$x = h_1 \circ h_2 \circ \ldots \circ h_\ell(0) \qquad \text{with } h_k \in \mathcal{U}_{j_k|i_k}\ ,\ i_1 = 0, j_\ell = 1,\ i_{k+1} = j_k \text{ for all } 1 \leq k \leq \ell - 1\ . \tag{11}$$

The number ℓ which is the height of the continued fraction is exactly equal to the number of iterations of the Ordinary Algorithm. Shallit analyses this algorithm, and proves that the worst-case configuration mainly involves the two LFT's b and c

$$b(x) := \frac{1}{1+x} \qquad c(x) := \frac{2}{2+x}$$

that are the "minimal" elements of $\mathcal{U}_{1|0}$ and $\mathcal{U}_{0|1}$. The last step involves the minimal element f of $\mathcal{U}_{1|1}$ and the first step involves the minimal element of $\mathcal{U}_{0|0}$ (for odd values of the height),

$$f(x) := \frac{2}{3+x} \qquad i(x) := \frac{1}{2+x}.$$

More precisely, the worst–case configuration depends on the parity of the height; it is

$$(bc)^\ell bf(0) \ \text{ for even heights} \qquad i(bc)^\ell bf(0) \ \text{ for odd heights.}$$

The maximum number of iterations on odd integers less than N is thus equivalent to

$$\log_\gamma N \qquad \text{with} \quad \gamma = (\frac{5+\sqrt{17}}{2})^{1/2} > 2.$$

Then the Ordinary Algorithm has a polynomial–time complexity, and its maximum number of iterations is slightly less than the maximum number of iterations of the Odd Algorithm.

2.6. Common structure of the three algorithms. When it performs ℓ pseudo–euclidean divisions on the input (u, v), each of the algorithm builds a continued fraction of height ℓ for the rational u/v: as relations (3), (6), (11) show it, the algorithm decomposes the rational $x = u/v$ as

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

where the h_i's are LFT's and a is the last value of the rational. The value a equals 1 for the Even Algorithm, and equals 0 for the other two algorithms. The precise form of the possible LFT's also depends on the algorithm.

In the case of the Even or the Odd algorithms, there exist special sets of LFT's in the first step for the Odd algorithm (5), or in the final step for the Even Algorithm (2). All the other steps are generic, in the sense that they use the same set of LFT's, that we call the generic set.

On the contrary, the Ordinary Algorithm has a Markovian flavour, since each step has two possible states, 0 and 1. There are four sets of LFT's, described in (7, 8, 9, 10) and each of these sets "brings" rationals from state i to state j.

2.7. A numerical example. For the pair $(13, 75)$, we obtain three different continued fractions:

$$\frac{13}{75} = \cfrac{1}{6 - \cfrac{1}{4 + \cfrac{1}{2+1}}} \qquad \frac{13}{75} = \cfrac{1}{5 + \cfrac{2}{3 - \cfrac{1}{5+0}}} \qquad \frac{13}{75} = \cfrac{1}{5 + \cfrac{2}{2 + \cfrac{1}{1 + \cfrac{2}{3+0}}}}.$$

3 Generating functions, dynamical operators and Tauberian Theorems.

Here, we describe the general tools for analysing the three algorithms. We first introduce the generating functions relative to the height of the continued fraction and we relate them to the dynamical operator $\mathbf{H}_s$ associated to the algorithm. In this way, the two Dirichlet series that intervene in the analysis, called $F(s)$ and $G(s)$, are expressed in terms of the Ruelle operator. The average number of steps involves partial sums of coefficients of these two Dirichlet series, and Tauberian Theorems are a classical tool that transfers analytical properties of Dirichlet series into asymptotic behaviour of their coefficients.

3.1. Generating functions. We consider the following sets

$$\Omega := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v\}, \quad \widetilde{\Omega} := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v, \gcd(u,v) = 1\},$$

$$\Omega_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N\},$$

$$\widetilde{\Omega}_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N,\ \gcd(u,v) = 1\},$$

for the possible inputs of the Jacobi symbol algorithm, and we denote by $\Omega^{[\ell]}$, $\widetilde{\Omega}^{[\ell]}$, $\Omega_N^{[\ell]}$, $\widetilde{\Omega}_N^{[\ell]}$ the subsets of Ω, $\widetilde{\Omega}$, Ω_N, $\widetilde{\Omega}_N$ for which the algorithm performs exactly ℓ pseudo–divisions. Equivalently, the height of the continued fraction is equal to ℓ. We study the average number of steps $\widetilde{S}_N$ of the algorithm on $\widetilde{\Omega}_N$ and the average number of steps S_N of the algorithm on Ω_N. These quantities satisfy

$$\widetilde{S}_N := \frac{1}{|\widetilde{\Omega}_N|} \sum_{\ell \ge 0} \ell\ |\widetilde{\Omega}_N^{[\ell]}| \qquad S_N := \frac{1}{|\Omega_N|} \sum_{\ell \ge 0} \ell\ |\Omega_N^{[\ell]}| \tag{12}$$

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

$$\widetilde{S}(s,w) := \sum_{\ell \ge 1} w^\ell \sum_{n > 1} \frac{\widetilde{\nu}_n^{[\ell]}}{n^s}, \qquad S(s,w) := \sum_{\ell \ge 1} w^\ell \sum_{n > 1} \frac{\nu_n^{[\ell]}}{n^s}. \tag{13}$$

Since the algorithms perform the same steps on two pairs (u, v) and (ru, rv) for any odd integer r, the Riemann series $\widetilde{\zeta}$ relative to odd numbers

$$\widetilde{\zeta}(s) := \sum_{v \text{ odd}} \frac{1}{v^s} = (1 - \frac{1}{2^s})\zeta(s) \tag{14}$$

relates sums over $\Omega^{[\ell]}$ and sums over $\widetilde{\Omega}^{[\ell]}$

$$\sum_{n>1} \sum_{(u,n)\in\widetilde{\Omega}^{[\ell]}} \frac{1}{n^s} f(\frac{u}{n}) = \frac{1}{\widetilde{\zeta}(s)} \sum_{n>1} \sum_{(u,n)\in\Omega^{[\ell]}} \frac{1}{n^s} f(\frac{u}{n}) \qquad \text{so that} \quad \sum_{n>1} \frac{\widetilde{\nu}_n^{[\ell]}}{n^s} = \frac{1}{\widetilde{\zeta}(s)} \sum_{n>1} \frac{\nu_n^{[\ell]}}{n^s},$$

and the two generating functions satisfy the equality

$$S(s, w) = \widetilde{\zeta}(s)\widetilde{S}(s, w). \tag{15}$$

It is then sufficient to study $\widetilde{S}(s, w)$. We introduce the two sequences

$$a_n := \sum_{\ell\geq 1} \widetilde{\nu}_n^{[\ell]} \qquad b_n := \sum_{\ell\geq 1} \ell \; \widetilde{\nu}_n^{[\ell]}, \tag{16}$$

that intervene in the evaluation of the quantity $\widetilde{S}_N$,

$$\widetilde{S}_N = \frac{\sum_{n\leq N} \sum_{\ell\geq 0} \ell \; \widetilde{\nu}_n^{[\ell]}}{\sum_{n\leq N} \sum_{\ell\geq 0} \widetilde{\nu}_n^{[\ell]}} = \frac{\sum_{n\leq N} b_n}{\sum_{n\leq N} a_n} \tag{17}$$

So, the ratio $\widetilde{S}_N$ involves partial sums of coefficients of the two Dirichlet series,

$$F(s) = \sum_{n>1} \frac{a_n}{n^s}, \qquad G(s) = \sum_{n>1} \frac{b_n}{n^s}, \tag{18}$$

that can be easily expressed in terms of $\widetilde{S}(s, w)$ since

$$F(s) = \widetilde{S}(s, 1), \qquad G(s) = \frac{d}{dw}\widetilde{S}(s, w)|_{w=1}. \tag{19}$$

3.2. Ruelle operators (generalities). We show now how the Ruelle operators associated to the algorithms intervene in the evaluation of the generating function $\widetilde{S}(s, w)$. We denote by $\mathcal{L}$ a set of LFT's. For each $h \in \mathcal{L}$, $D[h]$ denotes the denominator of the linear fractional transformation (LFT) h, defined for $h(x) = (ax + b)/(cx + d)$ with a, b, c, d coprime integers by

$$D[h](x) = |cx + d| = \sqrt{\frac{|\det h|}{|h'(x)|}}, \tag{20}$$

that can be extended on a natural way on a complex neighborhood of $]0, 1]$. The Ruelle operator $\mathbf{L}_s$ relative to the set $\mathcal{L}$ is defined as

$$\mathbf{L}_s[f](x) := \sum_{h\in\mathcal{L}} \frac{1}{D[h](x)^s} \; f \circ h(x). \tag{21}$$

More generally, when given two sets of LFT's, $\mathcal{L}$ and $\mathcal{K}$, the set $\mathcal{LK}$ is formed of all $h \circ g$ with $h \in \mathcal{L}$ and $g \in \mathcal{K}$. The multiplicativity of D, defined in (20)

$$D[h \circ g](x) = D[h](g(x)) \, D[g](x), \tag{22}$$

proves that the operator $\mathbf{K}_s \circ \mathbf{L}_s$ uses all the LFT's of $\mathcal{LK}$

$$\mathbf{K}_s \circ \mathbf{L}_s[f](x) := \sum_{h\in\mathcal{LK}} \frac{1}{D[h](x)^s} \; f \circ h(x). \tag{23}$$

3.3. Ruelle operators (basic case). The first two algorithms uses the same set $\mathcal{H}$ at each generic step. In this case, the Ruelle operator associated to $\mathcal{H}$ can be viewed as a "generating" operator, since the ℓ-th iterate of $\mathbf{H}_s$ generates all the LFT's used in ℓ (generic) steps of the algorithm. This explains why usual generating functions can be easily expressed with the Ruelle operator associated to the generic set $\mathcal{H}$.

Proposition 1. *The double generating function $\widetilde{S}(s,w)$ of the sequence $(\widetilde{\nu}_n^{[\ell]})$ defined as*

$$\widetilde{S}(s,w) := \sum_{\ell\geq 1} w^\ell \sum_{n>1} \frac{\widetilde{\nu}_n^{[\ell]}}{n^s}$$

can be expressed as a function of the Ruelle operators associated to the algorithm,

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

Here, the 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 $\mathbf{I}_s$ is related to the set $\mathcal{I}$ used at the initial step; The operator $\mathbf{J}_s$ is related to the set $\mathcal{J} := \mathcal{I}\cap\mathcal{F}$ and solely intervenes when the algorithm performs only one step; the value a is the final value of the rational u/v.
The average number of steps $\widetilde{S}_N$ of the algorithm on $\widetilde{\Omega}_N$ is a ratio where the numerators and the denominators involve the partial sums of the Dirichlet series $F(s)$ and $G(s)$, defined from the quasi–inverse $(I-\mathbf{H}_s)^{-1}$,

$$F(s) := \mathbf{J}_s[1](a) + \mathbf{F}_s\circ(I-\mathbf{H}_s)^{-1}\circ\mathbf{I}_s[1](a),$$

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

Remark. In the same vein, from the relation (15), the average number of steps S_N of the algorithm on Ω_N is a ratio where the numerators and the denominators involve the partial sums of the Dirichlet series $\widetilde{\zeta}(s)F(s)$, and $\widetilde{\zeta}(s)G(s)$.

Proof. Consider an element (u,v) of $\widetilde{\Omega}^{[\ell+2]}$ $(\ell\geq 0)$; 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 $\mathcal{I}\mathcal{H}^\ell\mathcal{F}$ such that $u/v = h(a)$. Then, from (23) and for any $\ell\geq 0$,

$$\mathbf{F}_s\circ\mathbf{H}_s^\ell\circ\mathbf{I}_s[f](a) = \sum_{h\in\mathcal{I}\mathcal{H}^\ell\mathcal{F}} \frac{1}{D[h](a)^s}\ f\circ h(a) = \sum_{n>1}\ \sum_{(u,n)\in\widetilde{\Omega}^{[\ell+2]}} \frac{1}{n^s} f(\frac{u}{n}). \tag{24}$$

In the same vein, when the algorithm performs only one division on this input (u,v), and finishes with the value a, there exists exactly one linear fractional transformation h in the set $\mathcal{J}$ such that $u/v = h(a)$ and

$$\mathbf{J}_s[f](a) = \sum_{h\in\mathcal{J}} \frac{1}{D[h](a)^s}\ f\circ h(a) = \sum_{n>1}\ \sum_{(u,n)\in\widetilde{\Omega}^{[1]}} \frac{1}{n^s} f(\frac{u}{n}). \tag{25}$$

When choosing $f=1$, and summing expressions of (25) and (24) over all the possible values of ℓ, we derive the expression of the double generating function $\widetilde{S}(s,w)$

$$\widetilde{S}(s,w) = \sum_{\ell\geq 1} w^\ell \sum_{n>1}\ \sum_{(u,n)\in\widetilde{\Omega}^{[\ell]}} \frac{1}{n^s} = w\mathbf{J}_s[1](a)\ +\ \sum_{\ell\geq 0} w^{\ell+2}\,\mathbf{F}_s\circ\mathbf{H}_s^\ell\circ\mathbf{I}_s[1](a)$$

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

When using expressions (19) of F and G in terms of S, and derivating with respect to w and/or taking $w=1$, we prove the proposition. ∎

3.4. Ruelle operators (Ordinary algorithm). Here, each step has two possible states, 0 and 1. There are four sets of LFT's, and each of these sets, denoted by $\mathcal{U}_{i|j}$ "brings" rationals from state j to state i. We denote by $\mathcal{U}_{i|j}^{<\ell>}$ the set of LFTs that brings rationals from state j to state i in ℓ steps. Then the union $\mathcal{U}^{<\ell>}$ of the four subsets $\mathcal{U}_{i|j}^{<\ell>}$ contains all the possible LFTs used in ℓ steps. 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 a LFT of set $\mathcal{U}_{1|0}^{<\ell>}$ and decomposes the rational $x = u/v$ as

$$x = h_1 \circ h_2 \circ \ldots \circ h_\ell(0) \qquad \text{with } h_k \in \mathcal{U}_{j_k|i_k}\ ,\ i_1 = 0, j_\ell = 1,\ i_{k+1} = j_k \text{ for all } 1 \leq k \leq \ell - 1\ .$$

We denote by $\mathbf{U}_{s,i|j}$ the Ruelle operator associated to the set $\mathcal{U}_{i|j}$, and by $\mathbf{U}_s$ the "matrix operator"

$$\mathbf{U}_s = \begin{pmatrix} \mathbf{U}_{s,0|0} & \mathbf{U}_{s,0|1} \\ \mathbf{U}_{s,1|0} & \mathbf{U}_{s,1|1} \end{pmatrix}. \tag{26}$$

By multiplicative properties (23), 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 $\mathcal{U}_{i|j}^{<\ell>}$. Moreover, the last value of the rational u/v equals 0, and the algorithm begins at the 0 state and finishes at the 1 state. Then

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

When choosing $f = 1$, and summing expressions of (27) over all the possible values of ℓ, we derive the expression of the double generating function $\widetilde{S}(s,w)$

$$\widetilde{S}(s,w) = \sum_{\ell \geq 1} w^\ell \sum_{n>1} \sum_{(u,n) \in \widetilde{\Omega}^{[\ell]}} \frac{1}{n^s} = (\,0 \quad 1\,)\, w\mathbf{U}_s\, (I - w\mathbf{U}_s)^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0)$$

Now, using the expression of F and G in terms of $\widetilde{S}$, we derivate this expression with respect to w and/or take $w = 1$, and we have shown:

Proposition 2. *The double generating function $\widetilde{S}(s,w)$ of the sequence $(\widetilde{\nu}_n^{[\ell]})$ in the Ordinary Algorithm defined as*

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

can be expressed as a function of the Ruelle operator associated to the algorithm,

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

The average number of steps $\widetilde{S}_N$ of the Ordinary Algorithm on $\widetilde{\Omega}_N$ is a ratio where the numerators and the denominators involve the partial sums of the Dirichlet series $F(s)$ and $G(s)$, defined from the quasi–inverse $(I - \mathbf{U}_s)^{-1}$,

$$F(s) := (\,0 \quad 1\,)\ \ \mathbf{U}_s(I - \mathbf{U}_s)^{-1}\ \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0),$$

$$G(s) := (\,0 \quad 1\,)\ \ \mathbf{U}_s\ (I - \mathbf{U}_s)^{-2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0).$$

Remark. As in the previous paragraph, the average number of steps S_N of the algorithm on Ω_N is a ratio where the numerators and the denominators involve the partial sums of the Dirichlet series $\widetilde{\zeta}(s)F(s)$, and $\widetilde{\zeta}(s)G(s)$.

3.5. Tauberian Theorems. Thus, in the two possible cases described in 3.2 or 3.3, the asymptotic evaluations of $\widetilde{S}_N$ and S_N (for $N \to \infty$) are possible if we can apply the following Tauberian theorem [8], [41] to the Dirichlet series $F(s)$, $\widetilde{\zeta}(s)F(s)$, $G(s)$, $\widetilde{\zeta}(s)G(s)$.

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 $\gamma \geq 0$

$$F(s) = \frac{A(s)}{(s-\sigma)^{\gamma+1}} + C(s),$$

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 show that the Tauberian Theorem applies to the Dirichlet series defined in (18) with $\sigma = 2$. For $F(s)$, it applies with $\gamma = 0$. For $G(s)$, it applies with $\gamma = 1$ or $\gamma = 2$. For the Even Algorithm, γ equals 2, and the average number of steps will be of order $\log^2 N$. For the Odd Algorithm or the Ordinary Algorithm, 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 (18).

Lemma 1. *The Tauberian Theorem applies to $F(s)$ and $\widetilde{\zeta}(s)F(s)$ with $\sigma = 2$ and $\gamma = 0$.*

Proof. First, the Dirichlet series $F(s)$ is closely linked to the $\widetilde{\zeta}$ function defined in (14), since

$$F(s) = \frac{1}{2\widetilde{\zeta}(s)} \sum_{v \text{ odd}} \frac{v-1}{v^s} = \frac{1}{2}[\frac{\widetilde{\zeta}(s-1)}{\widetilde{\zeta}(s)} - 1].$$

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

$$F(s) \simeq \frac{2}{\pi^2}\frac{1}{s-2} \tag{28}$$

4 Functional Analysis.

Here, we establish all the properties of the Ruelle operators that we need in the sequel. If the LFT's used by the algorithm satisfy some natural conditions, i.e. are sufficiently regular and contracting, then the Ruelle operator is compact (even nuclear in the sense of Grothendieck) on a suitable space of analytic functions. Furthermore, for real values of parameter s, it has dominant spectral properties. Then, there is a spectral gap between the dominant eigenvalue and the remainder of the spectrum. Finally, the quasi–inverse of the Ruelle operator which intervenes in the expression of generating function $\widetilde{S}$ fulfills the hypothesis of Tauberian Theorem.

In this section, we denote by $\mathbf{H}_s$ a Ruelle operator associated to a generic set $\mathcal{H}$, and by $\mathbf{U}_s$ a Ruelle markovian operator (associated to a Markovian process with two states). In this case, we add some notations

$$\mathcal{U} = \bigcup_{i,j=0}^{1} \mathcal{U}_{j|i} \qquad \mathcal{U}_i = \bigcup_{j=0}^{1} \mathcal{U}_{j|i}.$$

4.1. Component operators. Each component operator of the Ruelle operator $\mathbf{H}_s$ or $\mathbf{U}_s$ that we use is associated to an LFT h, it is denoted by $\mathbf{H}_{s,h}$. This is a so–called composition operator, defined by

$$\mathbf{H}_{s,h}[f] := \frac{1}{D[h]^s}\ f \circ h.$$

Here, each LFT h is a (real) analytic function which maps $]0,1[$ inside $]0,1[$ and satisfies $|h'(x)| \le 1$ on $]0,1[$. We consider a stronger condition $\mathcal{C}(\mathcal{V}, h)$.

Conditions $\mathcal{C}(\mathcal{V}, h)$.*There exists a suitable disk $\mathcal{V}$ that contains $[0,1]$ such that the following holds:*

(C_1) *h has an analytic continuation on $\mathcal{V}$,*

(C_2) *h maps the closure $\bar{\mathcal{V}}$ of the disk $\mathcal{V}$ inside $\mathcal{V}$,*

(C_3) *$|h'|$ has an analytic continuation on $\mathcal{V}$denoted by $\widetilde{h}$ and there exists $\delta < 1$ for which $0 < |\widetilde{h}(z)| \le \delta$ for all $z \in \mathcal{V}$.*

Under these conditions, the operator $\mathbf{H}_{s,h}$ acts on the space $A_\infty(\mathcal{V})$ formed with 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. Such operators are studied in an extensive way by several authors, in particular Shapiro [38], [39] who proves the following:

Property 1. Compacity of component operators. *If the conditions $\mathcal{C}(\mathcal{V}, h)$ hold, the operator $\mathbf{H}_{s,h}$: $A_\infty(\mathcal{V}) \to A_\infty(\mathcal{V})$ is compact.*

The operator has even stronger properties, since it is then nuclear of order 0 (in the sense of Grotehendieck [16], [17]. 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.

4.2. Compacity of the Ruelle operator. The previous properties are also true for the Ruelle operators $\mathbf{H}_s$ or $\mathbf{U}_s$ themselves provided that some uniformity and convergence be added in the conditions $\mathcal{C}(\mathcal{V}, h)$.

Conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$ *There exists a suitable disk $\mathcal{V}$ that contains $[0,1]$ such that the following holds:*

(C_1) *All the LFT's $h \in \mathcal{H}$ have an analytic continuation on $\mathcal{V}$,*

(C_2) *For all $h \in \mathcal{H}$, h maps the closure $\bar{\mathcal{V}}$ of the disk $\mathcal{V}$ inside $\mathcal{V}$,*

(C_3) *There exists $\delta(h) < 1$ for which $0 < |\widetilde{h}(z)| \le \delta(h)$ for all $z \in \mathcal{V}$,*

(C_4) *There exists $\sigma < 2$ for which the series*

$$\sum_{h\in\mathcal{H}} \left|\frac{\delta(h)}{\det(h)}\right|^{s/2}$$

converges on the plane $\Re(s) > \sigma$.

In the sequel of this Section, we always suppose that the conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$ or $\mathcal{C}(\mathcal{V}, \mathcal{U}, \sigma)$ are satisfied. In this case, the Ruelle operators $\mathbf{H}_s$ or $\mathbf{U}_s$ have quite good properties for s in the plane $\Re(s) > \sigma$. First, one can easily prove:

Property 2. Compacity. *If the conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$ are satisfied, the operator $\mathbf{H}_s$ acts on $\mathcal{A}_\infty(\mathcal{V})$ for $\Re(s) > \sigma$. If the conditions $\mathcal{C}(\mathcal{V}, \mathcal{U}, \sigma)$ are satisfied, the operator $\mathbf{U}_s$ acts on $\mathcal{A}_\infty(\mathcal{V})^2$ for $\Re(s) > \sigma$. Moreover, in each case, the operators are compact.*

In the same way as in previous paragraph, the operators $\mathbf{H}_s$ and $\mathbf{U}_s$ are also nuclear of order 0.

4.3. Dominant spectral properties for real s. Under the natural condition $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$, we have shown a first important property. We now define another condition under which we can establish other important properties.

Condition $\mathcal{D}(\mathcal{H})$. *For some integer A, the set $\mathcal{H}$ contains a subset*

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

When $s = \sigma$ is real, the operators $\mathbf{H}_s$ satisfy strong positivity properties related to the Perron-Frobenius theory [28].

Property 3. Dominant eigenvalue. *Under conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$, and $\mathcal{D}(\mathcal{H})$, and for real $s > \sigma$, the operator $\mathbf{H}_s$ has a unique dominant eigenvalue (of largest modulus) $\lambda(s)$. It is positive, has multiplicity 1. The same is true for the operator $\mathbf{U}_s$ under conditions $\mathcal{C}(\mathcal{V}, \mathcal{U}, \sigma)$, and $\mathcal{D}(\mathcal{U}_0)$ and $\mathcal{D}(\mathcal{U}_1)$.*

Proof. We follow the lines of Mayer's work [32] that we adapt in our context. Mayer himself uses a result due to Krasnoselsky [28].

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

$$\alpha\ u_0 \leq \mathcal{L}^p[f] \leq \beta\ u_0, \tag{29}$$

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 (29) 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 $\mathcal{D}$ 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 (29), 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. ∎

4.4. Spectral gap. With two previous properties, we can separate the dominant eigenvalue from the remainder of the spectrum.

Property 4. Spectral Gap. *Suppose that conditions $\mathcal{C}(\mathcal{V},\mathcal{H},\sigma)$ and $\mathcal{D}(\mathcal{H})$ hold. Then, for real $\rho > \sigma$,*
(i) 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.
(ii) The function $s \to \lambda(s)$ is strictly decreasing along the real line for $s > \sigma$.

The same is true for operator $\mathbf{U}_s$ under conditions $\mathcal{C}(\mathcal{V},\mathcal{U},\sigma)$ and $\mathcal{D}(\mathcal{U}_i), i = 0, 1$.

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

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

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

$$\mathbf{H}_\rho^\ell[f](z) = \lambda(\rho)^\ell \mathbf{P}_\rho[f](z) + \mathbf{Q}_\rho^\ell[f](z). \tag{30}$$

By Perturbation theory, the decomposition (30) extends for $\mathbf{H}_s$ to a (complex) neighborood of $s = \rho$, the dominant eigenvalue $\lambda(s)$ is analytic there, and $\mathbf{Q}_s$ has a spectral radius less than $\beta\lambda(s)$ (with $\beta < 1$. Then, one has, for all $\ell \geq 1$, and for $z \in \mathcal{V}$,

$$\mathbf{H}_s^\ell[f](z) = \lambda(s)^\ell \mathbf{P}_s[f] + \mathbf{Q}_s^\ell[f](z). \tag{31}$$

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).
Now, for proving (ii), 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 (30), we obtain

$$\lambda(\rho) = \lim_{\ell\to\infty} \Big(\sum_{h\in\mathcal{H}^\ell} \frac{1}{D[h](0)^\rho}\Big)^{1/\ell}. \tag{32}$$

¿From items (C_3) and (C_4) of conditions $\mathcal{C}(\mathcal{V},\mathcal{H},\sigma)$, there exists $\gamma < 1$ such that

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

and then, by multiplicativity of D (22), and for any ℓ

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

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 defined in paragraph 3.4, 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. ■

4.5. Maximum properties. We now exhibit an important property of the radius spectral $R(s)$ of the operator $\mathbf{H}_s$, when s varies on a vertical line $\Re(s) = \rho$: it attains its maximum for $s = \rho$.

Property 5. Maximum Properties. *Suppose that the conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$ and $\mathcal{D}(\mathcal{H})$ hold. Then, for any s on the vertical line $\Re(s) = \rho, s \neq \rho$, the spectral radius $R(s)$ of the operator $\mathbf{H}_s$ is strictly less than $\lambda(\rho)$. The same is true for the operator $\mathbf{U}_s$ under conditions $\mathcal{C}(\mathcal{V}, \mathcal{U}, \sigma)$ and $\mathcal{D}(\mathcal{U}_i), i = 0, 1$.*

Proof. 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

$$\mu(x) := \frac{f(x)}{f_\rho(x)} \tag{33}$$

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

$$|\lambda\, f(x_0)| = |\mathbf{H}_s[f](x_0)| = |\sum_{h\in\mathcal{H}} D[h](x_0)^{-s}\ f\circ h(x_0)| \leq \sum_{h\in\mathcal{H}} D[h](x_0)^{-\rho}\ |f\circ h(x_0)| \tag{34}$$

$$\leq \sum_{h\in\mathcal{H}} D[h](x_0)^{-\rho}\ f_\rho\circ h(x_0) = \lambda(\rho) f_\rho(x_0), \tag{35}$$

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 (34), (35) becomes a sequence of equalities. For any $h \in \mathcal{H}$, the equality

$$|f\circ h(x_0)| \ = \ f_\rho\circ h(x_0) \tag{36}$$

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

$$f\circ h(x_0)\ D[h](x_0)^{-it} \ = \ \theta\ |f\circ h(x_0)| \tag{37}$$

holds. Both relations (36), (37) 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 (36) proves that $|f(0)| \ = \ f_\rho(0) \neq 0$. Now, for $c \to \infty$, Relation (37) shows that the sequence

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

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

$$\mu_i(x) := \frac{f_i(x)}{f_{\rho,i}(x)} \tag{38}$$

are of modulus at most 1 on $[0,1]$, and the function μ_ℓ attains modulus 1 at point x_0. One always has

$$|\lambda\, f_\ell(x_0)| = |\sum_j D[h_{\ell|j}](x_0)^{-s}\ f_j \circ h_{\ell|j}(x_0)| \le \sum_j D[h_{\ell|j}](x_0)^{-\rho}\ |f_j \circ h_{\ell|j}(x_0)| \tag{39}$$

$$\le \sum_j D[h_{\ell|j}](x_0)^{-\rho}\ f_{\rho,j} \circ h_{\ell|j}(x_0) = \lambda(\rho) f_{\rho,\ell}(x_0), \tag{40}$$

and the definition of point x_0 and index ℓ proves the inequality $|\lambda| \le \lambda(\rho)$. Now, we suppose that, for $s = \rho + it, t \ne 0$, there exists $\lambda \in \mathrm{Sp}\,\mathbf{H}_s$ that satisfies $|\lambda| = \lambda(\rho)$. Then, the sequence of inequalities (39), (40) becomes a sequence of equalities. In particular, for any symbol j, we deduce the equalities

$$|f_j \circ h_{\ell|j}(x_0)| = f_{\rho,j} \circ h_{\ell|j}(x_0), \tag{41}$$

$$f_j \circ h_{\ell|j}(x_0)\ D[h_{\ell|j}](x_0)^{-it} \ = \ \theta\ |f_j \circ h_{\ell|j}(x_0)|, \tag{42}$$

that are in particular valid for the subset $\mathcal{D}_j$ of $\mathcal{U}_j$. Then, for $c \to \infty$, the sequence $h_{\ell|j}(x_0)$ tends to 0, and, equality (41) proves that $|f_j(0)| \ = \ f_{\rho,j}(0) \ne 0$. Now, for $c \to \infty$, Relation (42) shows that the sequence

$$D[h_{\ell|j}](x_0)^{-it} = (\frac{1}{c + x_0})^{it}$$

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

4.6. Special values at $s = 2$. Here, a "partition" condition plays an important rôle. This condition is natural when considering the algorithms as dynamical systems (cf paragraph 5.5). This condition is slightly different according to the case (generic or Markov).

Generic Conditions $\mathcal{Q}(\mathcal{H})$. *Let $\mathcal{H}_{[k]}$ be the subset of $\mathcal{H}$ defined as $\mathcal{H}_{[k]} := \{h \in \mathcal{H} \mid \det(h) = 2^k\}$.*
(Q_1) One of two conditions (i) or (ii) holds:
(i) $\mathcal{H} = \mathcal{H}_{[0]}$
(ii) For any $k \ge 1$, $\mathcal{H}_{[k]}$ is not empty and $\mathcal{H} = \bigcup_{k\ge 1} \mathcal{H}_{[k]}$
(Q_2) For any $k \ge 0$ for which $\mathcal{H}_{[k]}$ is not empty, the intervals $(h(\mathcal{I}), h \in \mathcal{H}_{[k]})$ form a pseudo-partition of the unit interval $\mathcal{I}$.

Markov Conditions $\mathcal{Q}(\mathcal{U})$. *Let $\mathcal{U}_{[k]}$ be the subset of $\mathcal{U}$ defined as $\mathcal{U}_{[k]} := \{h \in \mathcal{U} \mid \det(h) = 2^k\}$.*
(Q_1) Both conditions (i) and (ii) holds:
(i) $\mathcal{U}_0 = \mathcal{U}_{[0]}$
(ii) For any $k \ge 1$, $\mathcal{U}_{[k]}$ is not empty and $\mathcal{U}_1 = \bigcup_{k\ge 1} \mathcal{H}_{[k]}$
(Q_2) The intervals $(h(\mathcal{I}), h \in \mathcal{U}_{[k]})$ form a pseudo-partition of the unit interval $\mathcal{I}$.

These conditions prove that the Ruelle operators are density transformers and make explicit some of the dominant spectral objects at $s = 2$, namely the dominant eigenvalue and the dominant projector.

Property 6. Special values at $s = 2$. *Under conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$, $\mathcal{D}(\mathcal{H})$ and $\mathcal{Q}(\mathcal{H})$, the following holds:*
(i) The Ruelle operator $\mathbf{H}_s$ is at $s = 2$ a density transformer, i.e., for any function f positive on $[0,1]$, one has

$$\int_0^1 \mathbf{H}_2[f](x)dx = \int_0^1 f(x)dx.$$

(ii) At $s = 2$, the dominant spectral objects of $\mathbf{H}_s$ satisfy $\lambda(2) = 1$ and $e_2[f] = \int_0^1 f(x)dx$.
Under conditions $\mathcal{C}(\mathcal{V}, \mathcal{U}, \sigma)$, $\mathcal{D}(\mathcal{U}_i) (i = 0, 1)$, and $\mathcal{Q}(\mathcal{U})$, the following holds:

(i) The Ruelle operator $\mathbf{U}_s$ is at $s = 2$ a density transformer, i.e., for any function $f = (f_0, f_1)$ whose components are positive on $[0, 1]$

$$\int_0^1 (1 \quad 1)\, \mathbf{U}_2[f](x)dx = \int_0^1 [f_0(x) + f_1(x)]dx.$$

(ii) At $s = 2$, the dominant spectral objects of $\mathbf{U}_s$ satisfy $\lambda(2) = 1$ and $e_2[f] = \int_0^1 [f_0(x) + f_1(x)]dx$.

Proof. We begin by the generic case. From the expression of D given in (20), one of two expressions holds

$$\mathbf{H}_2[f](x) = \sum_{h \in \mathcal{H}_{[0]}} |h'(x)| f \circ h(x) \qquad \text{or} \quad \mathbf{H}_2[f](x) = \sum_{k \geq 1} \frac{1}{2^k} \sum_{h \in \mathcal{H}_{[k]}} |h'(x)| f \circ h(x).$$

In the first case, $\displaystyle\int_0^1 \mathbf{H}_2[f](x) = \sum_{h \in \mathcal{H}_{[0]}} \int_{h(\mathcal{I})} f(u)du = \int_0^1 f(u)du,$

in the second case, $\displaystyle\int_0^1 \mathbf{H}_2[f](x) = \sum_{k \geq 1} \frac{1}{2^k} \sum_{h \in \mathcal{H}_{[k]}} \int_{h(\mathcal{I})} f(u)du = \sum_{k \geq 1} \frac{1}{2^k} \int_0^1 f(u)du = \int_0^1 f(u)du,$

so that the operator is, in both generic cases, a density transformer. In the same vein, for the Markov case, one has,

$$(1 \quad 1)\, \mathbf{U}_2[f](x)dx = \sum_{h \in \mathcal{U}_0} |h'(x)| f_0 \circ h(x) + \sum_{k \geq 1} \frac{1}{2^k} \sum_{h \in \mathcal{U}_{[k]}} |h'(x)| f_1 \circ h(x).$$

Then

$$\int_0^1 (1 \quad 1)\, \mathbf{U}_2[f](x)dx = \sum_{h \in \mathcal{U}_0} \int_{h(\mathcal{I})} f_0(u)du + \sum_{k \geq 1} \frac{1}{2^k} \sum_{h \in \mathcal{U}_{[k]}} \int_{h(\mathcal{I})} f_1(u)du = \int_0^1 [f_0(u) + f_1(u)]du,$$

and the operator $\mathbf{U}_2$ is a density transformer.

It is clear that a compact density transformer has its dominant eigenvalue equal to 1. In the generic case, from decomposition (30) at $s = 2$, we obtain, for any fonction f positive on $\mathcal{I}$,

$$\int_0^1 f(x)dx = \int_0^1 \mathbf{H}_2^\ell[f](x)dx = e_2[f] \int_0^1 \psi_2(x)dx + O(\alpha)^\ell,$$

(with $\alpha < 1$) from which we deduce (when $\ell \to \infty$) the equality $e_2[f] = \int_0^1 f(x)dx$. The argument can be easily adapted to the Markov case from the relation

$$\int_0^1 [f_0(x) + f_1(x)]dx = \int_0^1 (1 \quad 1)\, \mathbf{U}_2^\ell[f](x)dx = e_2[f] \int_0^1 [\psi_{2,0} + \psi_{2,1}](x)dx + O(\alpha)^\ell. \qquad \blacksquare$$

4.7. Conclusion: Coming back to the Tauberian Theorem. In conclusion of this functional analysis study, we obtain the final result of this Section which will be basic when proving that different Diriclet series $G(s)$ fullfills the hypotheses of the Tauberian Theorem.

Theorem 1. *Let $\mathcal{H}$ be a set of LFTs for which conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, \sigma)$, $\mathcal{D}(\mathcal{H})$ and $\mathcal{Q}(\mathcal{H})$ hold. 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 $s = 2$. More precisely, there exists a neighborhood $\mathcal{R}$ of $s = 2$, an operator $\mathbf{R}_s$ and a complex function $\ell(s)$ with $\ell(2) = -1/\lambda'(2)$, both analytic in $\mathcal{R}$ such that, for any $s \in \mathcal{R}$,*

$$(I - \mathbf{H}_s)^{-p} = \frac{1}{(s-2)^p}\, \ell(s)^p\, \mathbf{P}_s + \mathbf{R}_s^p,$$

(Here $\mathbf{P}_s$ is the dominant projector). Near $s=2$, one has, for any function f positive on $\mathcal{J}$, and any $x \in \mathcal{J}$,

$$(I-\mathbf{H}_s)^{-p}[f](x) \sim \frac{1}{(s-2)^p}\,(\frac{-1}{\lambda'(2)})^p\,\psi_2(x)\int_0^1 f(x)dx, \tag{43}$$

where ψ_2 is the dominant eigenfunction of $\mathbf{H}_2$ defined by the normalization condition $\int_0^1 \psi_2(x)dx = 1$. Under conditions $\mathcal{C}(\mathcal{V},\mathcal{U},\sigma)$, $\mathcal{D}(\mathcal{U})$ and $\mathcal{Q}(\mathcal{U})$, the same is true for the powers $(I-\mathbf{U}_s)^{-p}$ of the quasi-inverse, and (43) is replaced by

$$(I-\mathbf{U}_s)^{-p}[f](x) \sim \frac{1}{(s-2)^p}\,(\frac{-1}{\lambda'(2)})^p\,\psi_2(x)\int_0^1 [f_0(x)+f_1(x)]dx,$$

where ψ_2 is the dominant eigenfunction of $\mathbf{U}_2$ defined by the normalization condition $\int_0^1[\psi_{2,0}(x)+\psi_{2,1}(x)]dx = 1$.

Proof. First, from Properties 4, 5 and 6, the operators $(I-\mathbf{H}_s)$ and $(I-\mathbf{U}_s)$ are invertible on the punctured plane $\{\Re(s)=2, s\neq 2\}$.
Second, we consider a (complex) neighborhood of $s=2$. One deduces from (31) expressions for the powers of the quasi–inverse

$$(I-\mathbf{H}_s)^{-p} = \frac{\mathbf{P}_s}{(1-\lambda(s))^p} + (I-\mathbf{Q}_s)^{-p}.$$

Since the spectral radius of $\mathbf{Q}_s$ is strictly than 1, $\mathbf{R}_s := (I-\mathbf{Q}_s)^{-1}$ is analytic there. Now, using the derivability of $s \longrightarrow \lambda(s)$ at $s=2$ and the equality $\lambda(2)=1$, we let

$$\ell(s) = \frac{s-2}{1-\lambda(s)} \qquad \text{so that} \quad \ell(2) = \frac{-1}{\lambda'(2)},$$

and ℓ is analytic near $s=2$. All these facts are also true in the Markov case. Now, the residues at $s=2$ are easily evaluated from special values at $s=2$. ∎

5 Average Analysis of the Algorithms

Now, we come back to the analysis of the three algorithms, and we successively study the Odd Algorithm, the Ordinary Algorithm, and finally the Even Algorithm. We explain how the method of inducing, that is classical in theory of dynamical systems, can be applied to analysing the Even Algorithm, as well as the Subtractive GCD Algorithm. We compare the Even Algorithm and the Subtractive GCD Algorithm. We then give an interpretation of our results in terms of continuous (and sometimes random) dynamical systems, that explains how the constants involved in our analysis can be related to entropy of these dynamical systems. We finish in describing a general framework for further analyses.

5.1. Analysis of the Odd Algorithm The dynamical generic operator is in this case

$$\mathbf{O}_s[f] := \sum_{\substack{c \text{ odd}\geq 3, \varepsilon=\pm 1,\\ k\geq 1, c\geq 2^k+1}} (\frac{1}{c+\varepsilon x})^s\, f(\frac{2^k}{c+\varepsilon x}),$$

and, since the first step uses solely elements of $\mathcal{O}_0$, the operator relative to the first step is

$$\mathbf{O}_{s,0}[f] := \sum_{\substack{c \text{ odd}\geq 1, \varepsilon=\pm 1,\\ \varepsilon=1 \text{ for } c=1}} (\frac{1}{c+\varepsilon x})^s\, f(\frac{1}{c+\varepsilon x}).$$

Then, the double generating function $S(w,s)$ associated to the Odd Algorithm satisfies

$$S(w,s) = w(I - w\mathbf{O}_s)^{-1} \circ \mathbf{O}_{s,0}\ [1](0),$$

and the Dirichlet series $F(s)$ and $G(s)$ are

$$F(s) := (I - \mathbf{O}_s)^{-1} \circ \mathbf{O}_{s,0}[1](0) \qquad G(s) := (I - \mathbf{O}_s)^{-2} \circ (2I - \mathbf{O}_s) \circ \mathbf{O}_{s,0}[1](0).$$

Now, the operator $\mathbf{O}_s$ satisfies all the hypotheses of Theorem 1, and properties of the powers of the quasi–inverse $(I - \mathbf{O}_s)^{-1}$ can be transfered to $F(s)$ and $G(s)$, that fullfill the hypotheses of the Tauberian Theorem. They decompose as

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

where $C(s)$ and $D(s)$ are analytic at $s = 2$. Constants A and B involve the spectral dominant objects of $\mathbf{O}_s$, i.e., dominant eigenvalue function $\lambda(s)$ and dominant eigenvector ψ_s, under the form

$$A = \frac{-1}{\lambda'(2)}\,\psi_2(0)\int_0^1 \mathbf{O}_{2,0}[1](x)dx \qquad B = \frac{1}{\lambda'(2)^2}\psi_2(0)\int_0^1 (2I - \mathbf{O}_2) \circ \mathbf{O}_{2,0}[1](x)dx. \tag{44}$$

Remark the equality

$$\int_0^1 (2I - \mathbf{O}_2) \circ \mathbf{O}_{2,0}[1](x)dx = \int_0^1 \mathbf{O}_{2,0}[1](x)dx = 1.$$

The comparison between expressions of the residue of F obtained in (44) and (28) proves

$$\frac{1}{\lambda'(2)}\psi_2(0) = \frac{1}{2\widetilde{\zeta}(2)} = \frac{2}{\pi^2}, \quad \text{i.e.} \quad \frac{B}{A} = \frac{-1}{\lambda'(2)} = \frac{2}{\pi^2\psi_2(0)}.$$

We obtain our first main result

Theorem 2. *The average number O_N or $\widetilde{O}_N$ of division steps performed by the Odd Algorithm on the set the set $\Omega_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N\}$, or on the set $\widetilde{\Omega}_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N, \gcd(u,v) = 1\}$ satisfies*

$$O_N \sim \widetilde{O}_N \sim \frac{2}{\pi^2\psi_2(0)}\ \log N$$

where ψ_2 is the fixed point of the operator $\mathbf{O}_2$

$$\mathbf{O}_2[f] := \sum_{\substack{c\ \text{odd} \ge 3,\, \varepsilon = \pm 1, \\ k \ge 1,\, c \ge 2^k + 1}} \left(\frac{1}{c + \varepsilon x}\right)^2\ f\left(\frac{2^k}{c + \varepsilon x}\right),$$

defined by the normalization condition $\int_0^1 \psi_2(t)dt = 1$.

5.2. Analysis of the Ordinary Algorithm. Now, the operator $\mathbf{U}_s$ satisfies all the hypotheses of the Theorem, and properties of the powers of the quasi–inverse $(I - \mathbf{U}_s)^{-1}$ can be transfered to $F(s)$ and $G(s)$, that fullfill the hypotheses of the Tauberian Theorem. They decompose as

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

where $C(s)$ and $D(s)$ are analytic at $s = 2$, and constants A and B involve the second component of the dominant eigenvector $\phi_2 = (\phi_{2,0}, \phi_{2,1})$ of $\mathbf{U}_2$, defined by the normalization condition $\int_0^1 [\phi_{2,0}(t) + \phi_{2,1}(t)]dt = 1$,

$$A = \frac{-1}{\lambda'(2)}\,\phi_{2,1}(0) \qquad B = \frac{1}{\lambda'(2)^2}\phi_{2,1}(0). \tag{45}$$

The comparison of expresssions of the residue of F at $s=2$ obtained in (28) and (45) at $s=2$ proves

$$\frac{1}{\lambda'(2)}\phi_{2,1}(0) = \frac{1}{2\widetilde{\zeta}(2)} = \frac{2}{\pi^2}, \quad \text{i.e.} \quad \frac{B}{A} = \frac{-1}{\lambda'(2)} = \frac{2}{\pi^2\phi_{2,1}(0)}.$$

We then obtain our second main result

Theorem 3. *The average number U_N or $\widetilde{U}_N$ of division steps performed by the Ordinary Algorithm on the set the set $\Omega_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N\}$, or on the set $\widetilde{\Omega}_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u < v \le N, \gcd(u,v)=1\}$ satisfies*

$$U_N \sim \widetilde{U}_N \sim \frac{2}{\pi^2\phi_{2,1}(0)}\ \log N$$

where $\phi_{2,1}$ is the second component of the fixed point $\phi_2 = (\phi_{2,0}, \phi_{2,1})$ of the "matrix "operator $\mathbf{U}_2$,

$$\mathbf{U}_2 \begin{pmatrix} f_0 \\ f_1 \end{pmatrix}(x) = \begin{pmatrix} \displaystyle\sum_{c\,\text{even}\,\ge 2} (\frac{1}{c+x})^2\ f_0(\frac{1}{c+x}) + \sum_{\substack{c\,\text{even}\,\ge 2,\\ k\ge 1, c\ge 2^k}} (\frac{1}{c+x})^2\ f_1(\frac{2^k}{c+x}) \\ \displaystyle\sum_{c\,\text{odd}\,c\ge 1} (\frac{1}{c+x})^2\ f_0(\frac{1}{c+x}) + \sum_{\substack{c\,\text{odd}\,\ge 3,\\ k\ge 1, c\ge 2^k+1}} (\frac{1}{c+x})^2\ f_1(\frac{2^k}{c+x}) \end{pmatrix}$$

defined by the normalization condition $\int_0^1 [\phi_{2,0}(t) + \phi_{2,1}(t)]dt = 1$.

5.3. Analysis of the Even Algorithm. The dynamical generic operator is in this case

$$\mathbf{E}_s[f] := \sum_{\substack{c\ge 1\\ \varepsilon=\pm 1}} (\frac{1}{2c+\varepsilon x})^s f(\frac{1}{2c+\varepsilon x}),$$

and, since the last step uses solely elements of $\mathcal{B}^+$, the operator relative to the last step is

$$\mathbf{B}_s^+[f] := \sum_{c\ge 1} (\frac{1}{2c+x})^s f(\frac{1}{2c+x}).$$

Then, the double generating function $S(w,s)$ associated to the Even Algorithm satisfies

$$S(w,s) = w\mathbf{B}_s^+\ (I - w\mathbf{E}_s)^{-1}\ [1](1).$$

Unfortunately, the operator $\mathbf{E}_s$ is not well–behaved, but only because of the branch a defined by $a(x) := 1/(2-x)$ which possesses an indifferent point, i.e., a fixed point where the absolute value of the derivative equals 1: the point 1 satisfies $a(1) = 1$ and $a'(1) = 1$. In this case, the component operator relative to this LFT,

$$\mathbf{A}_s[f](x) = (\frac{1}{2-x})^s f(\frac{1}{2-x}),$$

is not well–behaved, since the condition $\mathcal{C}(\mathcal{V}, a)$ is not satisfied, for any open disk $\mathcal{V}$ that contains interval $\mathcal{I}$. However, all the other conditions $\mathcal{C}(\mathcal{V}, b)$ are satisfied for elements $b \in \mathcal{E}$ distinct of a. Moreover, if we denote by $\mathcal{B}$ the set $\mathcal{E}\setminus\{a\}$ and by $\mathcal{B}^+$ the set relative to the last step, the stronger conditions $\mathcal{C}(\mathcal{V}, \mathcal{B}, 1)$ and $\mathcal{C}(\mathcal{V}, \mathcal{B}^+, 1)$ hold, so that the operator relative to the set $\mathcal{B}$,

$$\mathbf{B}_s[f] := \sum_{\substack{c\ge 1\\ \varepsilon=\pm 1\\ (c,\varepsilon)\ne(1,-1)}} (\frac{1}{2c+\varepsilon x})^s f(\frac{1}{2c+\varepsilon x}),$$

is compact, as well as the operator $\mathbf{B}_s^+$.
Since the last LFT is an element of $\mathcal{B}^+$, the set of LFT's used in the Even Algorithm is exactly $\mathcal{E}^* \mathcal{B}^+$ that can also be written as

$$(a^*\mathcal{B})^* \ a^*\mathcal{B}^+.$$

In this way, we use the method of inducing [5], we let

$$\mathcal{M} := a^*\mathcal{B}, \qquad \mathcal{M}^+ := a^*\mathcal{B}^+ \tag{46}$$

and we introduce two operators associated to $\mathcal{M}$ and $\mathcal{M}^+$, as Prellberg and Slawny suggest it in another context [33],

$$\mathbf{M}_{s,w} = \sum_{k\geq 0} w^{k+1} \ \mathbf{B}_s \ \mathbf{A}_s^k, \qquad \mathbf{M}_{s,w}^+ = \sum_{k\geq 0} w^{k+1} \ \mathbf{B}_s^+ \ \mathbf{A}_s^k,$$

so that the double generating function $S(s,w)$ can be expressed as

$$S(w,s) = \mathbf{M}_{s,w}^+ \ (I - \mathbf{M}_{s,w})^{-1}[1](1). \tag{47}$$

And, now, these operators are well-behaved:

Lemma. *The conditions $\mathcal{C}(\mathcal{V},\mathcal{M},1)$ and $\mathcal{C}(\mathcal{V},\mathcal{M}^+,1)$ hold, for some suitable open disk $\mathcal{V}$. The operators $\mathbf{M}_{s,w}$, $\mathbf{M}_{s,w}^+$ are well-defined for $|w|\leq 1$ and $\Re(s) > 1$, and, for these values of (s,w), they act on the set $\mathcal{A}_\infty(\mathcal{V})$ and are compact.*

Proof. Each element of $\mathcal{M}$ is of form $h = a^k \circ b$, with $b \in \mathcal{B}$ and it satisfies

$$h(x) = \frac{k(2c+\varepsilon x) - (k-1)}{(k+1)(2c+\varepsilon x) - k} \qquad \text{with } (c,\varepsilon) \neq (1,-1).$$

Since

$$|h'(x)| \leq (\frac{1}{2(k+1)(c-1)})^2 \qquad \text{for } c\geq 2, \qquad |h'(x)| \leq (\frac{1}{k+2})^2 \qquad \text{for } (c,\varepsilon) = (1,1),$$

the condition $\mathcal{C}(\mathcal{V},\mathcal{M},1)$ holds. ∎

When derivating (47) with respect to w, and/or evaluating at $w=1$, one obtains other forms for the Dirichlet series $F(s)$ and $G(s)$ which now involve the operators

$$\mathbf{M}_s := \mathbf{M}_{s,1} \qquad \mathbf{M}_s^+ := \mathbf{M}_{s,1}^+ \qquad \mathbf{N}_s := \frac{d}{dw}\mathbf{M}_{s,w}|_{w=1}, \qquad \mathbf{N}_s^+ := \frac{d}{dw}\mathbf{M}_{s,w}^+|_{w=1}$$

under the form

$$F(s) = \mathbf{M}_s^+ \ (I-\mathbf{M}_s)^{-1}[1](1), \qquad G(s) = \mathbf{N}_s^+ \circ (I-\mathbf{M}_s)^{-1}[1](1) + \mathbf{M}_s^+ \circ (I-\mathbf{M}_s)^{-1} \circ \mathbf{N}_s \circ (I-\mathbf{M}_s)^{-1}[1](1).$$

The following lemma describes the dominant spectral objects of $\mathbf{M}_2$ and the behaviour of $\mathbf{N}_s$ near $s=2$.

Lemma. *The operator $\mathbf{M}_2$ has a dominant eigenvalue equal to 1, and the dominant eigenvector is*

$$\psi_2(x) = \frac{1}{\log 3}[\frac{1}{3-x} + \frac{1}{1+x}].$$

The operators $\mathbf{N}_s$, $\mathbf{N}_s^+$ have a simple pole at $s=2$. More precisely, the function $J(s) := \int_0^1 \mathbf{N}_s[\psi_2](x)dx$ has a simple pole at $s=2$ with a residue equal to $1/\log 3$.

Proof. One has

$$\mathbf{M}_{s,w} \circ (I - w \ \mathbf{A}_s) = w \ \mathbf{B}_s \qquad \mathbf{A}_s + \mathbf{B}_s = \mathbf{E}_s.$$

At $s=2$, the operator $\mathbf{E}_2$ has an eigenvalue equal to 1, and the associated eigenvector is

$$\phi_2(x) := \frac{1}{1-x} + \frac{1}{1+x}.$$

This function ϕ_2 satisfies $\mathbf{E}_2[\phi_2] = \mathbf{A}_2[\phi_2] + \mathbf{B}_2[\phi_2] = \phi_2$. If we let

$$g := \mathbf{B}_2[\phi_2] = (I - \mathbf{A}_2)[\phi_2],$$

this function satisfies

$$\mathbf{M}_2[g] = \mathbf{M}_2 \circ (I - \ \mathbf{A}_2)[\phi_2] = \mathbf{B}_2[\phi_2] = g,$$

and g is an eigenvector relative to the eigenvalue 1. With normalization condition, we obtain the expression of ψ_2.

Consider an integral of the form

$$K = \int_0^1 \mathbf{N}_s[f]dx = \int_0^1 \sum_{\substack{k\geq 0,\\ b\in\mathcal{B}}} (k+1)\ \left(\frac{1}{D[a^k \circ b](x)}\right)^s f\circ a^k \circ b(x)dx.$$

Since the intervals $b(\mathcal{I})$ for $b\in\mathcal{B}$ form a partition of $[0,1/2]$, we let $u := b(x)$, and we use multiplicativity of D. We obtain

$$K = \int_0^{1/2} \sum_{k\geq 0} \frac{(k+1)u^{s-2}}{D[a^k](u)^s} f\circ a^k(u)du.$$

We consider now, near $s=2$, each of the integrals

$$J_1(s) := \int_0^1 \mathbf{N}_s[\frac{1}{1+x}]dx \qquad J_2(s) := \int_0^1 \mathbf{N}_s[\frac{1}{3-x}]dx.$$

Near $s=2$, they are both equivalent to

$$\int_0^{1/2} \sum_{k\geq 1} \frac{k}{2D[a^k](u)^s} du = \frac{1}{2}\sum_{k\geq 1}\int_0^{k/2} \frac{dv}{(k+1-v)^s} \simeq \frac{1}{2}\zeta(s-1),$$

so that the residue of $J(s)$ at $s=2$ equals $1/\log 3$. ■

Now, the operator $\mathbf{M}_s$ satisfies all the hypotheses of Theorem 1, and properties of the powers of the quasi–inverse $(I-\mathbf{M}_s)^{-1}$ can be transfered to $F(s)$ and $G(s)$, that fullfill 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 $s=2$. Constants A and B involve the spectral dominant objects of $\mathbf{M}_s$, i.e., dominant eigenvalue function $\lambda(s)$ and dominant eigenvector ψ_s, under the form

$$A = \frac{-1}{\lambda'(2)}\mathbf{M}_2^+[\psi_2](1) \qquad B = \frac{1}{\lambda'(2)^2}\mathbf{M}_2^+[\psi_2](1)\frac{1}{\log 3} \tag{48}$$

Remark the equality

$$\mathbf{M}_2^+[\psi_2](1) = \frac{1}{2}\mathbf{M}_2[\psi_2](1) = \frac{1}{2\log 3}.$$

The comparison between expressions of the residue of F obtained in (48) and (28) proves

$$\frac{1}{\lambda'(2)}\frac{1}{2\log 3} = \frac{1}{2\widetilde{\zeta}(2)} = \frac{2}{\pi^2}, \quad \text{i.e.} \quad \frac{B}{2A} = \frac{-1}{2\lambda'(2)\log 3} = \frac{1}{\pi^2},$$

so that the average number of iterations on $\Omega_N$is asymptotically equal to $B/(2A)\log^2 N$.

We can also study another parameter that is the number of steps that use LFTs of $\mathcal{B}$. These steps are related to so–called "large" pseudo–divisions of the form

$$v = 2cu + \varepsilon s \qquad \text{with } (c,\varepsilon) \neq (1,-1).$$

Since the operator $w\mathbf{M}_s$ only takes into account the steps that use elements of $\mathcal{B}$, the double generating function $S(s,w)$ associated to the parameter "number of large divisions used" can be expressed now as

$$S(w,s) = w\mathbf{M}_s^+ \, (I - w\mathbf{M}_s)^{-1}[1](1).$$

We then obtain our third result:

Theorem 4. *The average number E_N or $\widetilde{E}_N$ of division steps performed by the Even Algorithm on the set the set $\Omega_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u \le v \le N\}$, or on the set $\widetilde{\Omega}_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u \le v \le N, \gcd(u,v) = 1\}$ satisfies*

$$U_N \sim \widetilde{U}_N \sim \frac{1}{\pi^2}\log^2 N$$

The average number M_N or $\widetilde{M}_N$ of large division steps (i.e. divisions of the form $v = 2cu + \varepsilon r$ with $(c,\varepsilon) \neq (1,-1)$) performed by the Even Algorithm on the set the set $\Omega_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u \le v \le N\}$, or on the set $\widetilde{\Omega}_N := \{(u,v);\ u, v\,\text{odd},\ 0 < u \le v \le N, \gcd(u,v) = 1\}$ satisfies

$$M_N \sim \widetilde{M}_N \sim \frac{4\log 3}{\pi^2}\log N.$$

5.4. Comparing the Even algorithm and the subtractive GCD algorithm. Here, we explain why the Even Algorithm has the same kind of behaviour as the Subtractive GCD algorithm (here called the SGCD algorithm). We also provide alternative analyses of the SGCD algorithm and the classical GCD algorithm. First, we recall the SGCD algorithm.

Subtractive GCD Algorithm.

Input: a pair (u,v) satisfying $0 < u < v$,
Output: $G(u,v)$

begin
While $u \neq 0$ **do**
begin
  While $u < v$ **do**
    $v := v - u$;
  Exchange u and v;
end;
$G(u,v) := v$;
end.

We are interested here in evaluating the number of subtractions and the number of exchanges. Clearly, the number of exchanges is exactly equal to the number of steps performed by the classical GCD algorithm. Instead of the integer pair (u,v), one considers the rational $x = u/v$. The whole process finishes with the value 0, and uses only two LFTs, according as the subtraction is followed by an exchange or not:

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

The last step is particular, since it always use the LFT q, and, just before, it uses at least once LFT p. Finally, when the SGCD Algorithm runs on the integer pair (u,v), it finishes with the value 0, and it decomposes the rational $x = u/v$ in a unique way as

$$x = h_1 \circ h_2 \circ \ldots \circ h_\ell(0) \qquad \text{with } h_i \in \{p,q\} \text{ for } 1 \le i \le \ell-2,\ h_{\ell-1} = p,\ h_\ell = q. \tag{49}$$

Finally, the set of LFTs used by the algorithm is exactly $(p^*q)^*\, pp^*q$, so that we introduce

$$\mathcal{G} := p^*q = \{h \mid h(x) = \frac{1}{k+x}, k \ge 1\} \qquad \mathcal{G}^+ := pp^*q = \{h \mid h(x) = \frac{1}{k+x}, k \ge 2\}. \tag{50}$$

The Ruelle operator $\mathbf{T}_s$ a priori associated to the SGCD algorithm is $\mathbf{T}_s = \mathbf{P}_s + \mathbf{Q}_s$, where the operators $\mathbf{P}_s$ and $\mathbf{Q}_s$ are respectively associated to p and q,

$$\mathbf{P}_s[f](x) := (\frac{1}{1+x})^s f(\frac{x}{1+x}) \qquad \mathbf{Q}_s[f](x) := (\frac{1}{1+x})^s f(\frac{1}{1+x}).$$

However, the LFT p is not well-behaved because it has an indifferent point at $x = 0$ $(p(0) = 0, p'(1) = -1)$, so that the operator $\mathbf{P}_s$ has not good properties. In this case, the principle of inducing replaces the operator $\mathbf{T}_s$ by the operators

$$\mathbf{G}_{s,w} = \sum_{k \ge 0} w^{k+1}\ \mathbf{Q}_s\ \mathbf{P}_s^k, \qquad \mathbf{G}^+_{s,w} = \sum_{k \ge 1} w^{k+1}\ \mathbf{Q}_s\ \mathbf{P}_s^k,$$

so that the double generating function $S(s,w)$ associated to the parameter "number of subtractions" can be expressed as

$$S(w,s) = \mathbf{G}^+_{s,w}\ (I - \mathbf{G}_{s,w})^{-1}[1](0). \tag{51}$$

And, now, the operators $\mathbf{G}_{s,w}$, $\mathbf{G}^+_{s,w}$ are well-behaved: They are well-defined for $|w| \le 1$ and $\Re(s) > 1$. The conditions $\mathcal{C}(\mathcal{V},\mathcal{G},1)$, $\mathcal{C}(\mathcal{V},\mathcal{G}^+,1)$ hold, for some suitable open disk $\mathcal{V}$, as well as conditions $\mathcal{D}$ and $\mathcal{Q}$. For these values of (s,w), the operator $\mathbf{G}_{s,w}$, $\mathbf{G}^+_{s,w}$ act on the set $\mathcal{A}_\infty(\mathcal{V})$ and are compact.
When derivating (51) with respect to w, and/or evaluating at $w = 1$, one obtains other forms for the Dirichlet series $F(s)$ and $G(s)$ which involve now the operators

$$\mathbf{G}_s := \mathbf{G}_{s,1}, \qquad \mathbf{G}^+_s := \mathbf{G}^+_{s,1} \qquad \mathbf{K}_s := \frac{d}{dw}\mathbf{G}_{s,w}|_{w=1}, \qquad \mathbf{K}^+_s := \frac{d}{dw}\mathbf{G}^+_{s,w}|_{w=1},$$

under the form

$$F(s) = \mathbf{G}^+_s \circ\ (I - \mathbf{G}_s)^{-1}[1](0), \qquad G(s) = \mathbf{K}^+_s \circ\ (I - \mathbf{G}_s)^{-1}[1](0) + \mathbf{G}^+_s \circ (I - \mathbf{G}_s)^{-1} \circ \mathbf{K}_s \circ (I - \mathbf{G}_s)^{-1}[1](0).$$

The operator $\mathbf{G}_s$ is the famous operator associated to the classical continued fraction algorithm, so–called the Ruelle-Mayer operator

$$\mathbf{G}_s[f](x) = \sum_{k \ge 1} (\frac{1}{k+x})^s f(\frac{1}{k+x}).$$

The dominant spectral objects of $\mathbf{G}_2$ are well-known; for $s = 2$, $\mathbf{G}_2$ is a density transformer and

$$\psi_2(x) = \frac{1}{\log 2}\frac{1}{1+x} \qquad \frac{-1}{\lambda'(2)} = \frac{12}{\pi^2 \psi_2(0)} = \frac{12 \log 2}{\pi^2}.$$

Furthermore, it is easy to obtain the behaviour of $\mathbf{K}_s$ near $s = 2$: it has a simple pole at $s = 2$ and the function $J(s) := \int_0^1 \mathbf{K}_s[\psi_2](x)dx$ has a simple pole at $s = 2$ with a residue equal to $1/\log 2$.

The operator $w\mathbf{G}_s$ only takes into account the steps with exchanges, so that the double generating function $S(s,w)$ associated to the parameter "number of exchanges" can be expressed now as

$$S(w,s) = w\mathbf{G}^+_s\ (I - w\mathbf{G}_s)^{-1}[1](0).$$

We then obtain an easy alternative proof of classical results of Yao and Knuth [48], Heilbronn [18] and Dixon [9].

Theorem 5. *The average number T_N or $\widetilde{T}_N$ of subtractions performed by the subtractive Algorithm on the set the set $\{(u,v);\ 0 < u \le v \le N\}$, or on the set $\{(u,v);\ 0 < u \le v \le N, \gcd(u,v) = 1\}$ satisfies*

$$T_N \sim \widetilde{T}_N \sim \frac{6}{\pi^2} \log^2 N.$$

The average number G_N or $\widetilde{G}_N$ of exchanges performed by the classical Euclidean Algorithm on the set the set $\{(u,v);\ 0 < u \le v \le N\}$, or on the set $\{(u,v);\ 0 < u \le v \le N, \gcd(u,v) = 1\}$ satisfies

$$G_N \sim \widetilde{G}_N \sim \frac{12 \log 2}{\pi^2} \log N.$$

We now insist on the similar behaviour of the SGCD algorithm and the Even Algorithm. In both algorithms, there are two kinds of steps:

–good steps (exchanges in the SGCD Algorithm, large divisions in the Even Algorithm),

–bad steps (subtractions in the SGCD Algorithm, small divisions in the Even Algorithm).

In both cases, bad steps are related to transformations (p in the SGCD Algorithm, a in the Even Algorithm) that are not well-behaved. They have an indifferent point, so that the algorithms spend most of iterations in performing them. On the contrary, good steps are related to well-behaved transformations and the algorithms spend very few iterations in performing them.

The Non-Euclidean Algorithm is introduced by Vardi [47]. It uses the set

$$\mathcal{V} := \{h(x) := \frac{1}{k - x} \qquad k \ge 2\}$$

that also contains the "bad" LFT a. Vardi explains why the average number of iterations is of order $\log^2 N$ by comparing it to the SGCD algorithm. Our methods give an alternative proof of this fact by comparing it to the Even Algorithm.

5.5. The algorithms viewed as (random) dynamical systems. At $s = 2$, the Ruelle operators can be viewed as density transformers. We now make precise the dynamical systems to which they are associated: They are deterministic for the Even Algorithm, random for the Odd Algorithm, both random and markovian for the Ordinary Algorithm.

The Even pseudo–division is also defined for reals, so that the Even Algorithm can be extended to reals entries. Of course, it does not compute anymore the Jacobi symbol, but it creates a continued fraction with even quotients whose properties are studied by several authors [26], [37].

Extended Even Algorithm.

Input: $x \in]0,1[$.

Output: The even continued fraction expansion of x.

begin

While $x \neq 1$ **do**

  $b :=$ the even integer part of $\frac{1}{x}$; $\varepsilon :=$ Sign $[\frac{1}{x} - b]$; $x := \varepsilon\ [\frac{1}{x} - b]$

  Write (b, ε);

end.

We consider a density f on $\mathcal{I}$ before one generic iteration, and g is the density after the iteration of the algorithm; then, each term

$$\frac{1}{D[h](x)^2} f \circ h(x)dx = |h'(x)| f \circ h(x)dx$$

represents the part of $g(x)$ "brought" by the antecedent $h(x)$ of point x, and finally g exactly equals $\mathbf{E}_2[f]$. Then $\mathbf{E}_2$ is the density transformer associated to the Even Algorithm.

The other two pseudo-divisions are related to dyadic valuation, so that odd or ordinary continued fractions expansions are not defined for irrationals numbers. However, one can define random continued fraction in these two cases when choosing at random the dyadic valuation k of a real number, according to the law $\Pr[k = d] = 2^{-d}$ (for $d \geq 1$), that extends the natural law on even integers.

We first describe the random process associated to the Odd algorithm that produces a random continued fraction with odd quotients and numerators with powers of two.

Random–Odd Algorithm.

Input: $x \in]0, 1[$.

Output: The random continued fraction expansion of x with odd quotients.

begin

While $x \neq 0$ **do**

  begin

  If $x = \dfrac{u}{v}, (u, v \text{ odd})$ **then** perform the odd–euclidean division: $v = bu + \varepsilon\, 2^k\, s$; $x := \dfrac{s}{u}$

  else

    $b :=$ the odd integer part of $\dfrac{1}{x}$; $\varepsilon := \text{Sign}\,[\dfrac{1}{x} - b]$; $y := \varepsilon\, [\dfrac{1}{x} - b]$

    Choose k at random according to the law $\Pr[k = d] = 2^{-d}$ (for $d \geq 1$); $x := \dfrac{y}{2^k}$

  Write $(b, \varepsilon, 2^k)$;

  end;

end.

In this manner, we choose the LFT of determinant 2^k with probability 2^{-k}, and, for LFT's of determinant 2^k, the quantity

$$\frac{1}{D[h](x)^2} f \circ h(x)dx = \frac{|h'(x)|}{2^k} f \circ h(x)dx$$

represents exactly a random change of variables. Now, given a density f on $\mathcal{I}$ before one iteration, the function $\mathbf{O}_2[f]$ is the density after one iteration, so that $\mathbf{O}_2$ is a density transformer associated to this random algorithm.

In the same vein, the random process associated to the Ordinary algorithm produces a random continued fraction with ordinary quotients and numerators with powers of two.

Random–Ordinary Algorithm.

Input: $x \in]0, 1[$.

Output: The random continued fraction expansion of x with ordinary quotients.

begin

While $x \neq 0$ **do**

  begin

  If $x = \dfrac{u}{v}, (u, v \text{ odd})$ **then** perform the ordinary–euclidean division: $v = bu + 2^k\, s$; $x := \dfrac{s}{u}$

else

$b :=$ the integer part of $\frac{1}{x}$; $y := [\frac{1}{x} - b]$

If b is even **then** $k = 0$; $x := y$

else choose k at random according to the law $\Pr[k = d] = 2^{-d}$ (for $d \geq 1$); $x := \frac{y}{2^k}$

Write $(b, 2^k)$;

end;

end.

Now, given a density f on $\mathcal{I}$ before one iteration, the function $(1,1)\mathbf{U}_2[f, 0]$ is the density after one iteration, so that $\mathbf{U}_2$ is a density transformer associated to this random algorithm.

5.6. Relations with entropy of dynamical systems. Finally, this work deals with four sets of LFTs, the sets $\mathcal{O}, \mathcal{U}, \mathcal{M}, \mathcal{G}$ defined in (4), (7), (9), (8), (10), (46), (50). The first two sets $\mathcal{O}, \mathcal{U}$ are related to the Odd and Ordinary Algorithms. The third one $\mathcal{M}$ intervenes in the induced algorithm related to the Even Algorithm, where successive steps using LFT a are "gathered". The fourth one $\mathcal{G}$ is related to the classical GCD Algorithm, but it also intervenes in the induced algorithm related to the Subtractive GCD Algorithm, where successive steps using subtraction are "gathered" .

All these sets $\mathcal{O}, \mathcal{U}, \mathcal{M}, \mathcal{G}$ satisfy all the conditions $\mathcal{C}, \mathcal{D}, \mathcal{Q}$ defined in paragraphs 4.2, 4.3, 4.6, so that the Ruelle operators $\mathbf{O}_s$, $\mathbf{U}_s$, $\mathbf{M}_s$ ang $\mathbf{G}_s$ have "good " properties, according to Properties 2, 3, 4, 5, 6 and Theorem 1. At $s = 2$, they are density transformers, they all have dominant eigenvalues $\lambda_O(s)$, $\lambda_U(s)$, $\lambda_M(s)$, $\lambda_G(s)$ that equal 1 at $s = 2$. In each case, the values $-2\lambda'(2)$ equal the entropy h of the (possibly random) dynamical system associated to the (possibly random) algorithm,

$$-2\lambda'_O(2) = h(\mathcal{O}), \qquad -2\lambda'_U(2) = h(\mathcal{U}) \qquad -2\lambda'_M(2) = h(\mathcal{M}), \qquad -2\lambda'_G(2) = h(\mathcal{G}).$$

We have shown, in each case, how the values $-\lambda'(2)$ intervene in the average number of iterations of the (induced) algorithms. So, we get alternative values for multiplicative constants of Theorems 2, 3, 4, 5.

$$O_N \sim \frac{2}{h(\mathcal{O})} \log N \qquad U_N \sim \frac{2}{h(\mathcal{U})} \log N \qquad M_N \sim \frac{2}{h(\mathcal{M})} \log N \qquad G_N \sim \frac{2}{h(\mathcal{G})} \log N.$$

In this paper, we have proven four instances of the general result that we now state a bit informally:

Theorem 6. *Let $\mathcal{H}$ be a set of LFTs for which conditions $\mathcal{C}(\mathcal{V}, \mathcal{H}, 1), \mathcal{D}(\mathcal{H}), \mathcal{Q}(\mathcal{H})$ hold. Then, the Euclidean Algorithm associated to the (possibly random) dynamical system defined by $\mathcal{H}$ performs an average number of steps on rationals of numerator and denominator less than N which satisfies*

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

5.8. A general framework for further analyses. Our methods can be used for analysing other parameters of the continued fractions: the quotients, the powers of two, as we did it in [46] for the Binary continued fraction. These methods can also be easily adapted to evaluate the other moments of the random variables "number of iterations".

We obtain the following results: All the random variables "number of iterations" for which the average is of order $\log N$, have their moments of order ℓ that are of order $\log^\ell N$, with a constant which involves ℓ-th power of entropy. On the contrary, the random variables for which the average is of order $\log^2 N$ have their moments of order $\ell \geq 2$ which are of order $N^{\ell-1}$.

Acknowledgements. Brigitte Vallée wishes to thank Pierre Ducos for earlier discussions of 1992 related to this work, Thomas Prellberg for his explanations of the inducing method, and Dieter Mayer for his introduction to random dynamical systems.

References

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

[2] 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.

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

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

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

[6] 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

[7] 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

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

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

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

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

[12] 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

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

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

[15] Goldwasser, S. and Micali, S. Probabilistic encryption, *Journal of Computer and Systems Science* 28 (1984) 270–299.

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

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

[18] 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.

[19] Hensley, D. The distribution of badly approximable rationals and continuants with bounded digits II, *Journal of Number Theory* 34 (1990) 293-334

[20] Hensley, D. The Hausdorff dimensions of some continued fraction Cantor sets, *Journal of Number theory*33 (1989) pp 182-198

[21] Hensley, D. Continued Fraction Cantor sets, Hausdorff dimension, and functional analysis, *Journal of Number Theory* 40 (1992) pp 336-358.

[22] Hensley, D. A polynomial time algorithm for the Hausdorff dimension, *Journal of Number Theory* 58 (1996) pp 9–45.

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

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

[25] Kato, T. *Perturbation Theory for Linear Operators*, Springer-Verlag, 1980.

[26] Kraaikamp, C. and Lopes, A. The Theta group and the continued fraction expansion with even partial quotients. preprint, 1995

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

[28] Krasnoselskii, M. *Positive solutions of operator equations*, P. Noordhoff, Groningen, 1964.

[29] Lorch, E. R. *Spectral Theory*, Oxford University Press, New York, 1962.

[30] Mayer, D. H. On a ζ function related to the continued fraction transformation, *Bulletin de la Société Mathématique de France 104* (1976), 195–203.

[31] 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.

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

[33] 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

[34] Ruelle, D. *Thermodynamic formalism,* Addison Wesley (1978)

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

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

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

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

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

[40] Solovay, R. and Strassen, V. A fast Monte-Carlo test for primality, *SIAM J. Comp.* 6 (1977) 84-85; erratum 7 (1978) 118.

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

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

[43] Vallée, B. Algorithms for computing signs of 2×2 determinants: dynamics and average-case algorithms, Proceedings of the 8 th Annual European Symposium on Algorithms, ESA'97, pp 486–499, LNCS 1284, Springer Verlag.

[44] Vallée, B. Fractions continues à contraintes périodiques, Les Cahiers du GREYC 1997, Université de Caen, to appear in *Journal of Number Theory*.

[45] Vallée, B. Dynamical Sources in Information Theory: Fundamental intervals and Word Prefixes, Les cahiers du GREYC 1998, Université de Caen

[46] Vallée, B. Dynamics of the Binary Euclidean Algorithm: Functional Analysis and Operators., To appear in Algorithmica (1998), extended abstract in Proceedings of ANTS'98, LNCS.

[47] Vardi, I. Continued fractions, Preprint, chapter of a book in preparation.

[48] 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.