Sharp estimates for the main parameters of the Euclid Algorithm

Loïck Lhote, Brigitte Vallée

Sharp estimates for the main parameters of the Euclid Algorithm.

Loïck Lhote[1] and Brigitte Vallée[1]

GREYC, University of Caen, France
{loick.lhote, brigitte.vallee}@info.unicaen.fr

Abstract. We provide sharp estimates for the probabilistic behaviour of the main parameters of the Euclid algorithm, and we study in particular the distribution of the bit-complexity which involves two main parameters : digit–costs and length of continuants. We perform a "dynamical analysis" which heavily uses the dynamical system underlying the Euclidean algorithm. Baladi and Vallée [2] have recently designed a general framework for "distributional dynamical analysis", where they have exhibited asymptotic gaussian laws for a large class of digit–costs. However, this family contains neither the bit–complexity cost nor the length of continuants. We first show here that an asymptotic gaussian law also holds for the length of continuants at a fraction of the execution. There exist two gcd algorithms, the standard one which only computes the gcd, and the extended one which also computes the Bezout pair, and is widely used for computing modular inverses. The extended algorithm is more regular than the standard one, and this explains that our results are more precise for the extended algorithm. We prove that the bit–complexity of the extended Euclid algorithm asymptotically follows a gaussian law, and we exhibit the speed of convergence towards the normal law. We describe also conjectures [quite plausible], under which we can obtain an asymptotic gaussian law for the plain bit-complexity, or a sharper estimate of the speed of convergence towards the gaussian law.

1 Introduction

The Euclid algorithm computes the greatest common divisor (in short gcd) of u and v, with Euclidean divisions of the form $v = m \cdot u + r$ with $0 \leq r < u$. On an input (u, v), with $v_0 := v, v_1 := u$, it performs a sequence of Euclidean divisions

$$v_0 = m_1 \cdot v_1 + v_2, \quad \dots \quad v_i = m_{i+1} \cdot v_{i+1} + v_{i+2} \dots \quad v_{p-1} = v_p \cdot m_p + 0. \quad (1)$$

It stops when the remainder v_{p+1} is zero, and the last non-zero remainder v_p is the greatest common divisor d of u and v.

We wish to study the bit–complexity of the Euclid algorithm, i.e., the total number of binary operations performed during the execution of the Euclid algorithm. The (naive) bit–complexity of a Euclidean division $v = m \cdot u + r$ is $\ell(u) \cdot \ell(m)$, where $\ell(v)$ is the binary length of the integer v; it equals $\lfloor \lg v \rfloor + 1$, where lg denotes the logarithm in base 2. Then, the bit–complexity of the Euclid algorithm on the input (u, v) is

$$B(u,v) = \sum_{i=1}^{p} \ell(m_i) \cdot \ell(v_i), \qquad [\ p := P(u,v) \text{ is the number of iterations}] \quad (2)$$

The extended Euclid algorithm ouputs, at the same time, the Bezout pair (r, s) for which $d = rv + su$. It computes the sequence s_i defined by $s_0 = 0, \quad s_1 = 1, \quad s_i = s_{i-2} - s_{i-1} \cdot m_{i-1}, \quad 2 \leq i < p$. The last element s_p is the Bezout coefficient s. The bit–complexity D of this algorithm on (u, v) is

$$D(u,v) = \ell(m_p) \cdot \ell(v_p) + \sum_{i=1}^{p-1} \ell(m_i) \cdot [\ell(v_i) + \ell(s_i)]. \quad (3)$$

We introduce also a so–called "smoothed" version $\widetilde{D}, \widetilde{B}$ of costs D, B, where we replace the size $\ell(s_i), \ell(v_i)$ of s_i, v_i by their logarithms $\lg s_i, \lg v_i$,

$$\widetilde{D}(u,v) = \ell(m_p) \cdot \lg v_p + \sum_{i=1}^{p-1} \ell(m_i) \cdot [\lg v_i + \lg s_i], \qquad \widetilde{B}(u,v) = \sum_{i=1}^{p} \ell(m_i) \cdot \lg v_i. \quad (4)$$

We observe that all the costs of interest can be expressed as a sum of terms, each of them being a product of two factors: the first one involves the (bit-)size of digits, and the second one involves the size of the so-called continuants v_i, s_i.

1.1. Distributional analysis. We are interested here in studying the probabilistic behaviour of the gcd algorithm. The set Ω of inputs for the Euclid algorithm is $\Omega := \{(u,v) \in \mathbb{N}^2; \quad 0 \leq u < v\}$. For any (u, v) of Ω, the size of pair (u, v), denoted by $L(u, v)$, is just the binary length (or the size) of v, i.e., $L(u, v) := \ell(v)$. The subset Ω_n of inputs (u, v) with a fixed size n,

$$\Omega_n := \{(u,v) \in \Omega; \quad L(u,v) = n\}, \quad (5)$$

is endowed with the uniform probability $\mathbb{P}_n$. For a random variable R defined on Ω, its restriction to Ω_n is denoted by R_n, and we wish to analyze the asymptotic behaviour of R, i.e., the evolution of variables R_n when n becomes large.
The evolution of the mean values $\mathbb{E}[R_n]$ is of great interest and, more generally, the study of all moments $\mathbb{E}[R_n^\ell]$ provides a first understanding of the probabilistic behaviour of the algorithm: this is the average–case analysis. However, the distributional analysis, which describes the evolution of the distribution of variable R_n, provides a much more precise analysis of the algorithm: this is the ultimate purpose in analysis of algorithms. Very often, variables R_n have a distribution which tends to the gaussian law: this phenomenon is easily proved as soon as cost R_n is the sum of n elementary costs, which are independent, and possess the same distribution. However, in the "Euclidean world", the steps of (1) are not independent, and the distribution of numbers may evolve with the evolution of the algorithm. This is why asymptotic gaussian laws, even if they are widely expected, are often difficult to prove in this context.

We provide here such a distributional analysis, for the most precise parameter of the extended Euclid algorithm, its bit–complexity D. We are also interested in describing the evolution of the size of remainders v_i. There exist now many well-known results about the probabilistic behaviour of the Euclid algorithm, even if the last ones have been obtained recently. The first results on probabilistic

analysis of Euclid's algorithm are due to Heilbronn and Dixon who have shown, around 1970, that the average number of iterations is linear with respect to the size. In 1994, Hensley [6] performed the first distributional analysis, and proved that the number of steps has an asymptotic gaussian behaviour. However, his proof is not easily extended to other parameters of the algorithm. During the last ten years, the research team in Caen has designed a complete framework for analyzing an entire class of Euclidean algorithms, with a large class of parameters (see [10]). It is possible to obtain precise results on the average behaviour of the main parameters of the algorithm : the digits m_i, and the size of continuants v_i and s_i. Akhavi and Vallée have also analyzed the average bit–complexity [1]. These methods consider the underlying dynamical systems, and make a deep use of dynamical tools, like the transfer operator. However, all the analyses were "average–case analyses". There was a breakthrough two years ago, when Baladi and Vallée [2] extended the previous method for obtaining limit distributions, for a large class of costs, the so-called additive costs of moderate growth; they consider costs C defined on Ω and associated to an elementary cost c on digits,

$$C(u,v) := \sum_{i=1}^{p} c(m_i). \tag{6}$$

When $c(m)$ is $O(\log m)$, the cost c, and the cost C are said to be of moderate growth. This class of costs contains quite natural parameters, as the number of steps (for $c = 1$), the number of occurrences of a given digit m_0 (for $c(m) := [[m = m_0]]$), the total encoding length (when c equals the binary length ℓ), but NOT the bit–complexity. These bit–complexity costs are more difficult to deal with, because they involve both continuants and digits, in a multiplicative way. Here, we aim to study the distribution of the bit–complexity, and we wish to extend both the results of Akhavi and Vallée, about the average bit–complexity, and the distributional methods of Baladi and Vallée. We wish also to study the evolution of the size of remainders v_i.

As in previous works [2,3], we make a deep use of the weighted transfer operator relative to an elementary cost c and which depends on two parameters (s, w),

$$\mathbf{G}_{s,w,[c]}[f](x) := \sum_{m\geq 1} \frac{1}{(m+x)^{2s}} \cdot \exp[wc(m)] \cdot f\left(\frac{1}{m+x}\right). \tag{7}$$

When c is of moderate growth, the operator $\mathbf{G}_{s,w,[c]}$ admits (on a convenient functional space) a unique dominant eigenvalue, for (s, w) near $(1, 0)$. The logarithm of the dominant eigenvalue, called the pressure, and denoted by $\Lambda_{[c]}(s, w)$, plays a central work in [2], and also in the present paper. The particular case when c equals the binary length ℓ is crucial in study of bit–complexities.

1.2. Asymptotic gaussian laws. We prove here that many variables R defined on Ω follow asymptotically a gaussian law. We first provide a precise definition:

Definition [Asymptotic gaussian law.] *Consider a cost R defined on Ω and its restriction R_n to Ω_n. The cost R asymptotically follows a gaussian law if there exist three sequences a_n, b_n, r_n, with $r_n \to 0$, for which*

$$\mathbb{P}\left[(u,v) \in \Omega_n \;\middle|\; \frac{R_n(u,v) - a_n}{\sqrt{b_n}} \leq y\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2}\, dt + O(r_n)\,.$$

The sequence r_n defines the speed of convergence, denoted also by $r[R_n]$. The expectation $\mathbb{E}[R_n]$ and the variance $\mathbb{V}[R_n]$ satisfy $\mathbb{E}[R_n] \sim a_n, \quad \mathbb{V}[R_n] \sim b_n$. We say that the triple (a_n, b_n, r_n) is a characteristic triple for the gaussian law of R.

For instance, the result of Baladi and Vallée can be stated as follows.

Theorem 0. [Asymptotic gaussian Law for additive costs of moderate growth] (Baladi and Vallée). *Consider an additive cost C relative to an elementary cost c of moderate growth [defined in (6)].*
(i) On the set of integer inputs of size n, the cost C asymptotically follows a gaussian law, with a characteristic triple given by: $r[C_n] = O(n^{-1/2})$,

$$\mathbb{E}[C_n] = \mu(c) \cdot n + \mu_1(c) + O(2^{-n\gamma}), \qquad \mathbb{V}[C_n] = \rho(c) \cdot n + \rho_1(c) + O(2^{-n\gamma}),$$

Here γ is a strictly positive constant which does not depend on cost c.
(ii) The constants $\mu(c)$ and $\rho(c)$ involve the first five derivatives of order 1 and 2 of the pressure function $\Lambda(s, w) = \Lambda_{[c]}(s, w)$ of $\mathbf{G}_{s,w,[c]}$ at $(s, w) = (1, 0)$.

In the case when $c = \ell$, the constant $\rho(\ell)$ is (only) polynomial–time computable (see [8]) while $\mu(\ell)$ admits a closed form

$$\mu(\ell) = \frac{12}{\pi^2} \log \prod_{i=0}^{\infty} (1 + \frac{1}{2^i}).$$

1.3. Our main results. The "extended" cost D defined in (3) is easier to analyze because it is, in a sense, more regular than cost B. We prove here that the cost D follows asymptotically a gaussian law, with a characteristics triple which involves constants $\mu(\ell), \rho(\ell)$ of Thm 0 relative to the binary–length ℓ.

Theorem 1. [Asymptotic gaussian law for the extended bit–complexity.] *(i) On the set of integer inputs of size n, the bit complexity D of the extended Euclid algorithm follows asymptotically a gaussian law, with the characteristic triple*

$$\mathbb{E}[D_n] = \mu(\ell) \cdot n^2 \, [1 + O\left(\frac{1}{n}\right)], \;\; \mathbb{V}[D_n] = \rho(\ell) \cdot n^3 \, [1 + O\left(\frac{1}{n}\right)], \;\; r[D_n] = O(n^{-1/3}).$$

The smoothed bit–complexity $\widetilde{D}$ asymptotically follows a gaussian law with the same characteristic triple $[\mu(\ell) \cdot n^2, \, \rho(\ell) \cdot n^3, \, O(n^{-1/3})]$.
(ii) Under conjecture (C1), the speed of convergence $r[\widetilde{D}_n]$ is $O(n^{-1/2})$.

Conjecture $(C1)$ is described in 2.3. For the standard bit–cost B, defined in (2), we exhibit a precise estimate for the variance, and propose a conjecture $(C2)$, described in 2.4, under which we prove an asymptotic gaussian law.

Theorem 2. [Standard integer bit–complexity.] *(i) On the set of integer inputs of size n, the mean and the variance of the bit-complexity B satisfy*

$$\mathbb{E}[B_n] = \frac{1}{2} \mu(\ell) \cdot n^2 \, [1 + O\left(\frac{1}{n}\right)], \qquad \mathbb{V}[B_n] = \tau \cdot n^3 \, [1 + O\left(\frac{1}{n}\right)]. \tag{8}$$

Here τ is a strictly positive constant, which involves spectral objects of the operator $\mathbf{G}_{s,w,[c]}$. The same holds for the smoothed version $\widetilde{B}$.

(ii) Under Conjecture $(C2)$, the speed of convergence $r[\widetilde{B}_n]$ is $O(n^{-1/3})$, and the equality $4\tau = \rho(\ell)$ holds.

We are also interested in describing the evolution of the size of remainders v_i during the execution of the algorithm, and we consider the size of the remainder v_i at "a fraction of the depth". More precisely, for a real $\delta \in]0,1[$, we denote by $\ell^{[\delta]}$ the logarithm of v_i when i equals $\lfloor \delta P \rfloor$, [P is the number of iterations of the Euclid algorithm]. The following result shows that the remainders at a fraction of the depth asymptotically follow a log-normal law, and that the evolution of the sizes of continuants is very regular. This result constitutes a "discrete version" of the well-known result of [9] (sharpened by Vallée in [11]) who shows that the n-th continuant of a real $x \in \mathcal{I}$ asymptotically follows a gaussian law, when $\mathcal{I}$ is endowed with any density of class C^1.
This result also plays a central rôle in the analysis of the so–called Interrupted Euclidean algorithm which stops as soon as the remainder v_i is less than v_0^δ. An average–case analysis of the Interrupted algorithm is provided in [4], and the present results are a first [but crucial] step towards the distributional analysis of the algorithm. And the Interrupted algorithm is itself a basic procedure of the Lehmer Euclid algorithm [7], or the recursive Euclidean algorithms.

Theorem 3. [gaussian limit law for sizes of continuants at a fraction of the depth.] *Consider a rational δ of $]0,1[$. On the set of integer inputs of size n, the length $\ell^{[\delta]}$ follows asymptotically a gaussian law, with mean, variance and speed of convergence given by* $r[\ell_n^{[\delta]}] = O(n^{-1/2})$,

$$\mathbb{E}[\ell_n^{[\delta]}] = \mu_{[\delta]} \cdot n + \mu_1(\delta) + O(2^{-n\gamma}), \qquad \mathbb{V}[\ell_n^{[\delta]}] = \rho_{[\delta]} \cdot n + \rho_1(\delta) + O(2^{-n\gamma}).$$

Here γ is a strictly positive constant which depends on δ, and the constants $\mu_{[\delta]}$ and $\rho_{[\delta]}$ are related to the derivatives of the pressure function $\Lambda(s)$ at $s = 1$,

$$\mu_{[\delta]} = (1-\delta), \qquad \rho_{[\delta]} = \delta(1-\delta)\frac{|\Lambda''(1)|}{|\Lambda'(1)|} > 0\,.$$

1.4. Plan of the paper. Section 2 provides a description of the main steps for proving Theorems 1 and 2 and states Theorem 4, which will be a main tool in these proofs. Section 3 presents the transfer operators and explains their generating rôle. Then, it describes the main principles of the analytical study which provides a proof of Theorems 3 and 4. Finally, we describe the two conjectures and provide some hints towards a possible proof.

2 Main steps for Theorems 1 and 2.

Here, we explain how to obtain asymptotic gaussian laws for the bit–complexities. We prove Theorem 1, Assertion (i), describe conjectures $(C1)$ and $(C2)$ and explain how to prove Theorem 1 (ii) and Theorem 2 (ii) under these conjectures.

2.1. Expressions for continuants. Each division–step of the Euclid algorithm $v = m \cdot u + r$ uses a digit m and changes the old pair (u, v) into a new pair (r, u). Instead of integers, we consider rationals [the old rational $x = u/v$, and the new rational $y = r/u$] which both belong to the unit interval, and we look for a relation between y and x. One has

$$\frac{r}{u} = \frac{v - mu}{u} = \frac{v}{u} - \lfloor \frac{v}{u} \rfloor \qquad \text{so that} \quad y = T(x) \qquad \text{with} \quad T(x) := \frac{1}{x} - \lfloor \frac{1}{x} \rfloor.$$

With $T(0) = 0$, the map $T : [0,1] \to [0,1]$ is called the Gauss map and plays a fundamental rôle in the study of the Euclid algorithm. When the quotient is m, there exists also a linear fractional transformation (LFT) $h_{[m]}$ for which

$$x = h_{[m]}(y) \qquad \text{with} \quad h_{[m]}(y) = 1/(m+y) \ .$$

Of course, the LFT's $h_{[m]}$ are the inverse branches of T. On an input (u, v), the execution (1) creates a continued fraction of the form

$$\frac{u}{v} = h_{[m_1]} \circ h_{[m_2]} \circ \dots h_{[m_p]} = h(0). \tag{9}$$

When the algorithm performs p iterations, it gives rise to a continued fraction of depth p. Here, we show that the main parameters of the Euclid algorithm on the input (u, v) (quotients m_i, remainders v_i and continuants s_i) can be read on the continued fraction of the rational u/v. When the CFE of u/v is split at depth i, the LFT h defines three LFT's, the beginning LFT b_i, the middle LFT h_i and the ending LFT e_i, respectively defined as

$$b_i := h_{[m_1]} \circ h_{[m_2]} \circ \dots \circ h_{[m_{i-1}]}, \qquad h_i := h_{[m_i]} \qquad e_i := h_{[m_{i+1}]} \circ \dots \circ h_{[m_p]}.$$

Then, the i-th continuants admit expressions which involve LFT's e_i and b_i,

$$s_i^{-2} = |b_i'(0)|, \qquad v_i^{-2} = v_p^{-2} \cdot |e_i'(0)|. \tag{10}$$

2.2. Bit-complexity cost. This entails the following decompositions,

Proposition 1. *The bit–complexity costs $D, \widetilde{D}$ of the extended Euclidean algorithm decompose as* $\widetilde{D} = (L-1) \cdot Z + \widetilde{Y}$, $D = (L+1) \cdot Z + Y$,

$$\text{with} \quad \widetilde{Y} = -Y^{(1)} + O(Y^{(2)}) + Y^{(3)} + Y^{(4)}, \qquad Y = \widetilde{Y} + Y^{(5)}.$$

Here L is the size of the input, defined by $L(u,v) = \ell(v) = \ell(v_0)$ and

$$Z = \sum_{i=1}^{p-1} \ell(m_i), \qquad Y^{(1)} = \sum_{i=1}^{p-1} \ell(m_i) \cdot \lg m_i, \qquad Y^{(2)} = (\ell(m_p) + \lg v_p)^2, \tag{11}$$

$$Y^{(3)} = f \cdot \sum_{i=1}^{p-1} \ell(m_i), \qquad Y^{(4)} = \sum_{i=1}^{p-1} d_i \cdot \ell(m_i) \qquad Y^{(5)} = \sum_{i=1}^{p-1} f_i \cdot \ell(m_i), \tag{12}$$

$$\text{with} \quad d_i =: \lg \left| \frac{h_i'(0)}{h_i'(e_i(0))} \right| + \lg \left| \frac{b_i'(e_{i-1}(0))}{b_i'(0)} \right|, \ f := \{\lg v_0\}, \ f_i := -\{\lg v_i\} - \{\lg s_i\}.$$

Moreover, the so-called distortions d_i admit uniform lower and upper bounds.

We have then "splitted" the extended cost D into two costs: the "main" cost $X := L \cdot Z$ which will be proven to be (asymptotically) gaussian, and a "remainder" cost Y, which will be proven to be (asymptotically) more concentrated that the main cost. Then, the total cost $X + Y$ will be (asymptotically) gaussian: .

Proposition 2. *Consider two costs X and Y, defined on Ω and their restrictions X_n, Y_n to Ω_n. Suppose that X admits a gaussian limit law with speed of convergence $r[X_n]$ and the variances of X_n and Y_n satisfy $\mathbb{V}[Y_n] = \alpha_n \cdot \mathbb{V}[X_n]$, with*

$\alpha_n \to 0$. *Then, the random variable* $X+Y$ *follows asymptotically a gaussian limit law with a characteristic triple given by:* $r[X_n+Y_n] = r[X_n] + O(\alpha_n^{1/3})$,

$$\mathbb{E}[X_n+Y_n] = \mathbb{E}_n[X] + \mathbb{E}_n[Y], \quad \mathbb{V}[X_n+Y_n] = \mathbb{V}[X_n]\cdot[1+O(\alpha_n)].$$

2.3. Proof of Theorem 1. Theorem 1 (i) is easily deduced from Propositions 1 and 2. The "main" cost is $X := L\cdot Z$, where $L(u,v)$ is the size of pair (u,v), equal to $\ell(v)$. With results of Baladi and Vallée [Theorem 0], the cost Z follows an asymptotic gaussian law. Since $X_n = n\cdot Z_n$, the cost X follows itself an asymptotic gaussian law with the characteristic triple

$$\mathbb{E}[X_n] = n\cdot\mathbb{E}[Z_n] = O(n^2), \quad \mathbb{V}[X_n] = n^2\cdot\mathbb{V}[Z_n] = O(n^3), \quad r[X_n] = O(n^{-1/2}).$$

In Proposition 1, there appear three different kinds of costs: – (i) cost $Y^{(1)}$ – (ii) cost $Y^{(2)}$ which is an end-cost, [i.e., it depends only on variables used in the last step $\ell(m_p), \ell(v_p)$, and in a polynomial way.] – (iii) The other costs R [the distortion cost $Y^{(4)}$ and the two fractional costs $Y^{(3)}$, $Y^{(5)}$] deal with bounded sequence f_i, d_i. For these costs R, one has:

$$\mathbb{E}[R_n] = O(\mathbb{E}[Z_n]) = O(n), \qquad \mathbb{V}[R_n] \le \mathbb{E}[R_n^2] = O(\mathbb{E}[Z_n^2]) = O(n^2).$$

In the following Theorem 4, we will prove that the cost $Y = Y^{(1)}$ fulfills the concentration property, and that end-costs R are negligible i.e.,

$$\mathbb{E}[Y_n] = O(n), \quad \mathbb{V}[Y_n] = O(n), \qquad \mathbb{E}[R_n] = O(1), \quad \mathbb{V}[R_n] = O(1).$$

This leads to Theorem 1 [Assertion (i)], with a speed of convergence $O(n^{-1/3})$. If we wish to obtain a speed of convergence of order $n^{-1/2}$, we must study more carefully costs $Y^{(i)}$ for $j = 3, 4, 5$. The fractional cost $Y^{(5)}$ is clearly very difficult to study: this is why we have introduced the smoothed cost $\widetilde{D}$, which no longer involves $Y^{(5)}$. It is possible to generate the distortion cost $Y^{(4)}$ and the fractional cost $Y^{(3)}$ with some convenient transfer operator. However, we do not succeed in proving that the concentration property holds for them.

Conjecture $(C1)$: *The costs* $Y^{(3)}$ *and* $Y^{(4)}$ *satisfy the concentration property.* Under this conjecture, Theorem 1 (ii) is proven. ■

2.4. An asymptotic gaussian law for $\widetilde{B}$? For proving Theorem 2 (ii), we relate $w_i := v_i/v_p = (e_i'(0))^{1/2}$ to the approximate continuant $\overline{s}_i := b_i'(e_{i-1}(0))^{-1/2}$ and we introduce two (new) costs

$$A(u,v) := \sum_{i=1}^{p} \ell(m_i)\cdot \lg w_i, \qquad \overline{A}(u,v) := \sum_{i=1}^{p} \ell(m_i)\cdot \lg \overline{s}_i.$$

First, as in 2.3, the cost $(A+\overline{A})$ will be asymptotically gaussian with the same characteristic triple as $\widetilde{D}$. Second, since the cost A is close to costs $\widetilde{B}, B$, it is sufficient for Theorem 2 (ii) to prove that the decomposition $A = (1/2)(A+\overline{A}) - (1/2)(A-\overline{A})$ provides a new instance of application of Propositions 1 and 2. This is possible if the second cost $(A_n - \overline{A}_n)$ has a variance of order $O(n^2)$. Since $\mathbb{E}[A_n - \overline{A}_n]$ is of order $O(n)$ [see Proposition 3, Section 3], we study

$$\mathbb{E}[(A_n - \overline{A}_n)^2] = \mathbb{E}[A_n^2] + \mathbb{E}[\overline{A}_n^2] - 2\mathbb{E}[A_n\cdot\overline{A}_n],$$

where each term is of order $O(n^4)$. Proposition 3 proves a cancellation between the dominant terms, and entails for α_n an order of $O(1/n)$.

Conjecture $(C2)$: $\mathbb{E}[A_n^2]$, $\mathbb{E}[\overline{A}_n^2]$, $\mathbb{E}[A_n \cdot \overline{A}_n]$ *have the same terms of order* n^3.

This conjecture is plausible since it is based on a property of "semi-commutativity" which generalizes Proposition 3. Under $(C2)$, it is easy to prove Theorem 2 (ii).

2.5. Various kinds of costs. We are then led to study various costs C, and the behaviour of additive costs C heavily depends on the behaviour of cost c. We then introduce the Dirichlet series $A_c(s, w)$,

$$A_c(s, w) := \sum_{m \in \mathcal{M}} \frac{1}{m^{2s}} \exp[wc(m)],$$

closely related to the operator $\mathbf{G}_{s,w,[c]}$, which helps to define the behaviour of c.

Definition 1. (a) *A cost R is an end–cost if it depends only on variables used in the last step $\ell(m_p), \ell(v_p)$, and in a polynomial way.*
(b) *An elementary cost c and its associated additive cost C are of moderate growth if –$(b1)$ the bivariate generating function $A_c(s, w)$ is convergent for $\Re s > \sigma_0$ and $\Re w < \nu_0$ with $\sigma_0 < 1$ and $\nu_0 > 0$ – $(b2)$ it is analytic at $(1, 0)$,*
(c) *An elementary cost c and its associated additive cost C are of intermediate growth if –$(c1)$ its generating function $A_c(s, w)$ is convergent for $\Re s > \sigma_0$ with $\sigma_0 < 1$ and $\Re w \leq 0$, – $(c2)$ it is not analytic at $(s, w) = (1, 0)$, but, as a function of the real variable w, it admits derivatives of any order wrt w, at $w = 0^-$.*

Remark. The size cost $c = \ell$ is of moderate growth, while any power of the size of the form $c = \ell^\alpha$ (with $\alpha > 1$) defines a cost of intermediate growth.

The following theorem is one of the basic results of our paper. Note that Assertion (b) is already proven by Baladi and Vallée [2].

Theorem 4. *The following holds:*

(a) *An end cost R is negligible, i.e., the expectation $\mathbb{E}[R_n]$ and the variance $\mathbb{V}[R_n]$ are $O(1)$.*

(b) *An additive cost C of moderate growth is asymptotically gaussian with a characteristic triple of the form $[O(n), O(n), O(n^{-1/2})]$.*

(c) *An additive cost C of intermediate growth satisfies the concentration property, i.e., the expectation $\mathbb{E}[C_n]$ and the variance $\mathbb{V}[C_n]$ are $O(n)$.*

3 Dynamical Systems and Generating operators.

We explain here how dynamical systems allow to derive alternative forms for generating functions. This will be done via various extensions of the transfer operator, which plays here the rôle of a "generating operator".

3.1. Dynamical systems and transfer operators. A continuous extension of one step of the Euclid algorithm to real numbers x of $\mathcal{I} := [0, 1]$ is provided by the Gauss map $T : \mathcal{I} \to \mathcal{I}$, together with the set $\mathcal{H} := \{h_{[m]}; m \in \mathbb{N}\}$ of the branches of T^{-1}. The pair $(\mathcal{I}, T)$ defines a dynamical system. The set $\mathcal{H}^k$ is the

set of the inverse branches of the iterate T^k, and the set $\mathcal{H}^\star := \cup_k \mathcal{H}^k$ is the semi-group generated by $\mathcal{H}$.

If $\mathcal{I}$ is endowed with some initial density $f = f_0$, the time evolution governed by the map T modifies the density. The successive densities $f_1, f_2, \ldots, f_n, \ldots$ describe the global evolution of the system, and there exists an operator, the density transformer $\mathbf{G}$ which transforms f_0 into f_1. The weighted transfer operator $\mathbf{G}_{s,w,[c]}$ relative to some digit cost c,

$$\mathbf{G}_{s,w,[c]}[f](x) = \sum_{h\in\mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x),$$

is a perturbation of the density transformer $\mathbf{G}$ [obtained for $(s, w) = (1, 0)$]. When $w = 0$, we omit the variable w and the cost c, so that $\mathbf{G}_s := \mathbf{G}_{s,0,[c]}$. Now, if we extend cost c on $\mathcal{H}^\star$ by additivity, the quasi–inverse is of the form

$$(I - \mathbf{G}_{s,w,[c]})^{-1}[f](x) = \sum_{h\in\mathcal{H}^\star} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x) \ .$$

3.2. Generating functions. We consider a general parameter R defined on Ω, and we wish to study its distribution on Ω_n, when endowed with the uniform probability. Our final probabilistic tool [for distributional analyses] is the sequence of moment generating functions $\mathbb{E}[\exp(wR_n)]$,

$$\mathbb{E}[\exp(wR_n)] = \frac{R(n,w)}{R(n,0)}, \qquad \text{with} \quad R(n,w) := \sum_{(u,v)\in\Omega_n} \exp[wR(u,v)]. \tag{13}$$

We first consider the whole set Ω of inputs and our strategy consists in encapsulating all the moment generating functions $\mathbb{E}[\exp(wR_n)]$ in a Dirichlet series

$$S_R(s,w) := \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}} \exp[wR(u,v)] = \sum_{m\geq 1} \frac{1}{m^{2s}} r_m(w)\,, \tag{14}$$

where $r_m(w)$ is the cumulative value of $\exp[wR]$ on inputs (u, v) for which $v = m$. The series $S_R(s, w)$ is a bivariate generating function which depends on two parameters, s "marks" the input size, and w "marks" the cost of interest. This is a Dirichlet series with respect to s. The study of moments of order k,

$$\mathbb{E}[R_n^k] = \frac{R(n)^{[k]}}{R(n)^{[0]}}, \qquad \text{with} \quad R(n)^{[k]} := \sum_{(u,v)\in\Omega_n} R^k(u,v), \tag{15}$$

deals with a Dirichlet Series $S_R^{[k]}(s)$

$$S_R^{[k]}(s) := \frac{\partial^k}{\partial w^k} S_R(s,w)|_{w=0} = \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}} R^k(u,v) = \sum_{m\geq 1} \frac{1}{m^{2s}} r_m^{[k]}, \tag{16}$$

where $r_m^{[k]}$ is the cumulative value of R^k on inputs (u, v) for which $v = m$.

In both cases, the plain moment generating function, , and the plain moment of order k of R_n can be recovered from series $S_R(s, w)$ or $S_R^{[k]}(s)$ with (13,15), and relations

$$R(n,w) = \sum_{m=2^{n-1}}^{2^n-1} r_m(w), \qquad R(n)^{[k]} = \sum_{m=2^{n-1}}^{2^n-1} r_m^{[k]}. \tag{17}$$

We first look for an alternative expression for series $S_R(s,w)$ [defined in (14)] from which the position and the nature of its dominant singularity become apparent. With taking derivatives, we also obtain alternative expressions for $S_R^{[k]}(s)$ [defined in (16)]. Then, we transfer these informations on the asymptotic behaviour of coefficients of $S_R(s,w)$ or $S_R^{[k]}(s)$, which are closely related via (13,15,17) to our prime objects of interest $\mathbb{E}[\exp(wR_n)]$, $\mathbb{E}[R_n^k]$.

3.3. Alternative expressions for bivariate Dirichlet series. We will use transfer operators $\mathbf{G}_s, \mathbf{G}_{s,w,[c]}$ (or some of their extensions) as "generating" operators: Bivariate generating functions $S_R(s,w)$ can be expressed via quasi–inverses of these operators.

Additive costs. If C is the total cost relative to c, the quasi-inverse $(I-\mathbf{G}_{s,w,[c]})^{-1}$ "generates" the bivariate generating function of cost C (relative to coprime inputs). Furthermore, the Zeta function defined as $\zeta(2s) := \sum_{d\geq 1} d^{-2s}$ allows to deal with general inputs [not only coprime inputs]. Finally,

$$S_C(s,w) = \zeta(2s)\cdot (I-\mathbf{G}_{s,w,[c]})^{-1}[1](0). \tag{18}$$

Continuant at a fraction of the depth. We study the parameter $\ell^{[\delta]}$ which equals the logarithm of remainder v_i for $i=\lfloor \delta P\rfloor$. For an input (u,v) of Ω on which the algorithm performs p iterations, there exists LFT h of depth p such that $u/v = h(0)$. One decomposes h in two LFT's g and r of depth $\lfloor \delta p\rfloor$ and $p-\lfloor \delta p\rfloor$ such that $h=g\circ r$, and if δ is a rational of the form $\delta = c/(c+d)$, then

$$S_{\ell^{[\delta]}}(s,w) = \zeta(2s-2w)\cdot \sum_{j=0}^{c+d-1} \mathbf{G}_{s-w}^{j-\lfloor \delta j\rfloor} \circ \Big(\sum_{k\geq 0} \mathbf{G}_{s-w}^{dk}\circ \mathbf{G}_s^{ck}\Big)\circ \mathbf{G}_s^{\lfloor \delta j\rfloor}[1](0). \tag{19}$$

The operator $\mathbb{G}_{s,w} := \sum_{k\geq 0} \mathbf{G}_{s-w}^{dk}\circ \mathbf{G}_s^{ck}$ is called a pseudo–quasi-inverse; of course, since $\mathbf{G}_s$ and $\mathbf{G}_{s,w}$ do not commute, this is not a "true" quasi-inverse. However, we study this operator when w is near to 0, and we can hope that the properties of $\mathbb{G}_{s,w}$ will be close to properties of a true quasi-inverse.

3.4. Alternative expressions for Dirichlet series $S_R^{[j]}(s)$. In other cases of cost R, we look for alternative expressions for the series $S_R^{[i]}(s)$ for $i=1,2$. We denote by $W_{[c]}$ the derivation wrt w (at $w=0$), and by Δ the derivation wrt s,

$$W_{[c]}\mathbf{G}_s = \frac{\partial}{\partial w}\mathbf{G}_{s,w,[c]}|_{w=0}, \qquad \Delta := \frac{1}{\log 2}\frac{d}{ds}\mathbf{G}_s.$$

Then, the operators $W_{[c]}$ or Δ operate themselves on transfer operators. Our Dirichlet series of interest can be written as a sequence of occurrences of the quasi–inverse $(I-\mathbf{G}_s)^{-1}$, separated by occurrences of the form $A\mathbf{G}_s$ where A is a monomial of the (commutative) algebra $\mathcal{A}$ generated by $\{\Delta, W_{[c]}\}$. Then, we adopt shorthand notations where we omit the quasi-inverses, the zeta function, the function 1, and the point 0: we only take into account the operators "between" the quasi inverses.

Additive costs C of intermediate growth. In this case, it is not possible to deal directly with the transfer operator $\mathbf{G}_{s,w,[c]}$. However, the univariate series $S_C^{[j]}(s)$ admit alternative expressions of the form

$$S_C^{[1]} = [W_{[c]}] \qquad S_C^{[2]} = [W_{[c]}^2] + 2[W_{[c]}, W_{[c]}]. \tag{20}$$

Bit-Complexity Costs $A, \overline{A}$. Here, we omit also the index $[\ell]$ in $W_{[\ell]}$ and we obtain, in the same vein $\quad S_A^{[1]} = [\Delta, W], \qquad S_{\overline{A}}^{[1]} = [W, \Delta],$

$$(1/2)\, S_A^{[2]} \approx 2[\Delta,\Delta,W,W]+[\Delta,W,\Delta,W]+[\Delta^2,W,W]+[\Delta,\Delta W,W]+[\Delta,\Delta,W^2],$$

$$(1/2)\, S_{\overline{A}}^{[2]} \approx 2[W,W,\Delta,\Delta]+[W,\Delta,W,\Delta]+[W,W,\Delta^2]+[W,\Delta W,\Delta]+[W^2,\Delta,\Delta],$$

$$S_{A\overline{A}}^{[1]} \approx 2[W,\Delta,\Delta,W]+2[\Delta,W,W,\Delta]+[W,\Delta,W,\Delta]+[\Delta,W,\Delta,W]+$$

$$+[W,\Delta^2,W]+[\Delta,W,\Delta W]+[W,\Delta,\Delta W]+[\Delta W,\Delta,W]+[\Delta W,W,\Delta]+[\Delta,W^2,\Delta].$$

3.5. Analysis of Costs. With alternative expressions of Dirichlet series provided in Sections 3.3 and 3.4 at hand, we now perform the second step: we find the dominant singularities of these Dirichlet series and their nature, and then transfer these informations towards coefficients and obtain asymptotic expressions for their coefficients. We use, as a main tool, convenient "extractors" which express coefficients of series as a function of the series itself. There are two main "extractors" for Dirichlet series: Tauberian Theorems [which do not provide remainder terms] are well-adapted for the average–case analysis or the study of all the (non centered) moments [Thm 4 (a)] — the Perron Formula [which may provide remainder terms] constitutes an essential step, both in the studies of the variance [Thms 1, 2, 4(c)] and in distributional analyses [Thm 3].

Both extractors need informations on the quasi-inverse (QI), closely related to the dominant spectral properties of the transfer operator on the Banach space $\mathcal{C}^1(\mathcal{I})$. However, Tauberian Theorems "only" need informations on the QI on the domain $\Re s \geq 1$. For using with some success the Perron Formula, we need a more precise knowledge of the QI on vertical strips on the left of the vertical line $\Re s = 1$. Properties of the same vein are very often difficult to prove and intervene for instance in the proof of the Prime Number Theorem. The US Property [Uniformity on Strips] describes a convenient behaviour of the QI and informally says: "there exists a vertical strip $|\Re(s) - 1| < \alpha$ which contains only one pole of the QI; moreover, on the left line $\Re(s) = 1 - \alpha$, an adequate norm of the QI is bounded by $M \cdot |\Im s|^\xi$ (with $\xi > 0$ small). Baladi and Vallée [2] have generalized ideas due to Dolgopyat [5] and prove that the $US(s)$ Property holds for $(I - \mathbf{G}_s)^{-1}$, and that a uniform $US(s, w)$ Property (uniform wrt w) holds for $(I - \mathbf{G}_{s,w,[c]})^{-1}$ (when c is of moderate growth). Here, for Thm 3, we prove that a uniform $US(s, w)$ Property also holds for the "pseudo quasi–inverse".

For $(s, w) = (1, 0)$, and for any cost c, the operator $\mathbf{G}_{s,w,[c]}$ is just the density transformer $\mathbf{G}$, which possesses a unique dominant eigenvalue equal to 1 and an invariant function $\Psi(x) = (1/\log 2)(1/1 + x)$. Then, each occurrence of the quasi–inverse $(I - \mathbf{G}_s)^{-1}$ brings a pole at $s = 1$, with an explicit residue:

Proposition 3. *Any Dirichlet series denoted by an expression* $[A_1, A_2, \dots A_k]$ *[see Section 3.4], where each* A_i *is a monomial of the algebra generated by* $\{\Delta, W_{[c]}\}$, *has a pôle of order* $k + 1$ *at* $s = 1$, *with an expansion of the form*

$$[A_1, A_2, \dots, A_k](s) = \frac{1}{\log 2} \sum_{i \geq 0} a_i \cdot (|\lambda'(1)|(s-1))^{i-k-1},$$

$$\textit{with} \quad a_0 = \prod_{i=1}^{k} I[A_i \mathbf{G}],\ \ I[\mathbf{H}] := \int_I \mathbf{H}[\Psi](t)dt,\ \ \Psi(x) := (1/\log 2)(1/1+x).$$

Since the dominant constant a_0 depends only on the subset $\{A_1, A_2, \dots A_k\}$, this proves that, for additive costs $R = C$ or bit-complexities $R = A, \overline{A}$, there exists a relation of the form $b_0 = 2a_0^2$ between the dominant constant a_0 of $S_R^{[1]}$ and the dominant constant b_0 of $S_R^{[2]}$, which entails a cancellation in the variance.

Conjecture $(C2)$**.** It is based on a similar property which involves the Porter operator $\mathbf{Q}$ defined as the constant term in the expansion of $(I - \mathbf{G}_s)^{-1}$ near $s = 1$. Conjecture $(C2)$ says: *The following equality holds:*

$$\sum_{\substack{X,Y,X',Y' \in \{\Delta, W\} \\ X' \neq X, Y' \neq Y}} (-1)^{[X=Y]} \cdot I[X\mathbf{G}] \cdot I[Y\mathbf{G}] \cdot (I[X'\mathbf{G} \circ \mathbf{Q} \circ Y'\mathbf{G}] - I[X'Y'\mathbf{G}]) = 0.$$

Conjecture $(C1)$**.** It deals with two families of costs.

Costs $R = f \cdot C$. To prove that $\mathbb{V}R_n$ is $O(n)$, we use generating functions relative to moments of $\overline{R} := \lg v_0 \cdot C$. They can be expressed with $[\cdot, \dots \cdot, \cdot]$, and we have to prove cancellations between the constants, as in Conjecture $(C2)$.

Distortion costs. Generating functions for the distortion costs involve generalized transfer operators acting on functions with two variables as in [11]. The US properties are not yet proven to hold for such operators, and proving cancellations between constants needs to deal with their dominant spectral objects.

References

1. A. Akhavi, B. Vallée. Average bit–complexity of Euclidean Algorithms, Proceedings of ICALP'2000, Lecture Notes in Computer Science 1853, pp 373–387, Springer.
2. V. Baladi, B. Vallée. Euclidean Algorithms are Gaussian, Journal of Number Theory, Volume 110, Issue 2 (2005) pp 331–386
3. E. Cesaratto, B. Vallée. Reals with bounded digit averages, Proceedings of the Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probability, pp 473–490, M. Drmota et al., ed., Birkhauser Verlag, Trends in Mathematics, 2004.
4. B. Daireaux, B. Vallée. Dynamical analysis of the parameterized Lehmer-Euclid Algorithm, Combinatorics, Probability, Computing, pp 499–536 (2004).
5. D. Dolgopyat. On decay of correlations in Anosov flows, *Ann. of Math.* 147 (1998) 357–390.
6. D. Hensley. The number of steps in the Euclidean algorithm, Journal of Number Theory, 49, 2 (1994), 142-182
7. D.H. Lehmer. Euclid's algorithm for large numbers. Am. Math. Mon. (1938) 45 pp 227–233.
8. L. Lhote. Computation of a Class of Continued Fraction Constants Proceedings of Alenex–ANALCO04, pp 199–210
9. W. Philipp. Some metrical theorems in number theory II, Duke Math. J. 37 (1970) pp 447–488. Errata, ibid, 788.
10. B. Vallée. Euclidean Dynamics, to appear in Discrete and Continuous Dynamical Systems, 2005, web page: `www.info.unicaen.fr/~brigitte`
11. B. Vallée. Opérateurs de Ruelle-Mayer généralisés et analyse en moyenne des algorithmes de Gauss et d'Euclide, Acta Arithmetica 81.2 (1997), pp 101–144