EUCLIDEAN ALGORITHMS ARE GAUSSIAN

VIVIANE BALADI, BRIGITTE VALLÉE

Abstract. This study provides new results about the probabilistic behaviour of a class of Euclidean algorithms. It proves that the asymptotic distribution of a whole class of parameters associated to these algorithms is normal. Furthermore, it precisely compares the expected properties of continued fraction expansions of real numbers and rational numbers. This paper also focuses on the basic parameter of the algorithms, the number of steps. Hensley [23] already proved that there exists a Local Limit Theorem for this parameter, in the case of the standard Euclidean algorithm. This paper provides a new proof of this fact, both more natural and more concise, and, at the same time, an extension to two other Euclidean algorithms. The whole paper is based on the so–called dynamical analysis methodology, that views an algorithm as a dynamical system and introduces tools that come from dynamical systems into analysis of algorithms. The tranfer operator of the underlying dynamical system then plays a fundamental rôle, together with various other tools: Dirichlet series, Perron's formula, the Quasi-Powers theorem, saddle point method. Furthermore, this work introduces, for the first time, a new facet in dynamical analysis, distributional dynamical analysis, and it needs precise estimates on the transfer operators when the parameter varies along vertical complex lines. Such estimates are provided by adapting to the Euclidean context powerful methods due to Dolgopyat [15].

1. Introduction

According to Knuth [29, p. 335], "we might call Euclid's method the granddaddy of all algorithms, because it is the oldest nontrivial algorithm that has survived to the present day". Since addition and multiplication of rational numbers both require gcd calculations, the Euclid algorithm is a basic building block of computer algebra systems and multiprecision arithmetic libraries, and, in many such applications, most of the time is indeed spent in computing gcd's. However, although its design is quite clear, the Euclidean algorithm has not yet been completely analysed.

Continued fraction expansions. Continued fraction expansions of real numbers are the result of an iterative process that constitutes the continuous conterpart of the Euclidean algorithm. They can be viewed as trajectories of a specific dynamical system relative to the map $T : [0, 1] \to [0, 1]$ defined by

$$T(x) := \frac{1}{x} - \lfloor \frac{1}{x} \rfloor, \quad \text{for } x \neq 0, \quad T(0) = 0.$$

Date: July, 28th 2003.

Every real number x then admits a continued fraction expansion of the form

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{m_3 + \cfrac{1}{\ddots + \cfrac{1}{m_p + \dots}}}}}, \tag{1.1}$$

and the corresponding trajectory $\mathcal{T}(x) = (x, Tx, T^2x, \dots, T^nx, \dots)$ is entirely described by the (infinite) sequence of digits $(m_1, m_2, m_3, \dots, m_n, \dots)$. It is usual to consider the truncated trajectory $\mathcal{T}_n(x) := (x, Tx, \dots, T^nx)$ together with its truncated encoding $(m_1, m_2, \dots, m_n)$, which keeps track of the first n stages of the process, and then let the truncation degree n tend to ∞.

Associate a positive real value $c(m)$ to digit m and extend the function c additively to (truncated) trajectories, defining

$$C^{(n)}(x) := \sum_{i=1}^{n} c(m_i).$$

Some special cases are particularly interesting: if $c = c_m$ is the characteristic fonction of some fixed digit m, then $C^{(n)}$ equals the number of occurrences of digit m amongst the first n digits of the *CF*–expansion; if c is the binary length of the digit, then $C^{(n)}$ is closely related to the length of a standard encoding of the trajectory $\mathcal{T}_n$. We shall globally refer to c as a "digit–cost" and to C as the associated "total cost", since costs (in the usual sense of computational complexity) are usually expressible in this way.

Generally speaking, costs $C^{(n)}$ play the rôle of "observables" in physics, and it is of great importance to characterize their probabilistic behaviour. The first major question is: When does the sequence $(1/n)C^{(n)}(x)$ admit a pointwise limit as the truncation degree tends to ∞ ? A general answer to this question has been first provided by ergodic theorems. An important tool, introduced shortly afterwards, namely, the density transformer (also known as Perron–Frobenius operator), encapsulates all the important informations relative to the "dynamics" of the iterative process. Within the realm of continued fractions, studies of Lévy [33], Khinchin [28], Kuzmin [30], Wirsing [51], Babenko [2], and Mayer [34] describe the main properties of the density transformer. Two properties of the density transformer, namely, the existence of a unique dominant eigenvalue [Property *UDE*] together with the existence of a spectral gap that separates the (unique) dominant eigenvalue from the remainder of the spectrum [Property *SG*] entail the applicability of ergodic methods, which are both simple and powerful, while providing asymptotic estimates that hold almost eveywhere. It then becomes possible to quantify the probable behaviour of truncated trajectories as the truncation degree n tends to ∞: For each cost c of moderate growth, there exists a constant $\widehat{\mu}(c)$ such that, for almost all x, the total cost $C^{(n)}(x)$ is asymptotic to $n\widehat{\mu}(c)$.

Moreover, in order to study the speed of convergence and establish more precise results on the distribution of truncated trajectories, it proves very useful to deal with an extension of the density transformer, namely the transfer operator, that was originally introduced by Ruelle [40] in connection with the thermodynamic formalism and dynamical systems theory [4]. Such an operator can be built, relative to some digit–cost c; it depends on two parameters (s, w) and is denoted by $\mathbf{H}_{s,w}$. It coincides with the density transformer for $(s, w) = (1, 0)$. As soon as a supplementary property is fulfilled, namely Property $An_w(1, 0)$ expressing that the transfer operator $\mathbf{H}_{1,w}$ is analytic with respect to w when w is near 0, elementary perturbation theory [27] entails that the two properties UDE and SG extend to the operator $\mathbf{H}_{1,w}$ for w near 0. In the case when parameters are af moderate growth, this leads to a proof that the asymptotic distribution of total costs is Gaussian, i.e.,

$$\Pr\left[x \;\Big|\; \frac{C^{(n)}(x) - \widehat{\mu}(c)n}{\widehat{\delta}(c)\sqrt{n}} \le y\right] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt,$$

with speed of convergence $O(1/\sqrt{n})$. The constants $\widehat{\mu}(c)$ and $\widehat{\delta}^2(c)$ relative to the expectation and variance of C are expressed in terms of the first and second derivatives of the "dominant eigenvalue" function, $w \mapsto \lambda(1, w)$, relative to the transfer operator $\mathbf{H}_{1,w}$. This result is quite well known, (for instance, see [11] or [9] for a proof) but this paper provides an alternative proof mainly based on the use of moment generating functions (Theorem 1).

On the other hand, in relation to Euclid's algorithm, rational inputs give rise to trajectories that must be stopped as soon as they attain 0. Accordingly, continuous results are not well-suited to providing information on the behaviour of the main parameters that intervene in the continued fraction expansion of a rational number. However, this particular case of a rational input is central in computer science since it is closely related to the average–case analysis of the Euclidean Algorithm. As it is often the case, discrete problems are more difficult to solve than their continuous counterparts.

Average–case analysis of Euclidean Algorithms. There are several variants of the standard continued fraction algorithm summarized by (1.1). As a matter of fact, we study three of them specified at the beginning of the next section. The discrete counterparts of these continued fraction algorithms, i.e., the Euclidean algorithms, have been less extensively studied. An execution of a Euclidean algorithm on the input (u, v) gives rise to a rational trajectory $\mathcal{T}(u/v)$ which, as already said, is stopped as soon as it attains 0. The analysis of Euclidean algorithms then depends on an analysis of the probabilistic behaviour of such rational trajectories. Now, the reference parameter is no longer the truncation degree n, but it is related to the size N of the integer input (u, v) defined to be $N := \max(u, v)$.

The average–case complexity of Euclidean Algorithms is by now well–understood. The basic parameter is the number of steps P, which corresponds to the trivial digit–cost $c = 1$ and plainly measures the number of steps of the trajectory $\mathcal{T}(x)$. The standard Euclidean Algorithm has been first analysed in the average-case around 1969, independently by Heilbronn [22] and Dixon [14]. The centred algorithm has been studied by Rieger [39]. Recently, Vallée has proposed to use the transfer operator, only used so far for the continuous model, in order to analyse Euclidean Algorithms under the discrete model. This approach provides a unifying framework for the average–case analysis of a whole class of Euclidean algorithms [47]. First, this entails a classification of Euclidean algorithms in terms of their average number of steps. On a random input formed with two integers less than N, some of the Euclidean Algorithms are "fast", that is, of logarithmic average–case complexity $\mathrm{E}_N[P] = \Theta(\log N)$ while others are "slow", that is, of the "log-squared" type $\mathrm{E}_N[P] = \Theta(\log^2 N)$ on average. Next, it becomes possible to exhibit precisely the constant μ intervening in the mean value, and, in the case of the Fast Class, to relate it to the derivative $\lambda'_s(1,0)$ of the "dominant eigenvalue" function $s \mapsto \lambda(s,0)$ of the operator $\mathbf{H}_{s,0}$ at $s = 1$. This quantity $\lambda'_s(1,0)$ (by construction independent of cost c) equals the opposite of the entropy of the dynamical system [associated to the algorithm].

All the questions that were asked regarding truncated real trajectories can also be asked for rational trajectories, and they are clearly related to the probabilistic behaviour of the Euclidean Algorithms. It is in particular of great interest to study the total cost $C(x)$ of the rational trajectory relative to some digit-cost c. In [47], Vallée describes a general method that applies to the analysis (on average) of a whole class of digit-costs. When the costs are of moderate growth, and the Euclidean algorithms belong to the Fast Class, it is shown that the rational trajectories have an expected behaviour similar to what is observed on truncated real trajectories. Precisely, in the case of the Fast Class, the total cost of the rational trajectory $C(x)$ relative to some digit–cost c is as follows. If μ is the constant that intervenes in the mean value of the number of steps (that is, the mean number of steps is asymptotic to $\mu \log N$) and $\widehat{\mu}(c)$ is the constant that intervenes in the mean value of real (truncated) trajectories, then the mean value $\mathrm{E}_N[C]$ of cost C on rational trajectories whose inputs have numerator and denominator less than N satisfies

$$\mathrm{E}_N[C] \sim \widehat{\mu}(c) \cdot \mu \log N.$$

Technically, this method is an instance of what may be termed a "dynamical analysis": it relies on a description of relevant parameters by means of generating functions, an otherwise common tool in the average-case of algorithms [19, 21]. As it is usual in number theory contexts, the generating functions are Dirichlet series, and, here, bivariate Dirichlet series, that depend on two parameters s and w. The first parameter s, of Dirichlet type, "marks" the size of inputs, while the second one w, of the type usual in moment generating functions, "marks" the cost. These Dirichlet series, denoted by $S(s,w)$, are first proved to be closely related to the transfer operator $\mathbf{H}_{s,w}$, and the asymptotic extraction of coefficients can be achieved by means of Tauberian theorems [13, 44] provided they *can* be applied. For this purpose, two supplementary properties of the operator are needed: first, the analyticity of the transfer operator $\mathbf{H}_{s,w}$ with respect to (s,w) when (s,w) is near $(1,0)$, namely Property $An_{s,w}(1,0)$; second, Property *SMVL* which expresses that the spectral radius $R(s,0)$ of the transfer operator $\mathbf{H}_{s,0}$ admits a strict maximum on the vertical line $\Re(s) = 1$ at $s = 1$ (see Figure 1). This provides, in conjunction of the two other properties *UDE* and *SG* the convenient estimates on mean values of parameters related to rational trajectories.

Distributional analysis and dynamical methods. We have seen that, with respect to any cost of moderate growth, rational trajectories behave on average similarly to the way truncated real trajectories behave almost everywhere. It is then natural to ask whether this analogy extends to distributions: Is it true that the distribution of cost $C(x)$ on rational trajectories with an input x whose numerator and denominator are less than N is asymptotically Gaussian (when N tends to ∞) ? How to compare the distribution of some cost on truncated real trajectories and on rational trajectories? This paper provides a precise answer to all these questions, for three different algorithms that all belong to the Fast Class.

In classical analysis of algorithms, distributional analyses build upon the tools used for the average-case, that is, generating functions. However, these are then bivariate as one extra parameter needs to be taken care of. Furthermore, for distributional analyses, uniform estimates are needed, whereas only pointwise estimates are sufficient for dealing with average–case. In a sense, distributional analyses are obtained by (uniform) perturbation of average–case analyses, as Flajolet explains in [20]: "Parameters of combinatorial structures, provided they remain 'simple' enough, lead to local deformations (via an auxiliary variable w considered near 0) of the functional relations defining univariate counting generating functions. Under fairly general conditions, such deformations are amenable to perturbation theory and admit of uniform expansions near singularities".
This point of view applies to dynamical methods, and leads to a new facet of the domain—distributional dynamical analysis. This paper provides, for the first time, an instance of the method and applies it to the analysis of three Euclidean Algorithms of the Fast Class. Here the bivariate generating functions, $S(s, w)$ are, as already said, of Dirichlet type with respect to the variable s that marks the size N of the inputs, and we need "uniform" estimates of their coefficients (with respect to w). Previously, Tauberian Theorems sufficed for average-case analyses, since they provided point–wise estimates of coefficients. Now, they become useless for distributional analyses, since they are not capable of providing remainder terms.

We need a more precise "extractor" of coefficients, and the Perron formula is well-suited to this purpose, provided the transfer operator $\mathbf{H}_{s,w}$ itself satisfies good "uniform" properties when $(\Re s, \Re w)$ is near the reference point $(1, 0)$; specifically, Property *USMNVL*, described in Figure 1, is needed. It can be viewed as the uniform perturbation of Property *SMVL*. For this part of our treatment, we adapt powerful methods due to Dolgopyat [15], and Theorem 2 relates this *USMNVL* Property to the *UNI* Property of the dynamical system. The *UNI* Property expresses that the derivatives of inverse branches of the dynamical system are "not too often too close" and is proven to hold for the three Euclidean Dynamical Systems under consideration (Proposition 1). Dologopyat first introduced these tools in order to study the decay of correlations relative to flows. Later on, Pollicott [35, 36], Pollicott and Sharp [37] used them to find error terms in some asymptotic estimates, and Dolgopyat [16] deduced a class of limit theorems for dynamical systems. To best of our knowledge, this paper is the first instance where these powerful tools are applied for distributional analyses in discrete combinatorics. This paper can also be viewed in essence as a realization of Hensley's prophecy : "This suggests that [our results] might be sharpened considerably. Perhaps, what

UDE	Unique Dominant Eigenvalue at $(1,0)$
SG	Spectral Gap at $(1,0)$
$An_w(1,0)$	Analyticity with respect to w at 0 for $s=1$
$An_{s,w}(1,0)$	Analyticity with respect to (s,w) at $(1,0)$
$SMVL$	Strict Maximum of the spectral radius on the Vertical Line $\{(s,0); \Re(s)=1\}$ attained at $s=1$: $\forall s,\ s=\sigma+it,\ t\neq 0, s$ near to $1, R(s,0) < R(\sigma,0)$
$USMNVL$	Uniform Strict Maximum of the spectral radius on the Neighborhood of Vertical Bi-Line $\{(s,w); \Re(s)=1, \Re(w)=0\}$ attained on the real axes: $\exists\gamma<1,\ t_0>0, \forall(s,w),\ s=\sigma+it, w\nu+i\tau, \lvert t\rvert > t_0,\ (\sigma,\nu)$ near to $(1,0)$, $R(s,w) < \gamma R(\sigma,\nu)$

FIGURE 1. Main analytical properties of the transfer operator $\mathbf{H}_{s,w}$ and its spectral radius $R(s,w)$ (on a convenient functional space).

	Average–case	Distribution
Truncated real trajectories	$UDE + SG$	$UDE + SG$ $+\ An_w(1,0)$
Rational trajectories	$UDE + SG$ $+\ An_{s,w}(1,0) + SMVL$	$UDE + SG$ $+\ An_{s,w}(1,0) + USMNVL$

FIGURE 2. Properties of the transfer operator that are needed for analysing trajectories.

is needed is a deeper understanding of the function $\lambda(s)$ and the [transfer] operators".

Main results. This paper offers a general framework leading to new proofs of old results as well as many new results. The new proofs are more natural and more direct than the existing ones. The new results concern a whole class of algorithms and a broad class of cost functions.

We consider some digit–cost c of moderate growth and we first precisely describe the asymptotic distribution of the total cost $C(x)$ of an execution of a Euclidean Algorithm on the input x, when the input x is randomly drawn from rational inputs with denominator and numerator less than N. Theorem 3 proves that, on such inputs, the distribution of C is asymptotically Gaussian, with expectation and variance asymptotically proportional to $\log N$. Furthermore, the constants $\mu(c)$ and $\delta^2(c)$ that intervene in the main terms of the expectation and the variance are proven to involve the first five partial derivatives of the "dominant eigenvalue" function $(s,w)\mapsto\lambda(s,w)$ at $(s,w)=(1,0)$. Thus, they can be expressed as functions of the four constants $\mu,\delta,\widehat{\mu}(c),\widehat{\delta}(c)$. The derivative $\nu(c) := \lambda''_{sw}(1,0)$ is a covariance coefficient that reflects the correlation between the number of steps P and the cost C.

Three special instances of this result are of major interest.

(i) For a trivial digit–cost $c \equiv 1$, the total cost to be studied coincides with the number P of steps of the Euclidean Algorithm. We thus prove that the cost P asymptotically follows a Gaussian law, with a mean and a variance that are both proportional to $\log N$. This fact was already proved by Hensley [23], in the particular case of the Standard Euclidean Algorithm. However, as already said, our proof is both conceptually different and more concise. Furthermore, it applies to a whole class of Euclidean Algorithms.

(ii) When the digit–cost, $c = c_m$, is the characteristic function of the digit m, we obtain the total number of occurrences of digit m in a rational trajectory. We prove this quantity to be asymptotically Gaussian, with a mean-value asymptotic to $\mu \cdot \widehat{\mu}(c_m) \log N$. The constant $\widehat{\mu}(c_m)$ is explicit for the three Euclidean algorithms, and, in the case of the Standard Euclidean Algorithm, it is equal to

$$\widehat{\mu}(c_m) = \frac{1}{\log 2} \log \left(1 + \frac{1}{m(m+2)} \right)$$

(iii) When the digit–cost c is the binary length ℓ of the digit, we prove that the binary encoding of a rational trajectory is asymptotically Gaussian, with a mean-value equal to $\mu \cdot \mu(\ell) \log N$. The constant $\widehat{\mu}(\ell)$ is explicit for the three Euclidean algorithms. It is a variant of the Khinchine constant [28], and, in the case of the standard Euclidean Algorithm, it is equal to

$$\widehat{\mu}(\ell) = \frac{1}{\log 2} \log \prod_{k=0}^{\infty} (1 + \frac{1}{2^k}).$$

Theorem 3 also provides estimates on the speed of convergence towards the limit Gaussian law. It is proven that the optimal possible speed of convergence, of order $O(1/\sqrt{\log N})$, is attained.

Finally, Theorem 4 focuses on the number of steps P. Hensley has already established in [23] that this parameter follows a Local Limit Law. His proof used the transfer operator, but in way appreciably different from what we do here. In fact, Hensley approximates discrete measures on rationals by continuous measures, and he obtains distributional results on rational trajectories by approximating them by truncated real trajectories. Here, we use instead Dirichlet series in conjunction with Perron's formula as "extractor". Like before, estimates à la Dolgopyat are central to the asymptotic analysis and they lead, in conjunction with the saddle point method, to a very natural proof of Hensley's result.

Plan of the paper. Section 2 is an introductory section where we recall descriptions of Euclidean Algorithms and their underlying dynamical systems. We then describe the dynamical analysis methodology and we present our main tools —Dirichlet series, transfer operators, Perron's formula, quasi-powers theorem, saddle point method—. We explain there why it is necessary to get sharp estimates of the transfer operator when the parameter s varies on vertical lines. Section 3 is entirely devoted to obtaining such estimates. We adapt methods of Dolgopyat to our context, and we explain why the *UNI* Condition holds in the Euclidean framework. Then, in Sections 4 and 5, we present the proofs of our main results.

2. Dynamical Analysis Method and main tools.

This section can be viewed as a kind of skeleton for the whole paper. Here, we present the main mathematical objects to be used, and describe the main steps of the analysis, but without detailed proofs. After a description of the three main Euclidean algorithms to be studied (Subsection 2.1), we present the underlying dynamical systems (in 2.2), and introduce the weighted transfer operator (in 2.3). Subsection 2.4 explains (as a useful and exemplary parenthesis) the fundamental rôle which it plays in the distributional analysis of truncated real trajectories (Theorem 1). Next, we turn to the main objects: rational inputs and Euclidean Algorithms. We introduce Dirichlet series of moment generating functions (in 2.5), which we relate to the quasi-inverse of the weighted transfer operator (in 2.6) . The sequel of the Section introduces the following three Sections (3, 4, 5), and summarizes them in an informal way. We explain (in 2.7) why the Perron Formula is useful to extract coefficients, provided it can be applied. Subsection 2.8 introduces the *UNI* Condition which is needed for dealing with Perron's formula and yields the Property *USMNVL*. Then, we explain why this leads to a uniform quasi-powers approximation for the moment generating function of costs, which implies the asymptotic Gaussian distribution for costs on rational trajectories. We then come back to the most basic parameter of a Euclidean Algorithm, its number of steps, and describe how an application of the saddle opoint method entails a local limit theorem for this parameter.

2.1. **Euclidean Algorithms.** Three main Euclidean algorithms are to be analysed; each of them is related to a Euclidean division. The classical division $v = cu + r$ produces a positive remainder r such that $0 \le r < u$; The centered division between u and v $(v > u)$, of the form $v = cu + \epsilon r$, with $\epsilon = \pm 1$ produces a positive remainder r such that $0 \le r < u/2$. The odd division produces an odd quotient: it is of the form $v = cu + \epsilon r$, with $\epsilon = \pm 1, 0 \le r < u$, and c odd.
Instead of the integer pair (u, v), we consider the rational u/v; then, the division expressing the pair (u, v) as a function of the subsequent pair (r, u) is replaced by a linear fractional transformation (LFT in shorthand notation) h that expresses the rational u/v as a function of r/u. For the Centered algorithm (K), the rational u/v is any rational of $\mathcal{I} = [0, 1/2]$. In the two other cases, the rational u/v is any rational of $\mathcal{I} = [0, 1]$.
Finally, on the input (u, v), each of the algorithms performs a sequence of Euclidean divisions. On the input u/v, it performs a sequence of LFT's, and stops when the rational becomes zero. Except for the final step where it makes use of a special set $(\mathcal{F})$, it always uses the same set of LFT's that we call the generic set (see Figure 3).
Each algorithm then builds a specific continued fraction of height p for the rational u/v and decomposes the rational u/v as

$$(u/v) = h_1 \circ h_2 \circ \ldots \circ h_p(0) = h(0), \tag{2.1}$$

with $h_i \in \mathcal{H}, 1 \le i \le p-1$ and $h_p \in \mathcal{F}$. We are interested in various costs related to an execution of the algorithm; the basic one is the number of iterations P, defined by $P(u, v) := p$. However, we study general cost functions C, relative to some cost c (which plays the rôle of a digit–cost), of

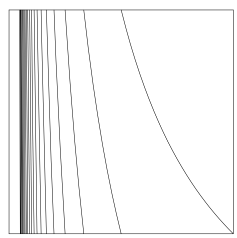
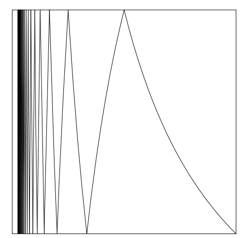
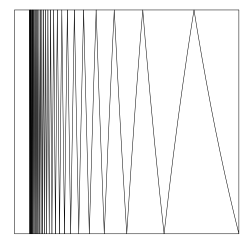
Alg., Interval	(G) $\mathcal{I} = [0,1]$	(K) $\mathcal{I} = [0,1/2]$	(O) $\mathcal{I} = [0,1]$
Division	$v = mu + r$ $0 \le r < u$	$v = mu + \epsilon r, \epsilon = \pm 1,$ $(m, \epsilon) \ge (2, +1)$ $0 \le r < \frac{u}{2}$	$v = mu + \epsilon r,$ m odd, $\epsilon = \pm 1,$ $0 \le r < u$
Generic set	$\mathcal{G} = \{\dfrac{1}{m+x}, \ m \ge 1\}$	$\mathcal{K} = \{\dfrac{1}{m+\epsilon x}, \epsilon = \pm 1,$ $(m,\epsilon) \ge (2,+1)\}$	$\mathcal{O} = \{\dfrac{1}{m+\epsilon x}, \ \epsilon = \pm 1,$ m odd,$(m,\epsilon) \ge (1,1)\}$
Final set $\mathcal{F}$.	$\mathcal{G} \bigcap \{m \ge 2\}$	$\mathcal{K} \bigcap \{\epsilon = 1\}$	$\mathcal{O} \bigcap \{m \ge 3, \epsilon = 1\}$
			

FIGURE 3. The three Euclidean dynamical systems : Standard, Centered, Odd.

the form

$$C(u,v) := \sum_{i=1}^{P(u,v)} c(h_i), \tag{2.2}$$

Such a cost C is said to be additive, and the cost c is called the digit–cost. Remark that the basic cost P is relative to the trivial digit–cost $c = 1$.

2.2. **Euclidean dynamical systems.** We now relate Euclidean algorithms with continued fractions algorithms that can be viewed as their continuous extensions. It proves rewarding to study continued fractions algorithms in the dynamical systems framework.

We consider here dynamical systems of the interval. For nice and elementary surveys, see [7, 4, 31]. A dynamical system (of the interval) $(\mathcal{I}, T)$ consists of an interval $\mathcal{I}$ together with a continuous map $T : \mathcal{I} \to \mathcal{I}$. which is called the shift mapping. Given an initial condition x in $\mathcal{I}$, the sequence $(x, Tx, T^2x, \ldots)$ of iterates of x under the action of T forms the orbit (or the trajectory) of the initial point x. The main study in dynamical systems concerns itself with the interplay between properties of the transformation T and properties of trajectories of points under iteration of the transformation. Here, Continued Fraction algorithms are important particular cases of what is usually called "piecewise complete maps of the interval":

Definition. *A mapping $T : \mathcal{I} \to \mathcal{I}$ is piecewise complete if there exist a (finite or denumerable) set $\mathcal{M}$, whose elements are called digits, and a (topological) partition $\{\mathcal{I}_m\}_{m \in \mathcal{M}}$ of the interval $\mathcal{I}$ into open subintervals $\mathcal{I}_m$*

such that the restriction of T to $\mathcal{I}_m$ is a bijective mapping of class $\mathcal{C}^2$ from $\mathcal{I}_m$ onto $\mathcal{I}$.

A special rôle is played by the set $\mathcal{H}$ of branches of the inverse function T^{-1} of T that are also naturally numbered by the index set $\mathcal{M}$: we denote by $h_{[m]}$ the inverse of the restriction $T|_{\mathcal{I}_m}$, so that $\mathcal{I}_m$ is exactly the image $h_{[m]}(\mathcal{I})$. The set of the inverse branches of the iterate T^k is exactly the set $\mathcal{H}^k$; its elements are of the form $h_{[m_1]} \circ h_{[m_2]} \circ \cdots \circ h_{[m_k]}$ and are called the inverse branches of depth k. The set $\mathcal{H}^\star := \cup_k \mathcal{H}^k$ is the semi-group generated by $\mathcal{H}$. For an inverse branch of depth k, of the form $h_{[m_1]} \circ h_{[m_2]} \circ \cdots \circ h_{[m_k]}$, the interval $h(\mathcal{I})$ gathers all real x for which the k first digits of the *CF* expansion are $(m_1, m_2, \ldots, m_k)$. Such an interval is called a fundamental interval of depth k.

The general framework of piecewise complete maps of intervals is very convenient for a discussion of continued fraction algorithms that are extensions of our three Euclidean algorithms. The interval $\mathcal{I}$ is then defined in Figure 3; the mapping T extends the map defined on rationals by the equality $T(u/v) = r/u$ where r is the remainder of the Euclidean division on (u, v) and it admits the common form

$$T(x) := \left| \frac{1}{x} - A(\frac{1}{x}) \right|,$$

where the function A depends on the algorithm and is defined as

$$\begin{array}{lcl} A_G(x) & := & \text{Integer part of } x; \\ A_K(x) & := & \text{Nearest integer to } x; \\ A_O(x) & := & \text{Nearest odd integer to } x. \end{array}$$

The three Euclidean dynamical systems, presented in Figure 3, then provide continuous extensions of the three Euclidean Algorithms. The inputs of the Euclidean Algorithms give rise to very particular trajectories that are stopped as soon as they attain $x = 0$.

2.3. **Transfer operators.** The behaviour of typical trajectories of dynamical systems is well explained by examining the flow of densities. We consider a general class of dynamical systems which we call the Good Class (see Section 3.1 for a precise definition of this notion). This is a subclass of the class of the last Section, which contains the three Euclidean dynamical systems of interest. For dynamical systems of the Good Class, the map T is strictly expanding, so that the inverse branches are strict contractions, and the contraction ratio, defined as

$$\rho = \left(\max\{|h'(x)|; h \in \mathcal{H}^n, x \in \mathcal{I}\}\right)^{1/n}$$

is strictly less than 1.

If the set $\mathcal{I}$ is endowed with some initial distribution relative to some probability density $f = f_0$, repeated applications of the map T modify the density, and the successive densities $f_1, f_2, \ldots, f_n, \ldots$ describe the global evolution of the system at time $t = 0, 1, 2, \ldots, n, \ldots$. Since the laws governing change do not change with time, there exists an operator $\mathbf{H}$ for which $f_1 = \mathbf{H}[f_0]$, $f_2 = \mathbf{H}[f_1]$, and more generally $f_n = \mathbf{H}[f_{n-1}] = \mathbf{H}^n[f_0]$ for all n. This operator is called the density transformer, or the Perron-Frobenius operator. In the case of a piecewise complete dynamical system of the Good Class, it is defined as

$$\mathbf{H}[f](x) := \sum_{h\in\mathcal{H}} |h'(x)|\ f\circ h(x)$$

and its iterate of order n satisfies an analogous relation, but now with respect to the set $\mathcal{H}^n$,

$$\mathbf{H}^n[f](x) := \sum_{h\in\mathcal{H}^n} |h'(x)|\ f\circ h(x)$$

Ruelle [40, 41] shows that it proves useful to deal with a more general operator, the transfer operator, that depends on some complex parameter s and is denoted by $\mathbf{H}_s$. It satisfies

$$\mathbf{H}_s[f](x) := \sum_{h\in\mathcal{H}} |h'(x)|^s\ f\circ h(x), \qquad \mathbf{H}_s^n[f](x) := \sum_{h\in\mathcal{H}^n} |h'(x)|^s\ f\circ h(x).$$

When the dynamical system belongs to the Good Class, this operator is well-defined as soon as parameter s has a real part $\sigma := \Re s$ which satisfies $\sigma > \sigma_0$, for some $\sigma_0 < 1$. It extends the density transformer since one has $\mathbf{H}_1 = \mathbf{H}$.

We consider a cost function c on $\mathcal{H}$ and we extend it to a cost function c on $\mathcal{H}^\star$ by means of the relation

$$c(h_1\circ h_2\circ\ldots\circ h_p) := \sum_{i=1}^p c(h_i).$$

Once some cost relative to the inverse branch h has been fixed, one defines the weighted composition operator relative to h, that depends on two (complex) parameters s and w,

$$\mathbf{R}_{s,w,h}[f](x) := \exp[wc(h)]\cdot|h'(x)|^s. \tag{2.3}$$

The multiplicative property of derivatives, and the additive property of costs,

$$(h\circ g)'(x) = h'(g(x))\cdot g'(x) \qquad c[h\circ g] = c(h)+c(g)$$

entail together a composition property on operators: given two inverse branches h and g, the operator $\mathbf{R}_{s,w,h\circ g}$ associated to the branch $h\circ g$ satisfies

$$\mathbf{R}_{s,w,h\circ g} = \mathbf{R}_{s,w,g}\circ\mathbf{R}_{s,w,h}. \tag{2.4}$$

In the sequel, we only consider costs of moderate growth that satisfy the following condition $\mathcal{MG}$:

Condition $\mathcal{MG}$ [Moderate Growth]. *The series*

$$\sum_{h\in\mathcal{H}} \exp[wc(h)]\cdot|h'(x)|^s$$

is convergent when $(\Re s, \Re w)$ *belongs to some real neighbourhood* $\Sigma\times W$ *of* $(1,0)$.

Denote by σ_1 the infimum of interval Σ (remark that $\sigma_0\le\sigma_1<1$), by ν_1 the supremum of W, by Σ_1 the interval $]\sigma_1,+\infty]$, and by W_1 the interval $]-\infty,\nu_1[$. When $(\Re s,\Re w)$ belongs to $\Sigma_1\times W_1$, the weighted transferoperator

$$\mathbf{H}_{s,w} := \sum_{h\in\mathcal{H}} \mathbf{R}_{s,w,h}$$

is well–defined and acts on the Banach space $C^1(\mathcal{I})$ endowed with the norm $||.||_{1,1}$ defined by

$$||f||_{1,1} = ||f||_0 + ||f||_1, \qquad \text{with} \quad ||f||_0 := \sup |f|, \quad ||f||_1 := \sup |f'|.$$

In the sequel, we consider trajectories that are relative to some dynamical system $(\mathcal{I}, T)$ of the Good Class. Moreover, we study properties of trajectories with respect to some digit–cost c of moderate growth. In this case, the triple $(\mathcal{I}, T, c)$ will be said to be of $\mathcal{GMG}$-type.

The three Euclidean dynamical systems belong to the Good Class with $\sigma_0 = 1/2$. Furthermore, all the inverse branches of any depth are linear fractional transformations (LFT's) with determinant equal to ± 1, and the denominator $D[h]$ of the LFT h, is closely related to the derivative h': For $h(x) = (ax + b)/(cx + d)$ with a, b, c, d coprime integers, one has

$$D[h](x) := |cx + d| = |\det h|^{1/2}\,|h'(x)|^{-1/2}.$$

Then, the operator $\mathbf{R}_{s,w,h}$ relative to h can be alternatively defined by

$$\mathbf{R}_{s,w,h}[f](x) = \exp[wc(h)]\frac{1}{D[h](x)^{2s}}\; f \circ h(x). \tag{2.5}$$

Since, in the Euclidean instances studied here, the quantity $|h'_{[m]}|$ is $O(m^{-2})$, a sufficient condition for the cost c to be of moderate growth is

$$c(m) := c(h_{[m]}) = O(\log m).$$

Now, given a Euclidean algorithm (amongst $\mathcal{G}, \mathcal{K}, \mathcal{O}$), its generic set $\mathcal{H}$ of LFT's, its final set $\mathcal{F} \subset \mathcal{H}$, and a cost c of class $\mathcal{MG}$ defined on $\mathcal{H}$, the weighted transfer operators relative to $\mathcal{H}, \mathcal{F}$ and c are well–defined on the basic set by

$$\mathbf{H}_{s,w} := \sum_{h \in \mathcal{H}} \mathbf{R}_{s,w,h}. \qquad \mathbf{F}_{s,w} := \sum_{h \in \mathcal{F}} \mathbf{R}_{s,w,h}. \tag{2.6}$$

In particular, thanks to (2.4), the weighted transfer operator relative to set $\mathcal{H}^n$ is the n–th iterate $\mathbf{H}^n_{s,w}$ and the quasi–inverse $(I - \mathbf{H}_{s,w})^{-1}$ plays the rôle of the Ruelle operator relative to the semi-group $\mathcal{H}^\star$.

2.4. **Transfer operator and real trajectories.** We now describe the fundamental rôle that the operator $\mathbf{H}_{s,w}$ plays in the analysis of truncated real trajectories. For some real x of $\mathcal{I}$, we consider its truncated trajectory $\mathcal{T}_n(x) := (x, Tx, T^2x, \ldots T^n x)$, and, for some digit—cost c, the total cost $C^{(n)}$ on the truncated trajectory $\mathcal{T}_n(x)$ of x

$$C^{(n)}(x) := \sum_{i=1}^{n} c(h_i).$$

We are interested in the distribution of cost $C^{(n)}$ when the interval $\mathcal{I}$ is endowed with measure μ with density f. More precisely, we wish to study the asymptotic behaviour of this distribution when the truncation degree

n tends to ∞. A very convenient tool is the moment generating function $M_n(w)$ of cost $C^{(n)}$ which satisfies

$$M_n(w) := \mathrm{E}[\exp(wC^{(n)})] = \sum_{h\in\mathcal{H}^n} \exp[wc(h)]\cdot\mu[h(\mathcal{I})]$$
$$= \sum_{h\in\mathcal{H}^n} \exp[wc(h)]\cdot\int_{h(\mathcal{I})} f(t)dt.$$

With the change of variables $t = h(u)$, the moment generating function $M_n(w)$ can be expressed as a function of the n-th iterate of the transfer operator $\mathbf{H}_{1,w}$,

$$(2.7)\quad M_n(w) = \int_{\mathcal{I}} \sum_{h\in\mathcal{H}^n} \exp[wc(h)]\cdot|h'(u)|\cdot f\circ h(u)du = \int_{\mathcal{I}} \mathbf{H}^n_{1,w}[f](u)du.$$

This relation is fundamental for analysing cost C on truncated real trajectories. In the case of $\mathcal{GMG}$–type, and as soon as w is near 0, the operator $\mathbf{H}_{1,w}$, when acting on $C^I(\mathcal{I})$, has dominant spectral properties; it satisfies Properties *UDE* and *SG*. Moreover, Property $An_w(1,0)$ entails that the iterate of order n of $\mathbf{H}_{1,w}$ asymptotically behaves as the n-th power of its dominant eigenvalue $\lambda(1,w)$. Then, the moment generating function $M_n(w)$ behaves as a Quasi-Power, and, with the so-called Quasi-Powers Theorem, this leads to a Gaussian asymptotic distribution for cost $C^{(n)}$. This result is not new, and does not need the deep estimates à la Dolgopyat. However, our method is somewhat different from methods that are usually used in the literature (see for instance [9, 11]); in particular, the use of moment generating function gives rise to a very short and natural proof (See Section 4.7).

Theorem 1. *For any Euclidean Dynamical System $\mathcal{E}$ amongst $\mathcal{G},\mathcal{K},\mathcal{O}$, and any cost non constant c of moderate growth, the distribution of $C^{(n)}$ on truncated real trajectories of length n is asymptotically Gaussian and the speed of convergence to the Gaussian limit is $O(1/\sqrt{n})$. There exist two constants $\widehat{\mu}(c), \widehat{\delta}(c)$ such that, for any n, and any $y \in \mathbf{R}$, one has*

$$\Pr\left[x \;\middle|\; \frac{C^{(n)}(x) - \widehat{\mu}(c)n}{\widehat{\delta}(c)\sqrt{n}} \le y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-t^2/2}dt + O\left(\frac{1}{\sqrt{n}}\right).$$

The mean and the variance satisfy, for some strictly positive constant β (that depends on cost c),

$$\mathrm{E}\,[C^{(n)}] = \widehat{\mu}(c)\cdot n + \widehat{\mu}_1(c) + O(e^{-n\beta}),$$
$$\mathrm{Var}\,[C^{(n)}] = \widehat{\delta}^2(c)\cdot n + \widehat{\delta}_1(c) + O(e^{-n\beta}),$$

The constants $\widehat{\mu}(c), \widehat{\delta}(c)$ are closely related to the derivatives of function Λ defined by $\Lambda(1,w) = \log\lambda(1,w)$ at $w = 0$, via the equalities

$$\widehat{\mu}(c) = \Lambda'_w(1,0), \qquad \widehat{\delta}^2(c) = \Lambda''_{w^2}(1,0).$$

Moreover, the constants $\widehat{\mu}(c)$ are computable, since they admit the alternative expressions

$$\widehat{\mu}(c) = \sum_{m\in\mathcal{M}} c(m)\int_{\mathcal{I}_m} f_1(u)du,$$

which involve the invariant density f_1 of the associated dynamical system, given in Figure 4.

2.5. **Dirichlet generating functions.** We are interested in the same situation, but we restrict now our study to rational inputs x and we try to follow the same principles as for real truncated trajectories. We consider the following sets

$$\widetilde{\Omega} := \{(u,v);\ \frac{u}{v} \in \mathcal{I}\}, \qquad \Omega := \{(u,v); \gcd(u,v) = 1, \frac{u}{v} \in \mathcal{I}\},$$

$$\widetilde{\Omega}_N := \{(u,v) \in \widetilde{\Omega}; v \leq N\}, \qquad \Omega_N := \{(u,v) \in \Omega; v \leq N\},$$

as the possible inputs of the $\mathcal{E}$ algorithm. Precisely, we deal with the probabilistic models relative to sets $\Omega_N, \widetilde{\Omega}_N$ endowed with uniform probabilities. Then $\Pr_N, \widetilde{\Pr}_N$ denote the associated probabilities, and $\mathrm{E}_N, \widetilde{\mathrm{E}}_N$ denote the expectations. When given an input pair (u,v), the $\mathcal{E}$ algorithm decomposes the rational $x = u/v$ as in Equation (2.1) and its total cost $C(u,v)$, associated to some digit–cost c, is defined in (2.2). More precisely, we wish to study the distribution of this cost, and our main tool is its moment generating function on Ω_N or on $\widetilde{\Omega}_N$,

$$\mathrm{E}_N[\exp(wC)] := \frac{C_N(w)}{C_N(0)}, \qquad \widetilde{\mathrm{E}}_N[\exp(wC)] := \frac{\widetilde{C}_N(w)}{\widetilde{C}_N(0)} \tag{2.8}$$

that involves the cumulative values of $\exp(wC)$ on Ω_N or on $\widetilde{\Omega}_N$,

$$C_N(w) := \sum_{(u,v)\in\Omega_N} \exp[wC(u,v)], \qquad \widetilde{C}_N(w) := \sum_{(u,v)\in\widetilde{\Omega}_N} \exp[wC(u,v)]. \tag{2.9}$$

(Remark that $C_N(0) = |\Omega_N|, \widetilde{C}_N(0) = |\widetilde{\Omega}_N|$.) Following the ideas previously used in [46, 48, 49], we opt for Dirichlet series. Here, we replace this sequence of moment generating functions by a unique Dirichlet generating function, henceforth called the Dirichlet–moment generating function, that depends on two parameters s and w,

$$S(s,w) := \sum_{(u,v)\in\Omega} \frac{1}{v^s} \exp[wC(u,v)] = \sum_{n\geq 1} \frac{c_n(w)}{n^s}, \tag{2.10}$$

$$\widetilde{S}(s,w) := \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{v^s} \exp[wC(u,v)] = \sum_{n\geq 1} \frac{\widetilde{c}_n(w)}{n^s}, \tag{2.11}$$

where $c_n(w)$, $\widetilde{c}_n(w)$ are the cumulative values of $(u,v) \mapsto \exp[wC(u,v)]$ on pairs (u,v) of Ω or $\widetilde{\Omega}$ with fixed $v = n$. Then, the relations

$$C_N(w) = \sum_{n\leq N} c_n(w), \qquad \widetilde{C}_N(w) = \sum_{n\leq N} \widetilde{c}_n(w) \tag{2.12}$$

show that the moment generating function of cost C on Ω_N or on $\widetilde{\Omega}_N$ involve sums of coefficients of series $S_C(s,w), \widetilde{S}_C(s,w)$ under the form

$$\mathrm{E}_N[\exp(wC)] = \frac{\sum_{n\leq N} c_n(w)}{\sum_{n\leq N} c_n(0)}, \qquad \widetilde{\mathrm{E}}_N[\exp(wC)] = \frac{\sum_{n\leq N} \widetilde{c}_n(w)}{\sum_{n\leq N} \widetilde{c}_n(0)}.$$

In the same vein, and in the particular case when the cost of interest is the number P of iterations, we also wish to evaluate the distribution function itself, on Ω_N or in $\widetilde{\Omega}_N$,

$$\Pr_N[P = p], \qquad \widetilde{\Pr}_N[P = p].$$

In this case, the associated generating functions are

$$S_p(s) := \sum_{(u,v)\in\Omega} \frac{1}{v^s}\, |\{P(u,v) = p\}| = \sum_{n\geq 1} \frac{c_n^{(p)}}{n^s}, \tag{2.13}$$

$$\widetilde{S}_p(s) := \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{v^s}\, |\{P(u,v) = p\}| = \sum_{n\geq 1} \frac{\widetilde{c}_n^{(p)}}{n^s},$$

where $c_n^{(p)}$, $\widetilde{c}_n^{(p)}$ are the number of elements (u,n) of Ω or $\widetilde{\Omega}$ for which $P(u,n) = p$. Then, the relations

$$\Pr_N[P = p] = \frac{1}{|\Omega_N|} \sum_{n\leq N} c_n^{(p)}, \qquad \widetilde{\Pr}_N[P = p] = \frac{1}{|\widetilde{\Omega}_N|} \sum_{n\leq N} \widetilde{c}_n^{(p)} \tag{2.14}$$

show that the distribution of cost P on Ω_N or on $\widetilde{\Omega}_N$ involves sums of coefficients of series $S_p(s), \widetilde{S}_p(s)$ together with series $S(s,0), \widetilde{S}(s,0)$.

2.6. **Dirichlet series and transfer operator.** As we previously did for cost $C^{(n)}$ defined on truncated real trajectories, we aim to relate the moment generating function of costs on rational trajectories to the weighted transfer operator. An execution of the Euclidean algorithm on the input (u,v) of Ω performing p steps decomposes the rational (u/v) as

$$(u/v) = h_1 \circ h_2 \circ \ldots \circ h_p(0) = h(0), \tag{2.15}$$

with $h_i \in \mathcal{H}, 1 \leq i \leq p-1$ and $h_p \in \mathcal{F}$. Then, each Euclidean algorithm defines a bijection between the sets Ω and $\mathcal{H}^\star \times \mathcal{F}$, so that, the following relations hold between the LFT h and the pair (u,v)

$$v = D[h](0), \qquad C(u,v) = c(h), \qquad P(u,v) = p.$$

These relations provide an alternative form for the Dirichlet moment generating function $S(s,w), \widetilde{S}(s,w)$ that involves the weighted transfer operators $\mathbf{H}_{s,w}$ and $\mathbf{F}_{s,w}$,

$$S(2s,w) = \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0), \tag{2.16}$$

$$\widetilde{S}(2s,w) = \zeta(2s)\, \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0). \tag{2.17}$$

The second relation is obtained since each element (u',v') of $\widetilde{\Omega}$ can be written in a unique way as (du, dv) with $d = \gcd(u,v)$, $(u,v) \in \Omega$. Then, an execution of the $\mathcal{E}$ algorithm on (u',v') leads to the same decomposition as in (2.15) and the two costs $C(u',v')$ and $C(u,v)$ are equal.

For an element (u,v) of Ω, the relation $P(u,v) = p$ holds if and only if the decomposition $u/v = h(0)$ involves some LFT h that belongs to $\mathcal{H}^{p-1} \times \mathcal{F}$. Furthermore, two pairs (u,v) and (du, dv) admit the same decomposition

(2.15). This again entails fundamental relations between the Dirichlet series $S_p(s)$ and the p-th iterate of the plain transfer operator $\mathbf{H}_s$,

$$S_p(2s) = \mathbf{F}_s \circ \mathbf{H}_s^{p-1}[1](0), \quad \widetilde{S}_p(2s) = \zeta(2s)\mathbf{F}_s \circ \mathbf{H}_s^{p-1}[1](0). \tag{2.18}$$

The relations (2.16), (2.17) and (2.18) are the analogues for rational trajectories of the relation (2.7) for the truncated real trajectories.

However, in the case of rational trajectories, we have to extract the coefficients of these Dirichlet series. This is why the discrete problem is more difficult to solve than the continuous problem. The whole paper is devoted to this task. The sequel of this Section is now only introductory, and our intention, until the end of the Section, is to state our results as three Theorems (Theorems 2, 3, 4) and describe the main tools, together with the logical chain of ideas that link them, in an informal way.

2.7. **Perron's formula.** Our first main tool is the Perron Formula. The classical identity

$$\frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} \frac{y^{s+1}}{s(s+1)}ds = \begin{cases} 0 & \text{if } \ 0 < y \le 1, \\ y-1 & \text{if } \ y > 1, \end{cases}$$

holds for any $D > 0$. When applied term by term to a Dirichlet series of the form $F(s) = \sum_{n\ge 1} a_n n^{-s}$ for some D inside the domain of convergence of F, it entails the equality, well-known as Perron's formula of order 2 (see for instance [18]),

$$\Psi(T) := \sum_{n\le T} a_n(T-n) = \frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} F(s)\frac{T^{s+1}}{s(s+1)}ds, \tag{2.19}$$

which is a first step for estimating the sum $\sum_{n\le N} a_n$ of the coefficients of $F(s)$.

We recall that we wish to evaluate the sum $C_N(w)$ of the first N coefficients of series $S(2s,w)$, and we first apply this identity to the Dirichlet series $S(2s,w)$. In this way, we shall evaluate

$$\Psi_w(T) := \sum_{n\le T} c_n(w)(T-n) = \frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} S(s,w)\frac{T^{2s+1}}{s(2s+1)}ds, \tag{2.20}$$

and the relation

$$\frac{1}{N}\Psi_w(N) = \frac{1}{N}\sum_{Q\le N}\sum_{n\le Q} c_n(w)$$

shows that $\Psi_w(N)$ is just a Cesaro average of coefficients $C_Q(N)$ for Q less than N.

We work with values of w near 0. For any w fixed (near 0), the domain of convergence of $S(2s,w)$ (considered as a function of s) is limited by the rightmost values s for which the transfer operator $\mathbf{H}_{s,w}$ (when acting on $C^1(\mathcal{I})$) admits 1 as an eigenvalue. More precisely, the dominant singularity of $s \mapsto S(2s,w)$ arises at $s = \sigma(w)$ where $\sigma(w)$ is the unique complex s near 1 for which the dominant eigenvalue $\lambda(s,w)$ of the operator $\mathbf{H}_{s,w}$ equals 1: Under Condition $\mathcal{MG}$, the operator $\mathbf{H}_{s,w}$ is analytic in the neighbourhood of $(1,0)$; for $(s,w) = (1,0)$, the operator $\mathbf{H}_{s,w}$ is exactly the density transformer, and it has a dominant eigenvalue equal to 1. Then, by perturbation theory, the function "dominant eigenvalue" $(s,w) \mapsto \lambda(s,w)$ is well–defined and analytic when (s,w) is close enough to $(1,0)$. Moreover, when w (complex) is near to 0, the Complex Implicit Function Theorem shows that, for each w near 0, there exists a unique value $s = \sigma(w)$ near 1 for which $\lambda(\sigma(w), w) = 1$. The function $w \mapsto \sigma(w)$ is analytic for w close enough to 0.

Finally, the integral defined in (2.20) is well defined as soon as $D > \Re(\sigma(w)$.

2.8. **Dolgopyat's estimates.** Then, the idea is to change the integration contour into a contour that contains $\sigma(w)$ as a unique pole of $S(2s,w)$. This change (of the integration contour) is possible provided the integral defined in (2.20) exists when D is on the left of $\Re\sigma(w)$, being of the form $D = \Re\sigma(w) - \alpha$, with $\alpha > 0$. Moreover, since we need uniform estimates when w is a complex close enough to 0, we study these integrals for $\alpha > 0$ independent of w. Writing $s = \sigma + it$, a sufficient condition which implies the existence of the integral is that $S(2s,w)$ is $O(|t|^\epsilon)$ (for some $\epsilon, 0 < \epsilon < 1$) when $\Re(s) = \Re\sigma(w) - \alpha$, $|t| = |\Im s| \geq t_0 > 0$ and w near 0. Moreover, we look for such an estimate that is "uniform" when w is a complex close enough to 0. This condition is closely related to the Property *USMNVL*; it is often difficult to obtain for general Dirichlet series.

Here, we adapt to our context powerful methods due to Dolgopyat. The main idea of Dolgopyat is to introduce what is known as the *UNI* Condition on inverse branches of the dynamical system $(\mathcal{I}, T)$ (See Section 3.4). This condition allows us to control oscillatory integrals associated to iterates of transfer operators. Here, we work in the context of general dynamical systems (we do not consider only our three Euclidean cases, but we wish to apply later the general setting to our particular setting), and we have to adapt these ideas; first, we have to consider dynamical systems which may possess an infinite number of branches; second, we work with weighted transfer operators. We shall obtain in Section 3 the following result that deals with a general $\mathcal{GMG}$ setting. It provides convenient estimates and proves that, in this general setting, the *UNI* Condition implies the *USMNVL* Property.

Theorem 2. *Consider a general triple $(\mathcal{I}, T, c)$ of $\mathcal{GMG}$–type (Good Class and Moderate Growth). Denote by $\rho < 1$ the contraction ratio of the dynamical system $(\mathcal{I}, T)$. Suppose furthermore that UNI Condition holds for the Dynamical System $(\mathcal{I}, T)$. Consider, for a pair (s,w) with $s = \sigma + it, t \neq 0, w = \nu + i\tau$, the weighted transfer operator $\mathbf{H}_{s,w}$ relative to the triple $(\mathcal{I}, T, c)$,*

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

acting on the Banach space $C^1(\mathcal{I})$, with respect to the norm

$$||f||_{1,t} = \sup |f| + \frac{\sup |f'|}{|t|}\,.$$

Denote, for real s and w, its dominant eigenvalue by $\lambda(s,w)$. Then, for any $\epsilon > 0$, there exist a (real) neighborhood $\Sigma \times W$ of $(1,0)$, as well as constants M, γ with $M > 0$ and $0 < \gamma < 1$ such that, for all (s,w) with

$(\Re s, \Re w) \in \Sigma \times W$ and $|\Im s| \geq 1/\rho^2$, and for all n, the n-th iterate of the operator $\mathbf{H}_{s,w}$ satisfies

$$\|\mathbf{H}_{s,w}^n\|_{1,t} \leq M \cdot \gamma^n \cdot |t|^\epsilon \cdot \lambda(\sigma,\nu)^n \qquad (s = \sigma + it, w = \nu + i\tau)\,. \tag{2.21}$$

For any $\epsilon > 0$, there exist a (real) neighborhood $\Sigma \times W$ of $(1,0)$ and a constant $M_1 > 0$ so that, for all (s,w) with $(\Re s, \Re w) \in \Sigma \times W$ and $|\Im s| \geq 1/\rho^2$, the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ satisfies

$$\|(I - \mathbf{H}_{s,w})^{-1}\|_{1,t} \leq M_1 \cdot |t|^\epsilon\,. \tag{2.22}$$

This result is the central functional analytic result of the paper. Section 3 will be completely devoted to its proof. Note that the Spectral Radius Formula together with Inequality (2.21) implies that the inequality

$$R(s,w) \leq \gamma R(\sigma,\nu)$$

holds for $(\Re s, \Re w) \in \Sigma \times W$ and $|\Im(s)| \geq t_0$, so that the *UNI* Property entails the Property *USMNVL*. It proves also that the Dirichet series $S(2s,w)$ admit $s = \sigma(w)$ as a unique pole in a strip $|\Re s - \Re\sigma(w)| \leq \alpha$, for α sufficiently small.

Section 3 proves also that our three Euclidean dynamical systems satisfy the *UNI* Condition, so that estimates à la Dolgopyat can be used in our framework.

2.9. **Asymptotic normality and Quasi-powers Estimates.** Finally, Perron's Formula combined with Dolgopyat estimates provides an estimate for

$$\Psi_w(T) := \sum_{n \leq T} c_n(w)(T - n)$$

that involves the residue $A(w)$ of $S(2s,w)$ at the pole $s = \sigma(w)$, under the form (when $T \to \infty$),

$$\Psi_w(T) = \frac{A(w)}{\sigma(w)(2\sigma(w)+1)}\, T^{2\sigma(w)+1}\, [1 + \widehat{\epsilon}_w(T)].$$

Here, functions σ and A are analytic on a complex neighbourhood $\mathcal{W}$ of $w = 0$, and $\widehat{\epsilon}_w(T)$ tends to 0 when $T \to \infty$ uniformly on $\mathcal{W}$. Section 4 will proceed in three main steps for obtaining the proof of the asymptotic normality of cost C.

First Step. We first transfer the estimates for $\Psi_w(T)$ into estimates for the moment generating function of some "smoothed" version $\overline{C}$ of cost C,

$$\mathrm{E}_N[\exp(w\overline{C})] := [1 + \epsilon_w(N)] \exp D_N(w)$$

where

$$D_N(w) = 2(\sigma(w) - \sigma(0)) \log N + \log \frac{A(w)}{A(0)\sigma(w)},$$

and $\epsilon_w(T)$ uniformly tends to 0. Such an estimate is called a quasi-power. Now, we are in a general and useful framework in which an asymptotically Gaussian law can be expected. See Hwang's papers [24, 25, 26]. It proves, in this case, that the cost function $\overline{C}$ follows an asymptotic Gaussian distribution with a speed of convergence in $O(1/\sqrt{\log N})$. Moreover, this theorem provides precise informations about the asymptotic behaviour of the expectation $\mathrm{E}_N[\overline{C}]$ and the variance $\mathrm{Var}_N[\overline{C}]$.

Second Step. We are interested in studying the cost C, and not really its smoothed version $\overline{C}$. In fact, the distributions of $\overline{C}$ and C are sufficiently close, so that we can prove that the distribution of C is also asymptotically gaussian, with a speed of convergence in $O(1/\sqrt{\log N})$. However, this comparison does not provide any information about the asymptotic behaviour of the expectation $\mathrm{E}_N[C]$ and the variance $\mathrm{Var}_N[C]$.

Third Step. We get informations on moments by directly studying the Dirichlet generating functions $S^{(1)}(s)$ or $S^{(2)}(s)$ respectively relative to costs C and C^2. They are the first two derivatives of the Dirichlet moment generating function $S(s, w)$ with respect to w at $w = 0$. Taking the first two derivatives of the fundamental relation (2.16)

$$S(2s, w) = \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0), \tag{2.23}$$

provides alternative expressions of $S^{(1)}(s)$ and $S^{(2)}(s)$ that involve the quasi-inverse of the plain transfer operator $\mathbf{H}_s$, together with derivatives (with respect to w, at $w = 0$) of $\mathbf{F}_{s,w}$ and $\mathbf{H}_{s,w}$. Then, we use the same methods as previously (Perron's formula, Dolgopyat estimates) that entail asymptotic estimates for functions Ψ related to these series. The transfer between functions Ψ and Φ is now easier, since series $S^{(1)}(s)$ and $S^{(2)}(s)$ have positive coefficients.

We finally obtain our second main result whose detailed proof is found in Section 4.

Theorem 3. *Consider any Euclidean Algorithm amongst $\mathcal{G}, \mathcal{K}, \mathcal{O}$, and any cost c of moderate growth. Then the following holds:*
(a) The distribution of the total cost C on Ω_N or on $\widetilde{\Omega}_N$ is asymptotically Gaussian, and the speed of convergence to the Gaussian limit law is $O(1/\sqrt{\log N})$. There exist two constants $\mu(c)$ and $\delta(c)$ such that, for any N, and any $y \in \mathbf{R}$, one has

$$\Pr_N\left[(u, v) \;\middle|\; \frac{C(u, v) - \mu(c)\log N}{\delta(c)\sqrt{\log N}} \le y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-t^2/2}dt + O\left(\frac{1}{\sqrt{\log N}}\right).$$

The same holds for $\widetilde{\Pr}_N$ in $\widetilde{\Omega}_N$.
(b) The mean and the variance satisfy, for some strictly positive constant α (that depends on cost c),

$$\mathrm{E}_N[C] = \mu(c)\log N + \mu_1(c) + O(N^{-\alpha}),$$
$$\mathrm{Var}_N[C] = \delta^2(c)\log N + \delta_1(c) + O(N^{-\alpha}),$$

The same holds for $\widetilde{\mathrm{E}}_N$ and $\widetilde{\mathrm{Var}}_N$ in $\widetilde{\Omega}_N$.
(c) In the case $c = 1$, the constants μ and δ involve the first and second derivatives of the dominant eigenvalue function $\lambda(s) : s \mapsto \lambda(s, 0)$ at $s = 1$,

$$\mu = \frac{2}{|\lambda'(1)|}, \qquad \delta^2 = \frac{2}{|\lambda'^3(1)|}\left(\lambda''(1) - \lambda'^2(1)\right). \tag{2.24}$$

The constant $|\lambda'(1)|$ equals the entropy of the associated dynamical system. In the general case, the constants $\mu(c), \delta(c)$ are closely related to the constants $\widehat{\mu}(c), \widehat{\delta}(c)$ that intervene in the limit behaviour of cost c on truncated *real trajectories, via the equalities*

$$\mu(c) = \mu \cdot \widehat{\mu}(c), \qquad \delta^2(c) = \widehat{\mu}(c)^2 \cdot \delta^2 + \mu \cdot \widehat{\delta}^2(c) + \mu^2 \cdot \nu(c),$$

where $\nu(c)$ is a covariance factor which is strictly positive. The constants $\mu(c)$ are computable.

2.10. **Local Limit Theorem and saddle point method.** We summarize now Section 5 which deals with the particular case of a trivial digit–cost. Then, the total cost C equals the number of iterations P. We apply Perron's formula to the Dirichlet series $S_p(s)$, and with (2.13)

$$\Psi^{(p)}(T) := \sum_{n \le T} c_n^{(p)}(T-n) = \frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} \mathbf{F}_s \circ \mathbf{H}_s^{p-1}[1](a)\frac{T^{2s+1}}{s(2s+1)}ds.$$

Our objective is a local limit theorem, that is, an estimate of the probability $\Pr_N[P(u,v) = p]$ where parameters p and N are related by

$$p = p(N) = \lfloor \mu \log N + x\delta\sqrt{\log N}\rfloor$$

where x varies and μ and δ are the constants that intervene in the main terms of the asymptotic expressions of the mean and the standard deviation of parameter P (see Equation (2.24)). Moreover, it proves convenient to use the vertical line $\Re(s) = 1$ as integration contour. Finally, we deal with integrals of the form

$$I_N := \frac{1}{2i\pi}\int_{1-i\infty}^{1+i\infty} \mathbf{F}_s \circ \mathbf{H}_s^{p(N)}[1](0)\frac{N^{2s+1}}{s(2s+1)}ds,$$

and we need to describe their asymptotic behaviour for $N \to \infty$. As previously, Dolgopyat estimates will play a fundamental role here, since they prove (via Lemma 12) that the dominant part of the integral is brought by the neighborhood of the real axis. Then, near the real axis, a large power of the transfer operator $\mathbf{H}_s^{p(N)}$ behaves as its dominant eigenvalue $\lambda(s)^{p(N)}$. Finally, the principal function to be studied near $s = 1$ is $\lambda(s)^{p(N)}N^{2s+1}$: it is split into two factors

$$\exp[(\mu\Lambda(s) + (2s+1))\log N] \cdot \exp[\delta x \Lambda(s)\sqrt{\log N}], \quad \text{with } \Lambda(s) := \log\lambda(s).$$

Now, by definition of constant μ, the point $s = 1$ is a saddle point for the first factor, and the general principles of saddle point analysis [10] provide the estimates for $\Psi^{(p(N))}(N)$. We then return to

$$\Phi^{(p(N))}(N) := \sum_{n \le N} c_n^{(p(N))}$$

by the same elementary arguments as previously.

We then obtain our third main result (proven in Section 5). It includes Hensley's earlier result [23] as a special case (Standard Algorithm), and provides an alternative proof for this result.

Theorem 4. *For any Euclidean Algorithm $\mathcal{E}$ amongst $\mathcal{G}, \mathcal{K}, \mathcal{O}$, the number P of iterations on Ω_N or $\widetilde{\Omega}_N$ follows a Local Limit Law. If μ and δ^2 are* *the constants that appear in the main terms of the mean and the standard deviation of P, one has*

$$\Pr_N[P = \lfloor \mu \log N + \delta x \sqrt{\log N} \rfloor] = \frac{e^{-x^2/2}}{\delta\sqrt{2\pi \log N}} \left[1 + O\left(\frac{(\log\log N)^{3/4}}{(\log N)^{1/4}} \right) \right]$$

uniformly for $|x| \le \theta(\log\log N)^{1/2}$ *with* θ *sufficiently small. The same holds for* $\widetilde{\Pr}_N$ *in* $\widetilde{\Omega}_N$.

3. *UNI* Property and Dolgopyat estimates.

This whole Section is entirely devoted to obtaining estimates à la Dolgopyat. We prove here Theorem 2. We use here many ideas due to Dolgopyat [15]; however, we have to adapt these ideas to our context that deals with an infinite number of branches, and weighted transfer operators. We first present the so-called $\mathcal{GMG}$–class of triple $(\mathcal{I}, T, c)$ formed with a dynnamical system $(\mathcal{I}, T)$ of the Good Class and a cost c of moderate growth; we describe the main properties of the weighted transfer operator in this context. We then introduce two technical tools, the normalized operator, and a slightly unusual norm, the $(1,t)$–norm. We establish four preliminary results in Lemmata 1, 2, 3 and 4. Subsection 3.4 is central since it shows how the *UNI* Condition entails a convenient estimate for the L^2–norm of the operator. Then, some bound transfers are operated in 3.5, and lead to the *USMNVL* Property. Finally Subsection 3.6 proves that the *UNI* Condition holds for the three Euclidean Systems to be studied.

3.1. **Classical properties of weighted transfer operators.** We consider a general class of dynamical systems that contains the three Euclidean dynamical systems to be studied. This is the so-called Good Class.

Definition [Good Class]. *The dynamical system* $(\mathcal{I}, T)$ *belongs to the Good Class if the following is true:*

(i) $T : \mathcal{I} \to \mathcal{I}$ *is a piecewise complete mapping that is furthermore uniformly expanding. Piecewise uniform expansion means that there are* C *and* $\widehat{\rho} < 1$ *so that* $|h'(x)| \le C\widehat{\rho}^n$ *for every inverse branch* h *of* T^n*, all* n *and all* $x \in \mathcal{I}$*. The smallest* ρ *that satisfies this property is called the contraction ratio, and is defined as*

$$\rho = \left(\max\{|h'(x)|; h \in \mathcal{H}^n, x \in \mathcal{I}\}\right)^{1/n}. \tag{3.1}$$

(ii) T *is topologically mixing: For any pair of non-empty sets* (V, W)*, there exists* n_1 *such that* $T^{-n}V \cap W \neq \emptyset$ *for* $n \ge n_1$.

(iii) There is a distortion constant $\widehat{K} > 0$ *so that every inverse branch* h *of* T *satisfies* $|h''(x)| \le \widehat{K}|h'(x)|$ *for all* $x \in \mathcal{I}$*. (This is automatic when there are finitely many branches, and is just Renyi's condition expressed in terms of the inverse branches otherwise);*

(iv) There is a $\sigma_0 < 1$ *such that* $\sum_{h \in \mathcal{H}} \sup |h'|^\sigma < \infty$ *for all* $\sigma > \sigma_0$.

We recall that $\mathcal{H}^n$ is the set of the inverse branches of T^n, and $\mathcal{H}^\star$ the set of all inverse branches (of any depth). We consider the distortion constant $K = \widehat{K}/(1-\rho)$. Then, $|h''(x)| \le K|h'(x)|$ for all $x \in \mathcal{I}$ and for all $h \in \mathcal{H}_\star$. This bounded distortion property will play an important rôle in the sequel;

since the ratio $h''(x)/h'(x)$ is the derivative of $\log|h'(x)|$, there exists $L := e^K$ for which

$$\frac{1}{L} \le \frac{|h'(x)|}{|h'(y)|} \le L \qquad \text{for all } x, y \in \mathcal{I}, \text{ for all } h \in \mathcal{H}^\star. \tag{3.2}$$

We recall that a digit–cost c is said to be of moderate growth if the series

$$\sum_{h\in\mathcal{H}} \frac{\exp[wc(h)]}{D[h](x)^{2s}} = \sum_{h\in\mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s, \qquad x \in \mathcal{I}$$

is convergent when $(\Re s, \Re w)$ belongs to some (real) neighbourhood $\Sigma \times W$ of $(1,0)$. The class of such costs is denoted by $\mathcal{MG}$. In the following, we study trajectories created by some dynamical system of the Good Class, with a digit-cost which belongs to the Class $\mathcal{MG}$. In this case, the system $(\mathcal{I}, T, c)$ is said to be $\mathcal{GMG}$ –Good and Moderate Growth–.

Our object of study is the weighted transfer operator relative to a $\mathcal{GMG}$ system,

$$\mathbf{H}_{s,w}[f](x) = \sum_{h\in\mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x)\,. \tag{3.3}$$

For a complete study about transfer operators, see [3].

Denote by σ_1 the infimum of interval Σ (remark that $\sigma_0 \le \sigma_1 < 1$), by ν_1 the supremum of W, by Σ_1 the interval $]\sigma_1, +\infty]$, and by W_1 the interval $]-\infty, \nu_1[$. When $(\Re s, \Re w)$ belongs to $\Sigma_1 \times W_1$, the operator $\mathbf{H}_{s,w}$ is well–defined and acts on the Banach space $C^1(\mathcal{I})$ with

$$\|f\|_{1,1} = \|f\|_0 + \|f\|_1, \qquad \text{with} \qquad \|f\|_0 := \sup|f|, \|f\|_1 := \sup|f'|;$$

(see the computations in the proof of Lemma 3).

Moreover, for $(\Re s, \Re w) \in \Sigma_1 \times W_1$, Properties $(i), (iii)$ and (iv) prove that the operator $\mathbf{H}_{s,w}$ is quasi–compact. We recall now this notion: For an operator $\mathbf{L}$, denote by $\mathrm{Sp}(\mathbf{L})$ its spectrum, by $R(\mathbf{L})$ its spectral radius, i.e., the supremum of moduli $|\lambda|$ when λ is an element of $\mathrm{Sp}(\mathbf{L})$, and by $R_e(\mathbf{L})$ its essential spectral radius, i.e., the smallest positive number r such that any eigenvalue λ of $\mathrm{Sp}(\mathbf{L})$ with modulus $|\lambda| > r$ is an isolated eigenvalue of finite multiplicity. For compact operators, the essential radius equals 0. An operator $\mathbf{L}$ is quasi-compact if the strict inequality $R_e(\mathbf{L}) < R(\mathbf{L})$ holds. Then, except for the part of the spectrum inside the closed disk of radius $R_e(\mathbf{L})$, the operator behaves just like a compact operator (in the sense that its spectrum consists of isolated eigenvalues of finite multiplicity).

For real values of parameters $s = \sigma$ and $w = \nu$, $(\sigma, \nu) \in \Sigma_1 \times W_1$, Property (ii) then implies Properties UDE : the operator $\mathbf{H}_{\sigma,\nu}$ has a unique (simple) eigenvalue $\lambda(\sigma, \nu)$ of maximal modulus. Furthermore, there exists an associated eigenfunction $f_{\sigma,\nu}$ which is strictly positive. Moreover, since the operator is quasi-compact, Property SG is also fulfilled, and there is a spectral gap between $\lambda(\sigma,\nu)$ and the rest of the spectrum. Then, the subdominant spectral radius

$$r_{\sigma,\nu} := \sup\{|\lambda|; \lambda \in \mathrm{Sp}(\mathbf{H}_{\sigma,\nu}), \lambda \ne \lambda(\sigma,\nu)\}$$

satisfies $r_{\sigma,\nu} < \lambda(\sigma,\nu)$. The eigenvector $\widehat{\mu}_{\sigma,\nu}$ of the adjoint operator $\mathbf{H}^*_{\sigma,\nu}$ relative to $\lambda(\sigma,\nu)$ is Lebesgue measure for $(\sigma,\nu) = (1,0)$, and for all (σ,ν)

near $(1,0)$ a positive Radon measure $\widehat{\mu}_{\sigma,\nu}$. We may assume $\widehat{\mu}_{\sigma,\nu}[1]=1$ and $\widehat{\mu}_{\sigma,\nu}[f_{\sigma,\nu}]=1$ so that $\mu_{\sigma,\nu}:=f_{\sigma,w}\widehat{\mu}_{\sigma,w}$ is a probability measure.

Note that $\mathbf{H}_{s,w}:C^1(\mathcal{I})\to C^1(\mathcal{I})$ depends analytically on (s,w) for $(\Re s,\Re w)\in\Sigma_1\times W_1$. Then Property $An_{s,w}(1,0)$ holds and Perturbation Theory applies. The quantities $\lambda(s,w)^{\pm 1}$, $f^{\pm 1}_{s,w}$, and $f'_{s,w}$ are well-defined and analytic when (s,w) is near $(1,0)$. On the other hand, the quantities $\lambda(\sigma,\nu)^{\pm 1}$, $f^{\pm 1}_{\sigma,\nu}$, and $f'_{\sigma,\nu}$ depend analytically on (σ,ν) for $(\sigma,\nu)\in\Sigma_1\times W_1$, and therefore are uniformly bounded in any compact subset of $\Sigma_1\times W_1$. Also, for any $(\sigma,\nu)\in\Sigma_1\times W_1$, the map $\sigma\mapsto\lambda(\sigma,\nu)$ is strictly decreasing. Furthermore, the maps $\sigma\mapsto\log\lambda(\sigma,\nu)$ and $\nu\mapsto\log\lambda(\sigma,\nu)$ are both convex. When the system has an infinite number of branches, the first one is strictly convex, as is the second one provided it is associated to a non–constant digit–cost (see for instance [45] for a proof).
In the main quantities relative to the plain operator $\mathbf{H}_s:=\mathbf{H}_{s,0}$, we omit the second index $w=0$. For instance, $\lambda(s)$ replaces $\lambda(s,0)$, etc .... Since $\mathbf{H}_{1,0}=\mathbf{H}_1$ is the density transformer, one has $\lambda(1)=1$, and the function f_1 is invariant under action of $\mathbf{H}_1$. Furthermore, the derivative $\lambda'(1)=\lambda'_s(1,0)$ equals the opposite of the entropy of the dynamical system. Thanks to Rohlin's formula, this derivative is also equal to

$$\lambda'(1)=\int_{\mathcal{I}}\mathbf{H}'_1[f_1](x)dx.$$

When w is a (complex) number near 0, relations $\lambda(1,0)=1$, $\lambda'_s(1,0)\neq 0$, combined with the complex implicit function theorem, prove that there exists a unique complex $s=\sigma(w)$ near 1 for which $\lambda(\sigma(w),w)=1$. The function $w\mapsto\sigma(w)$ is analytic near 0, its real restriction is real and satisfies $\sigma(0)=1$.

It is clear that the three Euclidean systems, whose main characteristics are recalled in Figure 4, together with a digit–cost c that satisfies $c(m)=O(\log m)$ form a $\mathcal{GMG}$ system. The weighted transfer operator, when acting on $C^1(\mathcal{I})$, then fulfills all the properties which we just described. In particular, in the $\mathcal{GMG}$ setting, Properties *UDE*, *SG*, $An_w(1,0)$, $An_{s,w}(1,0)$ described in Figure 1 hold. The last two Properties that are mentionned in Figure 1, namely Property *SMVL* and its "uniform" version Property *USMNVL*, are studied next.

3.2. **The *SMVL* Property for the three Euclidean systems.** We are mainly interested in the behaviour of the operator $\mathbf{H}_{s,w}$ when s varies on vertical lines $\Re s=\sigma$ and $w=\nu$ is a real fixed. The first result is the following.

Lemma 1. *Consider any Euclidean Algorithm amongst $\mathcal{G},\mathcal{K},\mathcal{O}$, and any cost of the Class $\mathcal{MG}$. When s varies on the vertical line $\Re s=\sigma, s\neq\sigma$, and $w=\nu$ is a fixed real number, the spectral radius $R(s,\nu)$ of the transfer operator $\mathbf{H}_{s,\nu}$ satisfies $R(s,\nu)<R(\sigma,\nu)$.*

Proof. Let λ be an eigenvalue of $\mathbf{H}_{s,\nu}$ and f denote an eigenfunction relative to λ. We recall that $f_{\sigma,\nu}$ denotes the dominant eigenfunction relative to $\lambda(\sigma,\nu)$ which is strictly positive on $\mathcal{I}$. Moreover, one can assume that the function $f(x)/f_{\sigma,\nu}(x)$ is of modulus at most 1 on $\mathcal{I}$ and attains modulus 1

Alg.	(G)	(K)	(O)
Interval	$[0,\, 1]$	$[0,\, 1/2]$	$[0,\, 1]$
Function $A(x)$	Integer part of x	Nearest integer to x	Odd integer nearest to x
Contraction ratio	$\phi - 1$	$\sqrt{2}-1$	$\phi - 1$
Invariant function f_1	$\frac{1}{\log 2}\frac{1}{1+x}$	$\frac{1}{\log\phi}[\frac{1}{\phi+x}+\frac{1}{\phi^2-x}]$.	$\frac{1}{\phi-1+x}+\frac{1}{\phi^2-x}$
Entropy	$\frac{\pi^2}{6\log 2}$	$\frac{\pi^2}{6\log\phi}$	$\frac{\pi^2}{9\log\phi}$

FIGURE 4. The three Euclidean dynamical systems. The symbol ϕ denotes the golden ratio $\phi = (1+\sqrt{5})/2$ and the map T is defined from function A, with the formula $T(x) := |\frac{1}{x} - A(\frac{1}{x})|$.

at point x_0. One has

$$|\lambda\, f(x_0)| = |\mathbf{H}_{s,\nu}[f](x_0)| = |\sum_{h\in\mathcal{H}} \exp[\nu c(h)]\cdot |h'(x_0)|^{-s}\; f\circ h(x_0)| \tag{3.4}$$

$$\leq \sum_{h\in\mathcal{H}} \exp[\nu c(h)]\cdot |h'(x_0)|^{-\sigma}\;\; |f\circ h(x_0)| \tag{3.5}$$

$$\leq \sum_{h\in\mathcal{H}} \exp[\nu c(h)]\cdot |h'(x_0)|^{-\sigma}\;\; f_{\sigma,\nu}\circ h(x_0) = \lambda(\sigma,\nu) f_{\sigma,\nu}(x_0), \tag{3.6}$$

and the definition of x_0 implies the inequality $|\lambda| \leq \lambda(\sigma,\nu)$.
We then suppose that, for $s = \sigma + it, t \neq 0$, $w = \nu$ real, there exists $\lambda \in \mathrm{Sp}\,\mathbf{H}_{s,\nu}$ that satisfies $|\lambda| = \lambda(\sigma,\nu)$. Then the sequence of inequalities (3.4), (3.5), (3.6) becomes a sequence of equalities. First, for any $h\in\mathcal{H}$, the equality

$$|f\circ h(x_0)| \;=\; f_{\sigma,\nu}\circ h(x_0) \tag{3.7}$$

holds. Second, the sequence $a_h := \exp[\nu c(h)]|h'(x_0)|^{-\sigma-it} f\circ h(x_0)$ 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

$$f\circ h(x_0)\;\, |h'(x_0)|^{-it} \;=\; \theta\;\, |f\circ h(x_0)| \tag{3.8}$$

holds. Then, for $m\to\infty$, the sequence $h_{[m]}(x_0) = O(m^{-1})$ tends to 0, and, equality (3.7) proves that $|f(0)| = f_{\sigma,\nu}(0) \neq 0$. Now, for $m\to\infty$, Relation (3.8) shows that the sequence

$$|h'_{[m]}(x_0)|^{-it} = (m+\epsilon x_0)^{2it}$$

has a limit, which can be only true for $t = 0$. $\square$

As we already explained, the *SMVL* Property is not sufficient for our purposes of obtaining a distributional analysis of Euclidean Algorithms, and we wish to obtain the *USMNVL* Property for the weighted transfer operators.

We shall proceed in two main steps: we first prove in Sections 3.3, 3.4, 3.5 that the *UNI* Condition entails the *USMNVL* Property in a general setting of $\mathcal{GMG}$–type. Then, we return to the Three Euclidean Dynamical Systems in Section 3.6 and we prove that they satisfy the *UNI* Condition, and thus the *USMNVL* Property.

Remark about notations. *In the sequel, the notation $A(x) \ll B(x)$ means: A is less than B up to absolute multiplicative constants. This means that there exists some absolute constant k such that for every x of interest, one has $A(x) \le kB(x)$. It is synonymous to $A(x) = O(B(x))$ with a absolute O–term. On the other hand, we denote by Σ, W real intervals where variables $\sigma = \Re s, \nu = \Re w$ may vary. The symbol $\mathcal{W}$ denotes a complex neighborhood for the (complex) variable w.*

We now return to a general triple of $\mathcal{GMG}$–type until the end of Section 3.5.

3.3. **Preparatory material.** Following Dolgopyat, we here introduce some technical tools, the normalized operator and the $(1,t)$- norm.

Introduction of the normalized operator. It will be convenient to work with the normalised operators; we let $s = \sigma + it, \ w = \nu + i\tau$, and we consider the operator $\widetilde{\mathbf{H}}_{s,w}$ defined by

$$\widetilde{\mathbf{H}}_{s,w}[f] = \frac{1}{\lambda(\sigma,\nu) f_{\sigma,\nu}} \, \mathbf{H}_{s,w}[f_{\sigma,\nu} \cdot f] \,, \tag{3.9}$$

For $(\sigma,\nu) \in \Sigma_1 \times W_1$, the operator $\widetilde{\mathbf{H}}_{\sigma,\nu}$ acting on $C^1(\mathcal{I})$ has spectral radius 1, and fixes the constant function $\equiv 1$. Remark the inequality $||\widetilde{\mathbf{H}}_{s,w}[f]||_0 \le ||f||_0 \, \widetilde{\mathbf{H}}_{\sigma,\nu}[1] = ||f||_0$ that proves that the useful bound

$$||\widetilde{\mathbf{H}}_{s,w}||_0 \le 1. \tag{3.10}$$

On the other hand, it is easy to check that $\widetilde{\mathbf{H}}^*_{\sigma,\nu}$ preserves the probability measure $\mu_{\sigma,\nu} = f_{\sigma,\nu} \cdot \widehat{\mu}_{\sigma,\nu}$.

For $(\sigma,\nu) \in \Sigma_1 \times W_1$, the operator $\widetilde{\mathbf{H}}^n_{\sigma,\nu}$, and the probability measure $\mu_{\sigma,\nu}$ can be compared to their analogues for $(\sigma,\nu) = (1,0)$. This is the purpose of the following result. This result which is (only) needed when the number of branches is infinite is a refinement of results that are implicit in works of Dolgopyat and Pollicott.

Lemma 2. *For $(\sigma,\nu) \in \Sigma_1 \times W_1$, denote by $A_{\sigma,\nu}$ the constant*

$$A_{\sigma,\nu} := \frac{\lambda(2\sigma - 1, 2\nu)^{1/2}}{\lambda(\sigma,\nu)}.$$

Let $\mathcal{L}$ be a compact of $\Sigma_1 \times W_1$. For $\mathcal{J} \subset \mathcal{H}^k$, denote by J the subset $J = \cup_{h \in \mathcal{J}} h(\mathcal{I})$. Then, for $(\sigma,\nu) \in \mathcal{L}$,

$$\mu_{\sigma,\nu}[J] \ll A^k_{\sigma,\nu} \ |J|^{1/2}.$$

For $(\sigma,\nu) \in \mathcal{L}$, for any $f \in \mathcal{C}^1(\mathcal{I})$, for any integer $k \ge 1$,

$$||\widetilde{\mathbf{H}}^k_{\sigma,\nu}[f]||_0^2 \ll A^{2k}_{\sigma,\nu} \ ||\widetilde{\mathbf{H}}^k_{1,0}[|f|^2]||_0.$$

The absolute constants involved only depend on $\mathcal{L}$.

Remark. The constant $A_{\sigma,\nu}$ depends continuously on (σ,ν) and equals 1 for $(\sigma,\nu)=(1,0)$. Then, it tends to 1 when (σ,ν) tends to $(1,0)$.

Proof. The equality $\mu_{\sigma,\nu}[f]=\mu_{\sigma,\nu}[\widetilde{\mathbf{H}}^k_{\sigma,\nu}[f]]$, when applied to the characteristic function of some fundamental interval $h(\mathcal{I})$ of depth k proves that

$$\mu_{\sigma,\nu}[h(\mathcal{I})] \ll \frac{\exp[\nu c(h)]}{\lambda(\sigma,\nu)^k}\int_{\mathcal{I}} |h'(x)|^\sigma d\mu_{\sigma,\nu}(x).$$

Moreover, with the bounded distortion property (3.2), the ratio (two by two) of the three quantities $a(h), b(h), c(h)$,

$$a(h) := \int_{\mathcal{I}} |h'(x)|^\sigma d\mu_{\sigma,\nu}(x); \qquad b(h) := |h(\mathcal{I})|^\sigma; \qquad c(h) := \int_{\mathcal{I}} |h'(x)|^\sigma dx$$

admit some upper and lower bounds that do not depend on h, and are uniform when (σ,ν) varies on compact $\mathcal{L}$. Then, summing the inequalities

$$\mu_{\sigma,\nu}[h(\mathcal{I})] \ll \frac{\exp[\nu c(h)]}{\lambda(\sigma,\nu)^k}|h(\mathcal{I})|^\sigma$$

over some subset $\mathcal{J}\subset\mathcal{H}^k$, and applying the Cauchy–Schwarz inequality, one obtains, for $J := \cup_{h\in\mathcal{J}} h(\mathcal{I})$,

$$\mu_{\sigma,\nu}[J] \le \frac{1}{\lambda(\sigma,\nu)^k}\left(\sum_{h\in\mathcal{H}^k}\exp[2\nu c(h)]\cdot|h(\mathcal{I})|^{2\sigma-1}\right)^{1/2}\left(\sum_{h\in\mathcal{J}}|h(\mathcal{I})|\right)^{1/2}.$$

Then, dominant spectral properties, together with bounded distortion, entails the inequality

$$\sum_{h\in\mathcal{H}^k}\exp[2\nu c(h)]\cdot|h(\mathcal{I})|^{2\sigma-1} \ll \lambda(2\sigma-1,2\nu)^k$$

and, finally, the relation $|J|=\sum_{h\in\mathcal{J}}|h(\mathcal{I})|$ provides the final first result.

Consider now some function $f\in C^1(\mathcal{I})$. The relation

$$|\widetilde{\mathbf{H}}_{\sigma,\nu})^k[f](x)| \ll \frac{1}{\lambda(\sigma,\nu)^k}\sum_{h\in\mathcal{H}^k}\exp[\nu c(h)]\cdot|h'(x)|^\sigma\cdot|f\circ h(x)|$$

is valid if (σ,ν) belongs to compact $\mathcal{L}$, and, with Cauchy-Schwarz inequality,

$$\left(\sum_{h\in\mathcal{H}^k}\exp[\nu c(h)]\cdot|h'(x)|^\sigma\cdot|f\circ h(x)|\right)^2$$

$$\le \left(\sum_{h\in\mathcal{H}^k}\exp[2\nu c(h)]\cdot|h'(x)|^{2\sigma-1}\right)\cdot\left(\sum_{h\in\mathcal{H}^k}|h'(x)|\cdot|f|^2\circ h(x)\right).$$

The second factor is exactly $\mathbf{H}^k_{1,0}[|f|^2](x)$, that is less than $\widetilde{\mathbf{H}}^k_{1,0}[|f|^2](x)$ (up to absolute multiplicative constants). Thanks to dominant spectral properties, the first factor is easily related to $\lambda(2\sigma-1,2\nu)^k$. □

The following Lemma describes how the operator $\widetilde{\mathbf{H}}_{s,w}$ deals with the quasi–norm $||.||_1$:

Lemma 3. [Lasota-Yorke]. *For every compact $\mathcal{L}$ of $\Sigma_1 \times W_1$, there is a constant $C > 0$, so that for all $n \geq 1$, all (s, w) for which $(\Re s, \Re w) \in \mathcal{L}$ and all $f \in C^1(\mathcal{I})$:*

$$||\widetilde{\mathbf{H}}^n_{s,w} f||_1 \leq C\left(|s|\,||f||_0 + \rho^n\,||f||_1\right) .$$

Proof. The quantity $\widetilde{\mathbf{H}}^n_{s,w}[f]$ can be written as a sum

$$\widetilde{\mathbf{H}}^n_{s,w}[f] = \frac{\exp[wc(h)]}{\lambda(\sigma,\nu)^n} r_h(x) \qquad \text{with} \qquad r_h := |h'|^s \cdot \frac{1}{f_{\sigma,\nu}} \cdot (f_{\sigma,\nu} f) \circ h.$$

The Leibniz sum for the derivative of each term r_h contains three terms. We can bound the first for all s using the distortion assumption since

$$|s||h''|||h'|^{s-1}| \leq |s|\overline{K}||h'|^s| = |s|K|h'|^\sigma .$$

Compactness of $\mathcal{L}$ and continuity of $(\sigma, \nu) \mapsto f_{\sigma,\nu}, (\sigma, \nu) \mapsto f'_{\sigma,\nu}$ imply that the second term may be controlled by

$$\frac{|f'_{\sigma,w}|}{f^2_{\sigma,w}} \leq C_{\mathcal{L}} \frac{1}{f_{\sigma,w}} ,$$

for some $C_{\mathcal{L}} > 0$. Finally the last term can be estimated using

$$|(f_{\sigma,w} \cdot f)' \circ h||h'| \leq \rho^n[|f'_{\sigma,w} \cdot f| \circ h + (f_{\sigma,w} \cdot |f'|) \circ h] .$$

We can ensure

$$\overline{K}|s| + C_{\mathcal{L}} + \rho^n C_{\mathcal{L}} \leq C(\mathcal{L})|s|$$

so that the derivative $(\widetilde{\mathbf{H}}^n_{s,w}[f])'$ satisfies

$$||(\widetilde{\mathbf{H}}^n_{s,w}[f])'||_0 \leq C\left(|s|\,||\widetilde{\mathbf{H}}^n_{\sigma,\nu}[f]||_0 + \rho^n\,||\widetilde{\mathbf{H}}^n_{\sigma,\nu}[|f'|]||_0\right) .$$

The final result is obtained with (3.10).

□

Introduction of the $(1,t)$-norm. In the derivative of $x \mapsto \widetilde{\mathbf{H}}^n_{s,w} f(x)$, there appear two terms, one which contains a factor $|s|$, and the other one which contains a decreasing exponential in n. In order to suppress the effect of the factor $|s|$, Dolgopyat proposes to use the family of norms

$$\|f\|_{1,t} := \|f\|_0 + \frac{1}{|t|}\|f\|_1 = \sup|f| + \frac{1}{|t|}\sup|f'| , \quad t \neq 0 .$$

With this norm, Lemma 3, together with relation (3.10), we obtain the first (easy) result:

Lemma 4. *For any $t_1 > 0$, for every compact $\mathcal{L}$ of $\Sigma_1 \times W_1$, there is a constant $M_0 > 0$, so that for all $n \geq 1$, all (s, w) for which $(\Re s, \Re w) \in \mathcal{L}$ and $s = \sigma + it$ with $|t| \geq t_1$, one has*

$$||\widetilde{\mathbf{H}}^n_{s,w}||_{1,t} \leq M_0.$$

3.4. ***UNI* Condition and L^2–estimates.** We are interested in refinements of both Lemmas 1 and 4, and we ask the following question about the validity of Property *USMNVL*:

Consider a (real) neighborhood $\Sigma \times W$ of $(1, 0)$. Does there exist some constant $\gamma < 1$ for which the spectral radius of $\mathbf{H}_{s,w}$ satisfies $R(s, w) < \gamma R(\Re s, \Re w)$ when $(\Re s, \Re w)$ belongs to $\Sigma \times W$ and $|\Im s|$ sufficiently large ?

We thus search some bounds on $||\widetilde{\mathbf{H}}^n_{s,w}||_{1,t}$ that are decreasing exponential in n. In fact, the main idea of Dolgopyat is to deal first with the L^2–norm of some iterate $\widetilde{\mathbf{H}}^n_{s,w}[f]$, with an index n that depends on $t = \Im s$. Then, it is possible to operate two transfers: a first transfer from this L^2–bound into a bound for the sup-norm $||.||_0$, then a second transfer from this sup-norm–bound into a bound for the $(1, t)$-norm.

Dolgopyat first proves that

$$\int_{\mathcal{I}} |\widetilde{\mathbf{H}}^{n_0}_s[f](x)|^2 dx \le \gamma^{n_0} ||f||_{1,t},$$

for some $n_0 = O(\log |t|)$ and some absolute constant $\gamma < 1$, provided that the dynamical system satisfies the so-called *UNI* Condition. In this subsection, we extend this result to the weighted operator $\mathbf{H}_{s,w}$, in the case when the number of branches of $(\mathcal{I}, T)$ is possibly infinite.

Writing $s = \sigma + it$, $w = \nu + i\tau$, the quantity $S_n(x) = |\widetilde{\mathbf{H}}^n_{s,w}[f](x)|^2$ can be expressed as a sum over all pairs $(h, k) \in \mathcal{H}^n \times \mathcal{H}^n$,

$$S_n(x) = \frac{1}{\lambda(\sigma,\nu)^{2n}} \sum_{(h,k)\in\mathcal{H}^n\times\mathcal{H}^n} \exp[wc(h) + \overline{w}c(k)] \cdot \exp[it\Psi_{h,k}(x)] \cdot R_{h,k}(x)$$

with

$$\Psi_{h,k}(x) := \log \frac{|h'(x)|}{|k'(x)|}, \tag{3.11}$$

$$R_{h,k}(x) = |h'(x)|^\sigma |k'(x)|^\sigma \frac{1}{f^2_{\sigma,\nu}(x)} (f.f_{\sigma,\nu}) \circ h(x) (\overline{f}.f_{\sigma,\nu}) \circ k(x). \tag{3.12}$$

Using the decomposition $f = \Re f + i\Im f$, the term $R_{h,k}$ decomposes into four terms, each of which has the form

$$r_{h,k}(x) = \xi |h'(x)|^\sigma |k'(x)|^\sigma \frac{1}{f^2_{\sigma,w}(x)} (g \cdot f_{\sigma,w}) \circ h(x) (\ell \cdot f_{\sigma,w}) \circ k(x). \tag{3.13}$$

for two real functions $g, \ell \in \{\Re f, \Im f\}$ and $\xi \in \{\pm 1, \pm i\}$.

The functions $\Psi_{h,k}$ play an important rôle here, and the sum $S_n(x)$ will be split into two parts, $S^-_n(x)$ and $S^+_n(x)$ according to the properties of the function $\Psi_{h,k}$. The sum S^+_n gathers all the pairs (h, k) for which the derivative $|\Psi'_{h,k}(x)|$ has a lower bound sufficiently large. In this case, the Van Der Corput Lemma can be applied (Lemma 6), and gives the convenient bound for the corresponding integral I^+_n. The sum S^-_n gathers the other pairs, for which the derivative $|\Psi'_{h,k}(x)|$ has a small lower bound, and the condition *UNI* entails that there are not too many pairs (h, k) for which the infimum of $|\Psi'_{h,k}(x)|$ is too small. It provides the convenient bound for the corresponding integral I_n^- (Lemma 5).

For two inverse branches h et k of same depth, we consider

$$\Delta(h,k) = \inf_{x\in\mathcal{I}} |\Psi'_{h,k}(x)| = \inf_{x\in\mathcal{I}} \left|\frac{h''}{h'}(x) - \frac{k''}{k'}(x)\right|. \tag{3.14}$$

For any fixed h in $\mathcal{H}^n$, we denote by $J(h,\eta)$ the union of all the fundamental intervals $k(\mathcal{I})$ relative to inverse branches k of $\mathcal{H}^n$ for which $\Delta(h,k) \le \eta$,

$$J(h,\eta) := \bigcup_{\substack{k\in\mathcal{H}^n\\ \Delta(h,k)\le\eta}} k(\mathcal{I}) \qquad (h \in \mathcal{H}^n). \tag{3.15}$$

Then $J(h,\eta)$ is analogous to a ball of center h and radius η with respect to "distance" Δ. In his paper [15], Dolgopyat states various Conditions *UNI*, and he expressed them in terms of flows. Here, we use the strongest of these conditions and we adapt it it in our context. Informally, Property *UNI* expresses that there are not too many pairs (h,k) that are too close (with respect to Δ). More precisely, the Lebesgue measure of $J(h,\delta)$ can be studied precisely when δ is scaled similarly to the maximal length of fundamental intervals of depth n. This length is ρ^n (up to absolute constants) and plays the role of a reference scale.

***UNI* Condition.** *A dynamical system of the Good Class, with a contraction ratio ρ, fulfills the UNI Condition if the functions $\Psi_{h,k}$ satisfy the following:* *(a) For any a $(0 < a \le 1)$, for any n, for any $h \in \mathcal{H}^n$, the Lebesgue measure of the ball $J(h,\rho^{an})$ relative to distance Δ satisfies*

$$|J(h,\rho^{an})| \ll \rho^{an}.$$

(b) There exists $Q < \infty$ for which

$$\sup\{|\Psi''_{h,k}(x)|; (h,k) \in \mathcal{H}^\star \times \mathcal{H}^\star, x \in \mathcal{I}\} = Q.$$

Study of the L^2–norm: the close pairs. In this case, the main result is the following:

Lemma 5. *Suppose that Condition UNI(a) holds. For any compact $\mathcal{L}$ of $\Sigma_1 \times W_1$, for all (s,w) for which $(\sigma,\nu) \in \mathcal{L}$, for all n, for all $a, 0 < a \le 1$, the integral I_n^- of the sum S_n^- relative to pairs $(h,k) \in \mathcal{H}^n \times \mathcal{H}^n$ for which $\Delta(h,k)$ is less than ρ^{an} satisfies*

$$|I_n^-| = |I_n^-(s,w,f,a)| \ll \left(\rho^{a/2}A_{\sigma,\nu}\right)^n ||f||_0^2, \quad \textit{with} \;\; A_{\sigma,\nu} := \frac{\lambda(2\sigma-1,2\nu)^{1/2}}{\lambda(\sigma,\nu)}$$

Proof. Up to a positive constant that only depends on (σ,ν) and is uniform on any compact $\mathcal{L}$, $|I_n^-|$ is less than

$$\frac{||f||_0^2}{\lambda(\sigma,\nu)^{2n}} \sum_{\substack{(h,k)\in\mathcal{H}^n\times\mathcal{H}^n\\ \Delta(h,k)\le\rho^{an}}} \exp[\nu(c(h)+c(k))] \cdot \int_{\mathcal{I}} |h'(x)|^\sigma |k'(x)|^\sigma dx.$$

First, using bounded distortion property, for all pairs $(h,k) \in \mathcal{H}^\star \times \mathcal{H}^\star$, up to multiplicative absolute constants, one has

$$\int_{\mathcal{I}} |h'(x)|^\sigma |k'(x)|^\sigma dx \ll \left(\int_{\mathcal{I}} |h'(x)|^\sigma dx \right) \cdot \left(\int_{\mathcal{I}} |k'(x)|^\sigma dx \right)$$

Moreover, as in the beginning of Lemma 2, with the same bounded distortion property and the definition of $\mu_{\sigma,\nu}$ as an invariant probability for the normalized operator, the ratio (two by two) of the four quantities $a(h), b(h), c(h), d(h)$,

$$a(h) := \frac{\exp[\nu c(h)]}{\lambda(\sigma,\nu)^n} \int_{\mathcal{I}} |h'(x)|^\sigma dx; \qquad b(h) := \frac{\exp[\nu c(h)]}{\lambda(\sigma,\nu)^n} \int_{\mathcal{I}} |h'(x)|^\sigma d\mu_{\sigma,\nu}(x);$$

$$c(h) := \mu_{\sigma,\nu}[h(\mathcal{I})]; \qquad d(h) := \frac{\exp[\nu c(h)]}{\lambda(\sigma,\nu)^n} |h(\mathcal{I})|^\sigma$$

admit some upper bounds and lower bounds that do not depend on h, and are uniform when (σ,ν) varies on any compact. Up to a multiplicative constant, it is then sufficient to study the sum

$$\sum_{h \in \mathcal{H}^n} \mu_{\sigma,\nu}[h(\mathcal{I})] \left(\sum_{\substack{k \in \mathcal{H}^n \\ \Delta(h,k) \le \rho^{an}}} \mu_{\sigma,\nu}[k(\mathcal{I})] \right) = \sum_{h \in \mathcal{H}^n} \mu_{\sigma,\nu}[h(\mathcal{I})]\ \mu_{\sigma,\nu}[J(h,\rho^{an})].$$

Now, the first relation of Lemma 2, $\mu_{\sigma,\nu}[J] \le C_{\sigma,\nu} A^n_{\sigma,\nu}\ |J|^{1/2}$, that holds for any subset J that is a union of fundamental intervals of depth n, is applied to subset $J(h,\rho^{an})$. *UNI* provides an evaluation of its Lebesgue measure, and finally

$$|I_n^-| \ll \left(\rho^{a/2} A_{\sigma,\nu}\right)^n ||f||_0^2 .$$

□

Study of the L^2–norm: Application of the Van der Corput Lemma. Consider now the integral I_n^+ of the sum S_n^+ relative to pairs (h,k) for which $\Delta(h,k)$ is more than ρ^{an} for some $a, 0 < a < 1$.

Lemma 6. *Suppose that Condition UNI(b) holds. Denote by n_0 the integer defined as*

$$n_0 = \left\lceil \frac{1}{|\log \rho|} \log |t| \right\rceil . \tag{3.16}$$

Then, for any compact $\mathcal{L}$ of $\Sigma_1 \times W_1$, for any (s,w) for which $(\sigma,\nu) \in \mathcal{L}$ and $|t| \ge 1/\rho^2$, for any $a, 0 < a < 1/2$, the integral $I_{n_0}^+$ of the sum $S_{n_0}^+(x)$ relative to pairs $(h,k) \in \mathcal{H}^{n_0} \times \mathcal{H}^{n_0}$ for which $\Delta(h,k)$ is more than ρ^{an_0} satisfies

$$|I_{n_0}^+| = |I_{n_0}^+(s,w,f,a)| \ll \rho^{(1-2a)n_0} ||f||_{1,t}^2 .$$

Proof. In this case, we bound $|I_n^+|$ as

$$|I_n^+| \le \frac{1}{\lambda(\sigma,\nu)^{2n}} \sum_{\substack{(h,k) \in \mathcal{H}^n \times \mathcal{H}^n \\ \Delta(h,k) \ge \rho^{an}}} \exp[\nu(c(h)+c(k))] \cdot \left| \widehat{I}(h,k) \right| \tag{3.17}$$

where the integral

$$\widehat{I}(h,k) := \int_{\mathcal{I}} \exp[it\Psi_{h,k}(x)]\, R_{h,k}(x)dx$$

involves functions $\Psi_{h,k}, R_{h,k}$ defined in (3.11) (3.12) and decomposes in four integrals of the form $I(h,k)$

$$I(h,k) := \int_{\mathcal{I}} \exp[it\Psi_{h,k}(x)]\, r_{h,k}(x)dx$$

with $r_{h,k}$ defined in (3.13). We apply the Van Der Corput Lemma to each integral $I(h,k)$.

Van der Corput Lemma (cf [43]). *For some interval $\mathcal{I}$, consider the integral*

$$I = \int_{\mathcal{I}} \exp[it\Psi(x)]\, r(x)dx.$$

Suppose that Ψ satisfies the following

$$\Psi \in \mathcal{C}^2(\mathcal{I}), \quad |\Psi''(x)| \le Q, \quad |\Psi'(x)| \ge \Delta, \quad \textit{with} \quad (1/|t|) \le \Delta \le 1,$$

and there exist (ϵ, D) for which $||r||_0 \le \epsilon$, and $||r||_{1,1} \le \epsilon D$. Then

$$|I| \le \epsilon\, C(Q) \left[\frac{D+1}{|t|\Delta} + \frac{1}{|t|\Delta^2}\right].$$

Consider some pair (t,n) with $1/|t| \le \rho^{an}$ for some real $a, 0 < a < 1/2$. With

$$M(h,k) := \sup_{x\in\mathcal{I}} |h'(x)|^\sigma |k'(x)|^\sigma \frac{1}{f^2_{\sigma,\nu}(x)} f_{\sigma,\nu}\circ h(x) f_{\sigma,\nu}\circ k(x), \tag{3.18}$$

the norm $||r_{h,k}||_0$ satisfies

$$||r_{h,k}||_0 \le M(h,k)||g||_0\, ||\ell||_0 \le M(h,k)||g||_{1,t}\, ||\ell||_{1,t}.$$

Principles used in the proof of Lemma 3 for the function r_h apply to function $r_{h,k}$, and the norm $||r_{h,k}||_{1,1}$ satisfies

$$\begin{aligned}||r_{h,k}||_{1,1} &\ll M(h,k)\, [||g||_0\, (||\ell||_0 + \rho^n\, ||\ell||_1) + ||\ell||_0\, (||g||_0 + \rho^n\, ||g||_1)] \\ &\ll M(h,k)||g||_{1,t}||\ell||_{1,t}\, [1 + \rho^n |t|]\,.\end{aligned}$$

Then, the Van Der Corput Lemma can be applied for evaluating each integral $I(h,k)$ that thus satisfies, with the Van Der Corput Lemma, and Property *UNI(b)*,

$$|I(h,k)| \ll M(h,k)\; ||g||_{1,t}||\ell||_{1,t}\; \left[\frac{2+|t|\rho^n}{|t|\rho^{an}} + \frac{1}{|t|\rho^{2an}}\right].$$

Now, we choose as $n = n_0$ the smallest integer n for which $|t|\rho^{n_0} < 1$, so that n_0 is defined as in (3.16). If, moreover $a < 1/2$, and $\sqrt{|t|} \ge 1/\rho$, then $n_0 \ge 2$ and condition $\rho^{-an_0} \le \rho^{-(n_0-1)} \le |t|$ holds, so that Van Der Corput Lemma can be applied for this value n_0, and

$$I(h,k) \ll M(h,k)\; ||g||_{1,t}||\ell||_{1,t}\; \rho^{(1-2a)n_0}. \tag{3.19}$$

Returning to integral $\widehat{I}(h,k)$,

$$|\widehat{I}(h,k)| \ll M(h,k)\; ||f||^2_{1,t}.$$

Now, take some x_0 in $\mathcal{I}$. Then, from bounded distortion property (3.2), and definition of $M(h,k)$ in (3.18), the inequality

$$M(h,k) \ll |h'(x_0)|^\sigma |k'(x_0)|^\sigma \frac{1}{f^2_{\sigma,\nu}(x_0)} f_{\sigma,\nu} \circ h(x_0) f_{\sigma,\nu} \circ k(x_0)$$

holds and entails that

(3.20)

$$\frac{1}{\lambda(\sigma,\nu)^{2n}} \sum_{(h,k)\in\mathcal{H}^n\times\mathcal{H}^n} \exp[\nu(c(h)+c(k))]\, M(h,k) \ll \left(\widetilde{\mathbf{H}}^n_{\sigma,\nu}[1](x_0)\right)^2 = 1.$$

From (3.19, 3.20, 3.17), we obtain

$$|I^+_{n_0}| \ll \rho^{(1-2a)n_0}\, ||f||^2_{1,t}\,.$$

□

Study of the L^2–norm: the final result. Finally, consider the integer n_0 defined in (3.16) of Lemma 6 and some a with $(2/5) < a < (1/2)$. Then, since $a/2 > (1-2a) > 0$, there exists a (real) neighborhood of $(\sigma,\nu) = (1,0)$ for which $A_{\sigma,\nu}\rho^{a/2}$ is less than ρ^{1-2a}. Then, from both Lemmas 5 and 6, the L^2 norm of the n_0–th iterate, that equals to $I^+_{n_0} + I^-_{n_0}$ satisfies

$$\int_{\mathcal{I}} |\widetilde{\mathbf{H}}^{n_0}_{s,w}[f](x)|^2 dx \ll \rho^{(1-2a)n_0} ||f||^2_{1,t}\,. \tag{3.21}$$

3.5. **Transfers between various norms.** We now operate two successive transfers, first to the sup-norm, then to the $(1,t)$ -norm.

From the L^2-norm to the sup-norm. For comparing the two norms, the L^2 norm and the sup–norm, we use two main ingredients:

(i) the two operators, the transfer operator $\widetilde{\mathbf{H}}^n_{\sigma,\nu}$ and the reference operator $\widetilde{\mathbf{H}}^n_{1,0} = \widetilde{\mathbf{H}}^n_1$, are not too different when (σ,ν) is near $(1,0)$. We use the second inequality of Lemma 2,

$$||\widetilde{\mathbf{H}}^m_{\sigma,\nu}[g]||^2_0 \ll A^{2m}_{\sigma,\nu}\, ||\widetilde{\mathbf{H}}^m_{1,0}[|g|^2]||_0. \tag{3.22}$$

(ii) Since the normalized density transformer $\widetilde{\mathbf{H}}_1$ is quasi–compact with respect to the $(1,1)$–norm, and fixes the function 1, it satisfies

$$||\widetilde{\mathbf{H}}^m_1[|g|^2]||_0 \le \left(\int_{\mathcal{I}} |g|^2(t)dt\right) + O(r^m)||g^2||_{1,1}, \tag{3.23}$$

where r is the subdominant spectral radius of $\mathbf{H}_1$.

Consider an iterate $\widetilde{\mathbf{H}}^n_{s,w}$ of $\widetilde{\mathbf{H}}_{s,w}$. Then

$$||\widetilde{\mathbf{H}}^n_{s,w}[f]||^2_0 \ll ||\widetilde{\mathbf{H}}^{n-n_0}_{\sigma,\nu}[g]||^2_0 \qquad \text{with} \qquad g = |\widetilde{\mathbf{H}}^{n_0}_{s,w}[f]|.$$

Now, using (3.22) and (3.23) with $m := n - n_0$, together with the Bound (3.21) for the L^2–norm and finally Lemma 3 for evaluating $||g^2||_{1,1}$, one obtains

$$||\widetilde{\mathbf{H}}^n_{s,w}[f]||^2_0 \le A^{2(n-n_0)}_{\sigma,w}\left[\rho^{(1-2a)n_0} + r^{n-n_0}\,|t|\right] ||f||^2_{1,t}.$$

We now choose $n = n_1$ so that the two terms $\rho^{(1-2a)n_0}$ and r^{n-n_0} are almost equal,

$$n_1 = (1+\eta)n_0 \qquad \text{with} \qquad \eta := 2(1-a)\frac{\log \rho}{\log r} > 0, \tag{3.24}$$

then we consider a (real) neighborhood $\Sigma \times W$ of $(1,0)$ for which

$$\sup\left[\lambda(\sigma,\nu)^{1+\eta}, A^{\eta}_{\sigma,\nu}\right] < \rho^{-\epsilon} \qquad \text{with} \qquad 0 < \epsilon < \frac{1}{2} - a, \tag{3.25}$$

One finally obtains, for n_1, η defined in (3.24)

$$||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_0 \ll \rho^{n_1 b}\, ||f||_{1,t}, \qquad \text{with} \qquad b := \frac{1-2a-2\epsilon}{1+\eta}. \tag{3.26}$$

From the sup-norm to the $||.||_{1,t}$–norm. Consider now the index $n_2 = 2n_1$. Applying two times Lemma 3 and using the previous bound (3.26) yield the inequality

$$||\widetilde{\mathbf{H}}^{2n_1}_{s,w}[f]||_1 \ll |s|\, ||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_0 + \rho^{n_1}\, ||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_1$$

$$\ll |s|\, \rho^{n_1 b}||f||_{1,t} + \rho^{n_1}|t| \left(\frac{|s|}{|t|}||f||_0 + \rho^{n_1}\frac{||f||_1}{|t|}\right) \ll |t|\rho^{n_1 b}||f||_{1,t}$$

that finally yields the bound

$$||\widetilde{\mathbf{H}}^{n_2}_{s,w}||_{1,t} \ll \rho^{n_2 b/2}. \tag{3.27}$$

The last step. Since any n can be written as $n = kn_2 + \ell$ with $\ell < n_2$, the previous bound (3.27) and the inequality of Lemma 4, entail

$$||\widetilde{\mathbf{H}}^{n}_{s,w}||_{1,t} \le M_0\, ||\widetilde{\mathbf{H}}^{n_2}_{s,w}||^k_{1,t} \le M_0\, \rho^{bkn_2/2} \le M_0\, \rho^{bn/2}\, \rho^{-bn_2/2}.$$

Since $bn_2/2 = bn_1 = (1-2a-2\epsilon)n_0$, with n_0 defined in (3.16), we finally obtain

$$||\widetilde{\mathbf{H}}^{n}_{s,w}||_{1,t} \le M\, |t|^d\, \gamma^n$$

with

$$\gamma := \rho^{b/2}, \qquad \frac{b}{2} := \frac{1-2a-2\epsilon}{2(1+\eta)}, \qquad d := 1-2a-2\epsilon, \qquad M = \frac{M_0}{\rho^d}.$$

We consider now an arbitrarly small $\epsilon > 0$, and we relate now a with ϵ via the equality $a = 1/2 - 2\epsilon$. Then d, b and η depend only on ϵ via the relations

$$d = 2\epsilon, \qquad \eta = 2(1+2\epsilon)\frac{\log \rho}{\log r}, \qquad b/2 = \frac{\epsilon}{1+\eta}$$

Now, we return to the operator $\mathbf{H}_{s,w}$ itself, and we obtain the first Inequality of Theorem 2. Finally, since $\Sigma \times W$ is chosen as in (3.25), one has, for any $(\sigma,\nu) \in \Sigma \times W$,

$$\gamma\lambda(\sigma,\nu) \le \rho^{\frac{\epsilon}{1+\eta}} \cdot \rho^{-\frac{\epsilon}{2(1+\eta)}} = \rho^{\frac{\epsilon}{2(1+\eta)}} = \widehat{\gamma} < 1.$$

This proves the second Inequality of Theorem 2 with $M_1 := M/(1-\widehat{\gamma})$. Then Theorem 2 is proved.

3.6. ***UNI* Condition and Euclidean dynamical systems.** In order to apply the result of Theorem 2 to the three Euclidean Algorithms, we now prove that any of these three Euclidean Systems satisfies the *UNI* Condition.

For two LFT's h_1 and h_2, with $h_i(x) = (a_i x + b_i)/(c_i x + d_i)$, the derivative $\Psi'_{h_1,h_2}(x)$ equals

$$\left|\frac{h_1''}{h_1'}(x) - \frac{h_2''}{h_2'}(x)\right| = \left|\frac{c_1}{c_1 x + d_1} - \frac{c_2}{c_2 x + d_2}\right| = \frac{|c_1 d_2 - c_2 d_1|}{|(c_1 x + d_1)(c_2 x + d_2)|}.$$

Then, the "distance" $\Delta(h_1, h_2)$ between two LFT's only involves the denominator coefficients c_i, d_i, via their quotients c_i/d_i and satisfies, with the distortion Property

$$\Delta(h_1, h_2) := \left|\frac{c_1}{d_1} - \frac{c_2}{d_2}\right| \cdot \inf_{x \in \mathcal{I}} \left|\frac{h_1'(x) h_2'(x)}{h_1'(0) h_2'(0)}\right|^{1/2} \geq \frac{1}{L}\left|\frac{c_1}{d_1} - \frac{c_2}{d_2}\right|. \tag{3.28}$$

We thus relate the distance Δ between two LFT's h and k with the (true) distance between rationals $\widehat{h}(0)$ and $\widehat{k}(0)$ which deals with the so-called mirror LFT $\widehat{h}$ of h which we now define. This mirror operation is used in [42] by Schweiger who relates it to the natural extension and in [6] where the authors use the geometric notion of "folded" and unfolded".

The mirror of h is the LFT whose decomposition involves the same "digits" as h but in the inverse order. The mirror map is an involution that is first defined on $\mathcal{H}$. If h is of depth one, of the form $h(x) := \epsilon'/(m + \epsilon x)$, the mirror LFT $\widehat{h}$ is of the form $\widehat{h}(x) := \epsilon/(m + \epsilon' x)$. The mirror mapping is then extended to $\mathcal{H}^\star$ with the morphism property $\widehat{h \circ k} = \widehat{k} \circ \widehat{h}$. Then, if $h \in \mathcal{H}^p$ deals with the sequence $\epsilon_0, (m_1, \epsilon_1), (m_2, \epsilon_2) \ldots (m_p, \epsilon_p)$, the mirror LFT $\widehat{h}$ deals with the sequence $\epsilon_p, (m_p, \epsilon_{p-1}) \ldots (m_{p-1}, \epsilon_{p-2}) \ldots (m_1, \epsilon_0)$. Furthermore, classical properties of continuants entail that the LFT $\widehat{h}$ satisfies

$$\widehat{h}(x) = \frac{ax + c}{bx + d} \qquad \text{if} \qquad h(x) = \frac{ax + b}{cx + d},$$

so that Relation (3.28) proves that the distance $\Delta(h_1, h_2)$ between h_1 and h_2 satisfies

$$\Delta(h_1, h_2) \geq \frac{1}{L}\,|\widehat{h}_1(0) - \widehat{h}_2(0)|. \tag{3.29}$$

Then, $\Delta(h_1, h_2)$ is closely related to the true distance between $\widehat{h}_1(0)$ and $\widehat{h}_2(0)$.

We now prove that, for each of the three Euclidean Algorithms to be analysed, the set $\{\widehat{h}, h \in \mathcal{H}\}$ is the set of inverse branches of some dynamical system $(\widehat{\mathcal{I}}, \widehat{T})$ that belongs to the Good Class. This dynamical system, denoted by $\widehat{\mathcal{E}}$, is called the dual dynamical system of $\mathcal{E}$. In the case of the Classical Euclidean algorithm $\mathcal{G}$, since all the ϵ are equal to 1, these two dynamical systems $(\widehat{\mathcal{I}}, \widehat{T})$ and $(\mathcal{I}, T)$ are the same. In all the three cases, the dual dynamical systems are described and drawn in Figure 5. For each of the three algorithms, the dual algorithm $\widehat{\mathcal{E}}$ belongs to the Good Class. Moreover, the algorithms $\mathcal{E}$ and $\widehat{\mathcal{E}}$ have the same contraction ratio ρ. This

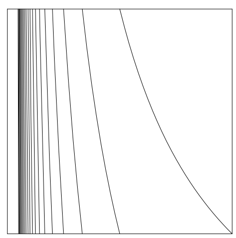
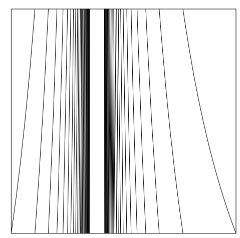
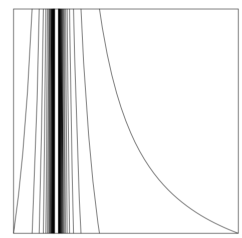
Alg.	$\widehat{G}$	$\widehat{K}$	$\widehat{O}$
Interval	$[0,1]$	$[\phi-2, \phi-1]$	$[\phi-2, \phi]$
Generic set	$\widehat{\mathcal{G}} = \{\frac{1}{m+x},\ m \geq 1\}$	$\widehat{\mathcal{K}} = \{\frac{\epsilon}{m+x}, \epsilon = \pm 1,$ $(m,\epsilon) \geq (2,+1)\}$	$\widehat{\mathcal{O}} = \{\frac{\epsilon}{m+x},\ \epsilon = \pm 1,$ m odd,$(m,\epsilon) \geq (1,1)\}$
Graphs			

FIGURE 5. The three dual Euclidean dynamical systems: Standard, Centered, Odd.

is due to the definition of ρ

$$\rho = \left(\max\{|h'(x)|; h \in \mathcal{H}^n, x \in \mathcal{I}\}\right)^{1/n}$$

together with distortion Property (for $\mathcal{E}$) and the fact that, in each system $\mathcal{E}$ and $\widehat{\mathcal{E}}$, the "worst branch" (for which the contraction ratio is attained) is a self-mirrored h (i.e., a LFT h for which $h = \widehat{h}$). Denote by $\widehat{L}$ the distortion constant of $\widehat{\mathcal{E}}$.

The existence of the dual system with nice properties entails the *UNI* Condition, as we now prove it.

Lemma 7. *A Euclidean system of the Good Class that admits a dual system that belongs to the Good Class satisfies the UNI Condition.*

Proof. We begin by proving *UNI (a)*. Denote by $\widehat{J}(h,\eta)$ the union of the intervals $\widehat{k}(\widehat{\mathcal{I}})$ relative to LFT's k that satisfy $\Delta(h,k) \leq \eta$. The proof has two steps. First, we describe the properties of $\widehat{J}(h,\eta)$, second, we relate $\widehat{J}(h,\eta)$ and $J(h,\eta)$.
First step. Relation (3.29) proves that, for $\Delta(h,k) \leq \eta$, the distance between the points $\widehat{h}(0)$ and $\widehat{k}(0)$ is less than $L\eta$. Second, since all the $\widehat{k}$ have the same depth n, the intervals $\widehat{k}(\widehat{\mathcal{I}})$ are disjoint. Then, the set $\widehat{J}(h,\eta)$ is an interval that contains $]\widehat{h}(0) - L\eta, \widehat{h}(0) + L\eta[$. Moreover, the length of a fundamental interval of depth n in the dual system has a length at most ρ^n(up to an

absolute constant), and, finally, the length of $\widehat{J}(h,\eta)$ satisfies

$$|\widehat{J}(h,\eta)| \ll L\eta + \rho^n. \tag{3.30}$$

Second Step. The fundamental intervals $h(\mathcal{I})$ and $\widehat{h}(\widehat{\mathcal{I}})$ have almost the same length: Thanks to bounded distortion property, and since $\widehat{h}'(0) = h'(0)$, for all $h \in \mathcal{H}^\star$

$$\frac{1}{(L\widehat{L})^{1/2}} \le \frac{|h(\mathcal{I})|}{|\widehat{h}(\widehat{\mathcal{I}})|} = \frac{|h(\mathcal{I})|}{|h'(0)|}\frac{|\widehat{h}'(0)|}{|\widehat{h}(\widehat{\mathcal{I}})|} \le (L\widehat{L})^{1/2}. \tag{3.31}$$

Then, the Lebesgue measures of $J(h,\eta)$ and $\widehat{J}(h,\eta)$ are closely related, and finally, one obtains, as soon as η satisfies $\eta \le \rho^{an}$ with $0 < a \le 1$,

$$|J(h,\eta)| \le (L\widehat{L})^{1/2}|\widehat{J}(h,\eta)| \ll (L\widehat{L})^{1/2}(L\eta + \rho^n) \ll \rho^{an}. \tag{3.32}$$

And, with (3.30, 3.32), Property *UNI(a)* holds.
We now deal with *UNI(b)*. The derivative $\Psi''_{h_1,h_2}(x)$ involves expressions of the form

$$\frac{h_i'''(x)h_i'(x)}{h_i'(x)^2}, \qquad \frac{h_i''(x)^2}{h_i'(x)^2},$$

for $i = 1,2$. Each term can be expressed as some multiple of

$$\frac{c_i^2}{(c_i x + d_i)^2} = \left(\frac{c_i}{d_i}\right)^2 \left|\frac{h_i'(x)}{h_i'(0)}\right|$$

which admits an absolute upper bound from distortion Property. □

Finally, Lemma 7 implies the following result:

Proposition 1. *Any dynamical system relative to a Euclidean Algorithm $\mathcal{G}, \mathcal{K}$, or $\mathcal{O}$ satisfies the* UNI *condition.*

This proves that Theorem 2 applies for the weighted transfer operator relative to any of the three Euclidean dynamical system and to any cost of moderate growth.

4. Limit Gaussian distribution for costs of moderate growth.

We now provide answers for the main first question, namely the distribution of costs on a rational trajectory, and prove Theorem 3. Section 2 has explained why the Dolgopyat estimates that we just obtained in Section 3 are so crucial. First, we have to extend these estimates on the neighborhood of the real axis (Subsection 4.1). Then, we give a precise description of our use of Perron's Formula (Subsection 4.2) that provides estimates (after a transfer operated in Subsection 4.4) for the moment generating function of a smoothed version $\overline{C}$ of our studied cost C. These estimates are uniform on a complex neighborhood of 0, so that the Quasi-Powers Theorem can be used, and entails asymptotic normality of cost $\overline{C}$, together with a bound for the speed of convergence and asymptotic estimates for the expectation and the variance of $\overline{C}$ (Subsection 4.5). In order to obtain results on cost C itself, a first comparison (done in Subsection 4.6) between distributions of cost C and $\overline{C}$ is sufficient to obtain the asymptotic normality of cost C with a speed of convergence in $O(1/\sqrt{\log N})$, which is the first part of Theorem 3. Then, in Subsection 4.7, we describe the asymptotic behaviour of the expectation and variance of cost C, which is the second part of Theorem 3. The last two subsections 4.8 and 4.9 are devoted to computing the constants that are involved in the asymptotic estimates on the total cost of rational trajectories, which is the third part of Theorem 3. In order to compare costs on rational trajectories with costs on truncated real trajectories, we prove, as a useful parenthesis, Theorem 1 (Subsection 4.8).

4.1. **Bounds for the quasi-inverse near the real axis.** Theorem 2 provides estimates for the norm $(1,t)$–norm of the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ of the transfer operator when $t = \Im s$ is sufficiently large. We now consider the case when s is close to the real axis.

Lemma 8. *Consider a (complex) neighborhood $\mathcal{W}$ of $w = 0$ on which the equation $\lambda(s,w) = 1$ defines an analytic function $s = \sigma(w)$ for which $\lambda(\sigma(w), w) = 1$. For any $\epsilon > 0$, there exist a (complex) neighborhood $\mathcal{W}' \subset \mathcal{W}$ of $w = 0$ and strictly positive numbers α_0, M_1, M_5, such that, for any α, with $0 < \alpha < \alpha_0$, for any (s,w) with $w \in \mathcal{W}'$, $\Re s = \Re\sigma(w) - \alpha, |t| \leq 1/\rho^2$, the $(1,1)$–norm of the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ satisfies*

$$||(I - \mathbf{H}_{s,w})^{-1}||_{1,1} \leq \frac{M_5}{\alpha},$$

and for any (s,w) with $w \in \mathcal{W}'$, $\Re s = \Re\sigma(w) - \alpha, |t| \geq 1/\rho^2$, the $(1,t)$–norm of the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ satisfies

$$||(I - \mathbf{H}_{s,w})^{-1}||_{1,t} \leq M_1 |t|^{\epsilon}.$$

Proof. Each vertical line $\Re s = \sigma$ is split into three regions : A first region near the real axis ($|t| \leq t_1$) where Perturbation Theory extends dominant spectral properties; An intermediate region ($t_1 \leq |t| \leq 1/\rho^2$) where Lemma 1 applies; The third region ($|t| \geq 1/\rho^2$) where Theorem 2 applies.

Near the real axis. When $(s,w) = (1,0)$, the transfer operator is the density transformer, and the operator is quasi–compact with a unique dominant eigenvalue $\lambda(1,0) = 1$, and a subdominant spectral radius $r < 1$. Moreover, the transfer operator satisfies the Property $An_{s,w}(1,0)$. Then, Perturbation Theory applies, with respect to variable (s,w), and there exists a (complex) neighborhood $\mathcal{A}$ of $(1,0)$ which satisfies the following:

(i) the operator $\mathbf{H}_{s,w}$ decomposes as

$$\mathbf{H}_{s,w} = \lambda(s,w)\, \mathbf{P}_{s,w} + \mathbf{N}_{s,w},$$

where $\mathbf{P}_{s,w}$ is the projector on the dominant eigensubspace, and $\mathbf{N}_{s,w}$ is an operator whose spectral radius equals the subdominant spectral radius $r_{s,w}$ of $\mathbf{H}_{s,w}$.

(ii) the mappings $(s,w) \mapsto \lambda(s,w)$, $(s,w) \mapsto \mathbf{P}_{s,w}$, $(s,w) \mapsto \mathbf{N}_{s,w}$ are well defined and analytic,

(iii) the spectral radius $r_{s,w}$ of the operator $\mathbf{N}_{s,w}$, i.e., the subdominant spectral radius of $\mathbf{H}_{s,w}$ is less than $1 - \delta$, so that the $(1,1)$–norm of the quasi–inverse $(I - \mathbf{N}_{s,w})^{-1}$ is bounded by some M_2.

(iv) Since $\lambda'_s(1,0)$ is not zero, Implicit Function Theorem applies, and there exists a portion of the "curve" $\mathcal{C} := \{(\sigma(w), w)\}$ that lies inside the neighborhood $\mathcal{A}$. Then, there exists some strictly positive α_2 for which a α_2–neighborhood of the curve $\mathcal{C}$, namely a subset of the form

$$\mathcal{A}_2 := \{(s,w); w \in \mathcal{W}_2, |\Re s - \Re\sigma(w)| \le \alpha_2, |t| \le t_2\}$$

lies inside $\mathcal{A}$ (for some complex neighborhood $\mathcal{W}_2$ of $w = 0$.

(v) Since $\lambda'_s(1,0)$ is not zero, the modulus of the function

$$(s,w) \mapsto \frac{\lambda(s,w) - 1}{s - \sigma(w)}$$

admits a strictly positive lower bound ℓ on $\mathcal{A}_2$.

When (s,w) is in $\mathcal{A}_2$, the quasi-inverse of $\mathbf{H}_{s,w}$ satisfies

$$(I - \mathbf{H}_{s,w})^{-1} = \frac{\lambda(s,w)}{1 - \lambda(s,w)}\,\mathbf{P}_{s,w} + (I - \mathbf{N}_{s,w})^{-1}. \tag{4.1}$$

It has a simple pole at each point $(s = \sigma(w), w)$ of the curve $\mathcal{C}$ with a residue $\mathbf{R}(w)$

$$\mathbf{R}(w) := \frac{-1}{\lambda'_s(\sigma(w), w)}\,\mathbf{F}_{\sigma(w),w}\mathbf{P}_{\sigma(w),w} \tag{4.2}$$

Consider now, for any α that satisfies $0 < \alpha \le \alpha_2$ the subset

$$\{(s,w);\ \Re s = \Re\sigma(w) - \alpha,\ |t| \le t_2,\ w \in \mathcal{W}_2\}.$$

Here, with Property (v), the dominant eigenvalue satisfies

$$|\lambda(s,w) - 1| \ge \ell|s - \sigma(w)| \ge \ell\alpha,$$

and, with (4.1) and Properties (ii), (iii), the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ satisfies on this set

$$||(I - \mathbf{H}_{s,w})^{-1}||_{1,1} \le \frac{M_4}{\alpha}$$

Second Region. Consider the reals $t_0 = 1/\rho^2$ of Theorem 2 and t_2 of the beginning of this Lemma. Thanks to Lemma 1, when (s,w) lies inside the compact subset

$$\mathcal{B} := \{(s,w);\ w \in \mathcal{W}_2,\ \Re s = \sigma(\Re w),\ t_2 \le |t| \le t_0, \},$$

the spectral radius $R(s, \Re w)$ satisfies

$$R(s, \Re w) < R(\Re s, \Re w) = 1,$$

so that there exists some $\beta > 0$ such that, for any $(s,w) \in \mathcal{B}$, the spectral radius $R(s, \Re w)$ of $\mathbf{H}_{s,\Re w}$ satisfies $R(s, \Re w) \le 1 - 2\beta$. The map σ is continuous, and $\sigma(w)$ is real for real w, so that the two real numbers $\sigma(\Re w)$ and $\Re\sigma(w)$ are close when w is near 0. Moreover, the map $(s,w) \mapsto R(s,w)$ is continuous, so that the two spectral radii $R(s,w)$ and $R(s,\Re w)$ are close when w is near 0. Finally, there exists some strictly positive α_3 and a complex neighborhood $\mathcal{W}_3$ of $w = 0$ such that the spectral radius $R(s,w)$ satisfies $R(s,w) \le 1 - \beta$ on

$$\mathcal{A}_3 := \{(s,w);\ w \in \mathcal{W}_3, |\Re s - \Re\sigma(w)| \le \alpha_3,\ t_2 \le |t| \le t_0, \}.$$

Then the quasi-inverse $(s,w) \mapsto (I - \mathbf{H}_{s,w})^{-1}$ is analytic on $\widetilde{\mathcal{B}}$ and thus, its $(1,1)$–norm is bounded by some M_3 on $\widetilde{\mathcal{B}}$.

Third Region. Consider some ϵ fixed, and the real neighborhood $\Sigma \times W$ that is mentionned in Theorem 2. There exists some strictly positive α_4 and a complex neighborhood $\mathcal{W}_4$ of $w = 0$ such that any (s, w) of

$$\mathcal{A}_4 := \{(s, w);\ w \in \mathcal{W}_4, |\Re s - \Re\sigma(w)| \le \alpha_4, |t| \le t_0, \}$$

has its real part (σ, ν) that belongs to $\Sigma \times W$.

We now choose $\alpha_0 := \min(\alpha_2, \alpha_3, \alpha_4)$, and $\mathcal{W}' := \mathcal{W}_2 \cap \mathcal{W}_3 \cap \mathcal{W}_4$. Taking $M_5 := \max(M_4, M_3)$, and M_1 as in Theorem 2, we obtain the result. $\square$.

We now apply this Lemma to Dirichlet series $S(s, w)$ or $\widetilde{S}(s, w)$ defined in (2.10, 2.11). With (2.16, 2.17) which relate these Dirichlet series to the quasi-inverse of the operator $\mathbf{H}_{s,w}$, we obtain the following result:

Corollary 9. *Consider a (complex) neighborhood $\mathcal{W}$ of $w = 0$ on which the equation $\lambda(s, w) = 1$ defines an analytic function $s = \sigma(w)$ for which $\lambda(\sigma(w), w) = 1$. For any $\epsilon > 0$, there exist a (complex) neighborhood $\mathcal{W}'$ of $w = 0$, and strictly positive numbers α_0, M, such that, for any α, with $0 < \alpha < \alpha_0$, for any (s, w) with $w \in \mathcal{W}'$, $\Re s = \Re\sigma(w) - \alpha$, the Dirichlet series $S(s, w)$ or $\widetilde{S}(s, w)$ satisfy*

$$|S(s, w)| \le \frac{M}{\alpha} \max(1, |t|^{\epsilon}), \qquad |\widetilde{S}(s, w)| \le \frac{M}{\alpha} \max(1, |t|^{\epsilon})$$

Furthermore, for each $w \in \mathcal{W}'$, the series $S(s, w)$, $\widetilde{S}(s, w)$ are meromorphic in the strip $|\Re s - \Re\sigma(w)| \le \alpha$; they have a unique pole inside this strip, at $s = \sigma(w)$.

Proof. For $S(s, w)$, this is an easy consequence of the previous Lemma. For $\widetilde{S}(s, w) = \zeta(s)S(s, w)$, we furthermore use that $\zeta(s)$ is bounded on vertical strips near $\Re s = 2$. $\square$

4.2. **Using Perron's Formula.** We choose some $\epsilon < 1/3$. Then Corollary 9 defines α_0 and $\mathcal{W}'$. We fix some $\alpha \le \alpha_0$ and some $w \in \mathcal{W}'$. The Dirichlet series $S(s, w)$ are viewed as functions of parameter s. Recall that the Dirichlet series $S(s, w)$ satisfies

$$S(2s, w) := \sum_{n \ge 1} \frac{c_n(w)}{n^{2s}} = \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0).$$

It is analytic in the punctured half–plane $\{\Re s \ge \Re\sigma(w), s \ne \sigma(w)\}$. The classical identity

$$\frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} \frac{y^{s+1}}{s(s+1)} ds = 0 \text{ if } \ 0 < y \le 1, \quad y - 1 \text{ if } \ y > 1$$

that holds for $D > 0$ is applied termised for each fixed w to the Dirichlet series

$$S(s, w) := \sum_{n \ge 1} \frac{c_n(w)}{n^s}.$$

It entails the equality, well-known as Perron's formula [18],

$$\Psi_w(T) := \sum_{n \le T} c_n(w)(T - n) = \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} S(2s, w) \frac{T^{2s+1}}{s(2s+1)} ds$$

$$= \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0) \frac{T^{2s+1}}{s(2s+1)} ds$$

provided $D > \Re\sigma(w)$.

We consider the strip $\mathcal{S}(w)$ that is limited by thetwo vertical lines $\Re s = \Re\sigma(w) - \alpha$ and $\Re s = \Re\sigma(w) + \alpha$. This strip contains $s = \sigma(w)$ as unique (simple) pole of $(I - \mathbf{H}_{s,w})^{-1}$. We consider the rectangle $\mathcal{U}(w)$ delimited by the strip $\mathcal{S}(w)$ and the two horizontal lines $\Im s = \pm U$. Then, by the Cauchy's residue theorem,

$$\frac{1}{2i\pi} \int_{\mathcal{U}(w)} S(2s, w) \frac{T^{2s+1}}{s(2s+1)} ds = \frac{A(w)}{\sigma(w)(2\sigma(w)+1)} T^{2\sigma(w)+1}$$

where $A(w)$ is related to residue $\mathbf{R}$ defined in (4.2) via the equality $A(w) = \mathbf{R}(w)[1](0)$.

We now let U tend to ∞. We first apply Perron's Formula with $D = \Re\sigma(w) + \alpha$. Now, Corollary 9 proves that the integral on the left vertical line $\Re s = \Re\sigma(w) - \alpha$ exists and satisfies

$$\int_{\Re\sigma(w)-\alpha-i\infty}^{\Re\sigma(w)-\alpha+i\infty} S(2s, w) \frac{T^{2s+1}}{s(2s+1)} ds = O\left(\frac{1}{\alpha} T^{2\Re\sigma(w)-2\alpha+1}\right),$$

(with a O–term which is uniform for $w \in \mathcal{W}'$, $\alpha \le \alpha_0$ and $T \to \infty$), while the integrals on the horizontal lines of $\mathcal{U}(w)$ tend to zero for $U \to \infty$. Finally, one obtains

$$\Psi_w(T) = \frac{A(w)}{\sigma(w)(2\sigma(w)+1)} T^{2\sigma(w)+1} \left[1 + O\left(\frac{T^{-2\alpha}}{\alpha}\right)\right], \tag{4.3}$$

with a O–term uniform for $w \in \mathcal{W}'$, $\alpha \le \alpha_0$ and $T \to \infty$. Since all these estimates involve constants that are uniform when α satisfies $0 < \alpha < \alpha_0$, we can choose $\alpha := \alpha(T)$ as a function of T, so that

$$\epsilon_{\alpha(T)}(T) = \frac{T^{-2\alpha}}{\alpha} = \frac{1}{\log T}.$$

This value $\alpha(T)$ satisfies

$$\frac{1}{2} \frac{\log\log T}{\log T} \le \alpha(T) \le \frac{\log\log T}{\log T},$$

and since $\alpha(T)$ tends to 0 for $T \to \infty$, it is less than α_0 for T sufficiently large. Then, we obtain

$$\Psi_w(T) = \frac{A(w)}{\sigma(w)(2\sigma(w)+1)} T^{2\sigma(w)+1} \left[1 + O\left(\frac{1}{\log T}\right)\right], \tag{4.4}$$

with a O–term which is uniform for $w \in \mathcal{W}'$ and $T \to \infty$.

For integer values $T = N$, the quantity

$$\frac{1}{N} \Psi_w(N) = \frac{1}{N} \sum_{Q \le N} \sum_{n \le Q} c_n(w)$$

is an average of quantities $\Phi_w(Q) := \sum_{n \le Q} c_n(w)$. Our eventual goal is to transfer the estimates on $\Psi_w(N)$ into estimates on $\Phi_w(Q)$. In order to do so, we introduce an intermediate model, called the smoothed model.

4.3. **Smoothed models and smoothed costs.** Associate to some positive function $t \mapsto \epsilon(T)$, with $\epsilon(T) \le 1$, the probabilistic model $\overline{\Omega}_N(\epsilon)$ defined as follows. For any integer N, choose first uniformy an integer Q between $N - \lfloor N\epsilon(N) \rfloor$ and N, then draw uniformly an element (u, v) of Ω_Q. Remark that for $\epsilon = 0$, the equality $\overline{\Omega}_N(0) = \Omega_N$. In order to insist on the difference of the two models, the total cost of the trajectory $\mathcal{T}(u/v)$ is now denoted by $\overline{C}(u, v)$, and the cost $\overline{C}$ can be viewed as a smoothed version of cost C. We first study this smoothed cost $\overline{C}$. Then, the weighted cumulative value of $\exp[w\overline{C}]$ on $\overline{\Omega}_N(\epsilon)$ equals

$$\overline{\Phi}_w(N) := \frac{1}{\lfloor N\epsilon(N) \rfloor} \sum_{Q=N-\lfloor N\epsilon(N)\rfloor}^{N} \sum_{n \le Q} c_n(w), \tag{4.5}$$

so that the moment generating function of cost $\overline{C}$ on $\overline{\Omega}_N(\epsilon)$ is exactly the ratio

$$\mathrm{E}_N[\exp(w\overline{C})] = \frac{\overline{\Phi}_w(N)}{\overline{\Phi}_0(N)}. \tag{4.6}$$

Remark that this quantity $\overline{\Phi}_w(N)$ equals $(1/N)\Psi_w(N)$ for $\epsilon = 1$. Generally speaking, it can be expressed as a function of $\Psi_w(T)$,

$$\overline{\Phi}_w(N) = \frac{1}{\lfloor N\epsilon(N) \rfloor} \left[\Psi_w(N) - \Psi_w(N - \lfloor N\epsilon(N) \rfloor)\right]. \tag{4.7}$$

We now proceed in three main steps. First, we transfer the estimates that we obtained for Ψ_w into estimates for $\overline{\Phi}_w$. Second, with these estimates, we obtain an asymptotic Gaussian law for the smoothed cost $\overline{C}$ relative to $\epsilon(N) = 1/\sqrt{\log N}$, with a speed of convergence of order $1/\sqrt{\log N}$. Third, for this choice of ϵ, we compare directly the two probabilistic models and we transfer the asymptotic Gaussian law for $\overline{C}$ into an asymptotic Gaussian law for C with the same speed of convergence of order $1/\sqrt{\log N}$.

4.4. **Transfer of estimates.** First, via (4.7), we transfer this bound (4.4) for Ψ into a bound for $\overline{\Phi}_w(T)$ defined in (4.5). We use here a general easy result, classical in number theory contexts, which will be also useful in Section 5.

Lemma 10. *Let $\mathcal{W}$ be a complex neighborhood of $w = 0$. Consider a sequence $c_n(w)$ for which the quantities $\Psi_w(T) := \sum_{n \le T} c_n(w)(T-n)$ admit, when $T \to \infty$, an estimate of the form*

$$\Psi_w(T) = F_w(T)\left[1 + O\left(G(T)\right)\right]$$

with

$$F_w(T) = B(w)T^{a(w)} \cdot (\log T)^{b(w)} \cdot \mathcal{P}(\log T)$$

for some $B(w), a(w), b(w)$ and some polynomial $\mathcal{P}$ which does not depend on w. Assume that

(i) the O−error term is uniform when w varies in $\mathcal{W}$,

(ii) the scalars $B(w), a(w), b(w)$ are uniformly bounded for $w \in \mathcal{W}$ and $a(w), B(w)$ satisfy $\Re a(w) > 1, B(w) \neq 0$ on $\mathcal{W}$,

(iii) the function $G : T \mapsto G(T)$ tends to 0 for $T \to \infty$ and is of moderate variation, i.e., there exists K so that, for any c with $1/2 \le c \le 2$, one has $|G(cT)| \le K|G(T)|$.

Then, the following is true, for $T \to \infty$:
(a) One has

(4.8)
$$\frac{1}{TG(T)^{1/2}}\left[\Psi_w(T) - \Psi_w\left(T - TG(T)^{1/2}\right)\right] = F'_w(T)\ \left[1 + O\left(G(T)^{1/2}\right)\right],$$

with a O–term which is uniform when w belongs to $\mathcal{W}$.
(b) Furthermore, if the sequence $G(T)^{-1}$ is $O(T)$, then the estimate (4.8) extends to integer parts, and the following estimate

(4.9)
$$\frac{1}{\lfloor TG(T)^{1/2}\rfloor}\left[\Psi_w(T) - \Psi_w\left(T - \lfloor TG(T)^{1/2}\rfloor\right)\right] = F'_w(T)\ \left[1 + O\left(G(T)^{1/2}\right)\right]$$

holds too, with a O–term which is uniform when w belongs to $\mathcal{W}$.
(c) Suppose now that the sequence $c_n(w)$ is positive for $w \in \mathcal{W} \cap \mathbf{R}$. Then, $\Phi_w(T)$ satisfy

$$\Phi_w(T) := \sum_{n \le T} c_n(w) = F'_w(T)\left[1 + O\left(G(T)^{1/2}\right)\right],$$

with a O–term which is uniform when w belongs to $\mathcal{W} \cap \mathbf{R}$.

Proof. Consider some sequence $\epsilon(T)$ which tends to 0. The estimate for $\Psi_w(T)$ and property (iii) then entail that

$$\frac{1}{T\epsilon(T)}[\Psi_w(T) - \Psi_w(T - T\epsilon(T))]$$
$$= \frac{1}{T\epsilon(T)}\left(F_w(T) - F_w(T - T\epsilon(T))\right) + \frac{1}{T\epsilon(T)} O\left(F_w(T)G(T)\right)$$
$$= F'_w(T)\left[1 + O\left(T\epsilon(T)\frac{F''_w(T)}{F'_w(T)}, \frac{1}{T\epsilon(T)}\frac{F_w(T)G(T)}{F'_w(T)}\right)\right].$$

Then, due to the special form of $F_w(T)$ and properties (ii) of a_w, b_w, the derivatives $F'_w(T), F''_w(T)$ satisfy

$$F'_w(T) = \Theta\left(T^{-1}F_w(T)\right), \qquad F''_w(T) = \Theta\left(T^{-2}F_w(T)\right)$$

with a uniform Θ, and, with the choice

$$T\epsilon(T) := \left(\frac{F_w(T)G(T)}{F''_w(T)}\right)^{1/2} = \Theta\left(TG(T)^{1/2}\right),$$

Part (a) of the Lemma is proven. For Part (b), remark that the difference of the left members of (4.8) and (4.9) can be split into two terms of order

$$F'_w(T) \cdot O\left(\frac{1}{TG(T)^{1/2}}\right) = F'_w(T) \cdot O\left(G(T)^{1/2}\right),$$

which proves Part (b).
For (c), remark that

$$\frac{1}{T\epsilon}[\Psi_w(T + T\epsilon) - \Psi_w(T)] = \frac{1}{T\epsilon}\int_T^{T+T\epsilon} \Phi_w(t)dt \ge \Phi_w(T), \tag{4.10}$$

$$\frac{1}{T\epsilon}[\Psi_w(T) - \Psi_w(T - T\epsilon)] = \frac{1}{T\epsilon}\int_{T-T\epsilon}^{T} \Phi_w(t)dt \le \Phi_w(T). \tag{4.11}$$

Then, with (4.10), and (4.8)

$$\Phi_w(T) \le F'_w(T) \ \left[1 + O\left(G(T)^{1/2}\right)\right]$$

With a similar argument based on (4.8) and the other inequality (4.11), we obtain finally the desired estimate for $\Phi_w(T)$. $\square$

We apply now Part (*b*) of the previous Lemma to our setting. With (4.4) and $G(T) = 1/\log T$, this provides the convenient estimate for the cumulative cost $\overline{\Phi}_w(N)$ on $\overline{\Omega}_N$,

$$\overline{\Phi}_w(N) = \frac{A(w)}{\sigma(w)} N^{2\sigma(w)} \left[1 + O\left(\frac{1}{\sqrt{\log N}}\right)\right]$$

where the O–term is uniform when w varies in a sufficiently small neighborhood of 0 and N tends to ∞. In the case $w = 0$, one has

(4.12)
$$\overline{\Phi}_0(N) = A(0)\, N^2 \left[1 + O\left(\frac{1}{\sqrt{\log N}}\right)\right], \qquad \text{with} \quad A(0) = -\frac{\mathbf{F}_1\mathbf{P}_1[1](0)}{\lambda'_1}.$$

Finally, with (4.6), we obtain the following estimate:

For any integer N, choose first uniformy an integer Q between $N - \lfloor N/\sqrt{\log N}\rfloor$ and N, then draw uniformly an element (u, v) of Ω_Q. The moment generating function of the relative smoothed cost $\overline{C}$ satisfies

$$\mathrm{E}_N[\exp(w\overline{C})] = \frac{A(w)}{A(0)\sigma(w)} N^{2\sigma(w)} \left[1 + O\left(\frac{1}{\sqrt{\log N}}\right)\right],$$

with a O–term which is uniform when N tends to ∞ and w is near 0.

4.5. **Asymptotic Gaussian law for the smoothed cost $\overline{C}$.** The moment generating function $\mathrm{E}_N[\exp(w\overline{C})]$ relative to the smoothed cost $\overline{C}$ satisfies

$$\mathrm{E}_N[\exp(w\overline{C})] = \exp(D_N(w)) \ \left[1 + O\left(\frac{1}{\sqrt{\log N}}\right)\right],$$

with

$$D_N(w) = 2[\sigma(w) - \sigma(0)] \log N + \log \frac{A(w)}{A(0)\sigma(w)},$$

and a O–term which is uniform when N tends to ∞ and w is near 0. This shows that the moment generating function $\mathrm{E}_N[\exp(w\overline{C})]$ is a "uniform" quasi-power when N tends to ∞ and w is a complex near 0.

At this stage, we need a process for extracting cumulative distribution function from the quasi-power form of moment generating functions. This is achieved in essence by a combination of Lévy's continuity theorem for characteristic functions and the Berry-Esseen inequality. This is encapsulated in the following Theorem due to Hwang [24, 25, 26]:

Theorem [Hwang]. *Assume that the moment generating function relative to some cost $\overline{C}$ on a sequence of probabilistic spaces $\overline{\Omega}_N$ are analytic in a complex neighborhood $\mathcal{W}'$ of 0, and satisfy on $\mathcal{W}'$ an expansion of the form*

$$\mathrm{E}_N[\exp(w\overline{C})] = \exp[\beta_N U(w) + V(w)] \left(1 + O(\frac{1}{\kappa_N})\right),$$

with $\beta_N, \kappa_N \to \infty$ as $N \to \infty$, and $U(w), V(w)$ analytic on $\mathcal{W}$. Assume also that $U''(0) \neq 0$. Then, the mean and the variance satisfy

$$\mathrm{E}_N[\overline{C}] = \beta_N U'(0) + V'(0) + O(\kappa_N^{-1}),$$

$$\mathrm{Var}_N[\overline{C}] = \beta_N U''(0) + V''(0) + O(\kappa_N^{-1}).$$

The distribution of $\overline{C}$ on $\overline{\Omega}_N$ is asymptotically Gaussian and the speed of convergence to the Gaussian limit is $O(\kappa_N^{-1} + \beta_N^{-1/2})$.

The application of this Theorem to our context is straightforward, with $\beta_N := \log N, \kappa_N = \sqrt{\log N}$,

$$U(w) = 2(\sigma(w) - \sigma(0)), \qquad V(w) = \log \frac{A(w)}{A(0)\sigma(w)}.$$

The function $\sigma(w)$ then plays a central rôle; we recall that is defined in an implicit way from the dominant eigenvalue $\lambda(s, w)$ of the weighted transfer operator $\mathbf{H}_{s,w}$ by the relation $\lambda(\sigma(w), w) = 1$. We will see in the sequel that function σ is always strictly convex, so that $U''(0) = 2\sigma''(0)$ is not zero and the Quasi-Powers Theorem can be applied and provides asymptotic estimates for the expectation and the variance of the smoothed cost $\overline{C}$. Finally, we proved the following:

Choose first uniformy an integer Q between $N - \lfloor N/\sqrt{\log N} \rfloor$ and N, then draw uniformly an element (u, v) of Ω_Q. The associated smoothed version $\overline{C}$ of cost C has a distribution that is asymptotically Gaussian, with a speed of convergence in $O(1/\sqrt{\log N})$. Moreover, there are asymptotic estimates for the expectation and the variance of the smoothed cost $\overline{C}$

$$\mathrm{E}_N[\overline{C}] = 2\sigma'(0) \log N + V'(0) + O(\frac{1}{\sqrt{\log N}}), \tag{4.13}$$

$$\mathrm{Var}_N[\overline{C}] = 2\sigma''(0) \log N + V''(0) + O(\frac{1}{\sqrt{\log N}}). \tag{4.14}$$

4.6. **Asymptotic normality of cost C.** We are interested in the behaviour of the random variable C on Ω_N. In fact, the distribution of the two variables C and $\overline{C}$ are sufficiently close so that that the distribution of C will be also asymptotically gaussian, with a speed of convergence in $O(1/\sqrt{\log N})$.

Lemma 11. *Suppose that the sequence $N\epsilon(N)$ tends to ∞. The distance between the two distributions on $\overline{\Omega}_N(\epsilon)$ and on Ω_N is $O(\epsilon(N))$. Furthermore, in $\overline{\Omega}_N(\epsilon)$, the ratio between the probability of the most probable elements and the probability of the least probable elements is of order $N\epsilon(N)$.*

Proof. First recall that the cardinality $|\Omega_N| \sim (6/\pi^2)N^2$, so that the probability $\mathrm{Pr}_N(u, v)$ for any $(u, v) \in \Omega_N$ satisfies

$$\mathrm{Pr}_N(u, v) \sim \frac{\pi^2}{6} \frac{1}{N^2}.$$

Denote by $N' := N - \lfloor N\epsilon(N) \rfloor$. Then the set Ω_N decomposes into two subsets, the ordinary subset $\mathcal{O}_N := \Omega_{N'}$ and the exceptional subset $\mathcal{E}_N :=$

$\Omega_N \setminus \mathcal{O}_N$. This last set is called exceptional since its total probability satisfies

$$\Pr_N(\mathcal{E}_N) = O\left(\frac{N^2 - N'^2}{N^2}\right) = O(\epsilon(N)).$$

Now, the probability $\overline{\Pr}_N(u,v)$ satisfies, for any $(u,v) \in \mathcal{O}_N$,

$$\overline{\Pr}_N(u,v) = \sum_{q=N'}^{N} \Pr_q(u,v) \cdot \Pr[Q=q] = \frac{1}{\lfloor N\epsilon(N) \rfloor} \sum_{q=N'}^{N} \frac{1}{|\Omega_q|}.$$

Then, the estimates

$$\sum_{q=N'}^{N} \frac{1}{|\Omega_q|} \sim \frac{\pi^2}{6} \sum_{q=N'}^{N} \frac{1}{q^2} \sim \frac{\pi^2}{6}\left[\frac{1}{N'} - \frac{1}{N}\right] \sim \frac{\pi^2}{6N^2} \lfloor N\epsilon(N) \rfloor \frac{N}{N'}$$

show that

$$\overline{\Pr}_N(u,v) \sim \frac{N}{N'} \Pr_N(u,v) \qquad \text{for any } (u,v) \in \mathcal{O}_N,$$

and

$$\overline{\Pr}_N(\mathcal{O}_N) \sim \frac{N}{N'} \Pr_N(\mathcal{O}_N) \sim \frac{N'}{N}.$$

Finally,

$$1 - \overline{\Pr}_N(\mathcal{O}_N) = \overline{\Pr}_N(\mathcal{E}_N) = O(\epsilon(N)).$$

Then, for any set $A \subset \Omega_N$, the difference $|\Pr_N(A) - \overline{\Pr}_N(A)|$ is less than

$$|\Pr_N(A \cap \mathcal{O}_N) - \overline{\Pr}_N(A \cap \mathcal{O}_N)| + |\Pr_N(A \cap \mathcal{E}_N) - \overline{\Pr}_N(A \cap \mathcal{E}_N)|.$$

The first term is $O(\epsilon)\Pr_N(A \cap \mathcal{O}_N) = O(\epsilon)$ while the second term is the difference of two positive quantities, each of them being $O(\epsilon)$: it is $O(\epsilon)$. $\square$

Then, the difference between the distributions of cost C and smoothed cost $\overline{C}$ relative to the choice $\epsilon(N) = 1/\sqrt{\log N}$ is $O(1/\sqrt{\log N})$. This shows that the asymptotic distribution of cost C is Gaussian, with a speed of convergence in $O(1/\sqrt{\log N})$. This proves the first Part of Theorem 3.

However, this comparison between distributions does not provide any precise transfer from the expectation $\mathrm{E}_N[\overline{C}]$ and the variance $\mathrm{Var}_N[\overline{C}]$ of the smoothed cost to the expectation $\mathrm{E}_N[C]$ and the variance $\mathrm{Var}_N[C]$ of cost C. This is due to the second part of Lemma 11, which proves that the distribution on $\overline{\Omega}_N(\epsilon)$ is not at all uniform. In order to obtain precise asymptotic estimates for the first two moments $\mathrm{E}_N[C]$ and $\mathrm{E}_N[C^2]$, we now use methods which are similar to those used previously (Perron's formula, estimates à la Dolgopyat). In passing, we make use of Equations (4.13, 4.14) in order to identify the constants appearing in the mean $\mathrm{E}_N[C]$ and the variance $\mathrm{Var}_N[C]$.

4.7. **Asymptotic behaviour of the expectation and variance of cost C.** We consider the Dirichlet series $S^{(1)}(s)$ and $S^{(2)}(s)$ that are the first two derivatives of $S(s,w)$ with respect to w at $w=0$. They satisfy

$$S^{(1)}(s) = \sum_{(u,v)\in\Omega} \frac{C(u,v)}{v^s} = \sum_{n\geq 1} \frac{c_n'(0)}{n^{2s}},$$

$$S^{(2)}(s) = \sum_{(u,v)\in\Omega} \frac{C^2(u,v)}{v^s} = \sum_{n\geq 1} \frac{c_n''(0)}{n^{2s}},$$

and they can be viewed as Dirichlet generating function for cost C and C^2. Taking the first two derivatives of the fundamental relation (2.16)

$$S(2s,w) = \mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0), \tag{4.15}$$

provides alternative expressions of $S^{(1)}(s)$ and $S^{(2)}(s)$ that involve the quasi-inverse of the plain transfer operator $\mathbf{H}_s$, together with derivatives (with respect to w, at $w=0$) of $\mathbf{F}_{s,w}$ and $\mathbf{H}_{s,w}$, as follows:

$$S^{(1)}(2s) = \mathbf{F}_s \circ (I - \mathbf{H}_s)^{-1} \circ \mathbf{H}_s^{(1)} \circ (I - \mathbf{H}_s)^{-1}[1](0) + \mathbf{F}_s^{(1)} \circ (I - \mathbf{H}_s)^{-1}[1](0)$$

$$S^{(2)}(2s) = \mathbf{F}_s \circ (I - \mathbf{H}_s)^{-1} \circ \mathbf{H}_s^{(1)} \circ (I - \mathbf{H}_s)^{-1} \circ \mathbf{H}_s^{(1)} \circ (I - \mathbf{H}_s)^{-1}[1](0)$$

$$+\mathbf{F}_s \circ (I - \mathbf{H}_s)^{-1} \circ \mathbf{H}_s^{(2)} \circ (I - \mathbf{H}_s)^{-1}[1](0) + \mathbf{F}_s^{(2)} \circ (I - \mathbf{H}_s)^{-1}[1](0)$$

These series are analytic in the punctured half– plane $\{\Re s \geq 1), s \neq 1\}$. At $s=1$, thanks to (4.1), the series $S^{(1)}(2s)$ admits a pole of order 2, and the series $S^{(2)}(2s)$ admits a pole of order 3. Perron's formulae apply,

$$\Psi_{(1)}(T) := \sum_{n\leq T} c_n'(0)(T-n) = \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} S^{(1)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds$$

$$\Psi_{(2)}(T) := \sum_{n\leq T} c_n''(0)(T-n) = \frac{1}{2i\pi} \int_{D-i\infty}^{D+i\infty} S^{(2)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds$$

provided $D > 1$.

We choose some $\epsilon < 1/3$. Then Corollary 9 defines α_0. We fix some $\alpha \leq \alpha_0$, and we consider the strip $\mathcal{S}$ that is delimited by the vertical lines $\Re s = 1-\alpha$ and $\Re s = 1+\alpha$. This strip contains $s=1$ as unique (simple) pole of $(I-\mathbf{H}_s)^{-1}$. We consider the rectangle $\mathcal{U}$ delimited by the strip $\mathcal{S}$ and the two horizontal lines $\Im s = \pm U$. Then, by the Cauchy Theorem,

$$\frac{1}{2i\pi} \int_{\mathcal{U}} S^{(1)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds = T^3(a \log T + b),$$

$$\frac{1}{2i\pi} \int_{\mathcal{U}} S^{(2)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds = T^3(c \log^2 T + d \log T + e),$$

for some constants a, b, c, d, e.

We now let U tend to ∞. We first apply Perron's Formula with $D = 1+\alpha$. Now, Corollary 9 proves that, on the left vertical line $\Re s = 1-\alpha$, the two Dirichlet series $S^{(1)}(2s)$ and $S^{(2)}(2s)$ satisfy

$$|S^{(1)}(2s)| \leq M^{(1)} \sup(1, |t|^{2\epsilon}), \qquad |S^{(2)}(2s)| \leq M^{(2)} \sup(1, |t|^{3\epsilon}).$$

Since we have chosen $\epsilon < 1/3$, the two integrals on the left vertical line $\Re s = 1-\alpha$ exist and satisfy

$$\int_{1-\alpha-i\infty}^{1-\alpha+i\infty} S^{(1)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds = O(T^{3-2\alpha}),$$

$$\int_{1-\alpha-i\infty}^{1-\alpha+i\infty} S^{(2)}(2s) \frac{T^{2s+1}}{s(2s+1)} ds = O(T^{3-2\alpha}),$$

while the integrals on the horizontal lines of $\mathcal{U}$ tend to zero for $U \to \infty$. Finally,

$$\Psi_{(1)}(T) = T^3(a \log T + b) \ \left[1 + O\left(\frac{1}{T^{2\alpha} \log T}\right)\right]. \tag{4.16}$$

$$\Psi_{(2)}(T) = T^3(c \log^2 T + d \log T + e) \ \left[1 + O\left(\frac{1}{T^{2\alpha} \log^2 T}\right)\right]. \tag{4.17}$$

Then Part (c) of Lemma 10 transfers these estimates to

$$\Phi_{(1)}(T) = T^2(a_1 \log T + b_1) \ \left[1 + O\left(\frac{1}{T^{\alpha} \sqrt{\log T}}\right)\right]. \tag{4.18}$$

$$\Phi_{(2)}(T) = T^2(c_1 \log^2 T + d_1 \log T + e_1) \ \left[1 + O\left(\frac{1}{T^{\alpha} \log T}\right)\right]. \tag{4.19}$$

This implies that, (if we change α into a smaller α)

$$\mathrm{E}_N[C] = a_2 \log N + b_2 + O\left(\frac{1}{N^\alpha}\right),$$

and

$$\mathrm{Var}_N[C] = c_2 \log^2 N + d_2 \log N + e_2 + O\left(\frac{1}{N^\alpha}\right).$$

In fact, we shall see in a moment the coefficient c_2 is zero ... The computations of coefficients a_2, b_2, c_2, d_2, e_2 as a function of spectral objects of the transfer operator and its derivative is possible, but quite involved. We are mainly interested in the constants a_2, c_2, d_2 that appear in the dominant terms of the expansions. Remark first that the expectation $\mathrm{E}_N[C]$ and the variance $\mathrm{Var}_N[C]$ are related to their analogues $\mathrm{E}_N[\overline{C}]$, $\mathrm{Var}_N[C]$ relative to the smoothed cost $\overline{C}$, via the relations

$$\mathrm{E}_N[\overline{C}] = \frac{1}{\lfloor N\epsilon(N) \rfloor} \sum_{Q=N-\lfloor N\epsilon(N)\rfloor}^{N} \mathrm{E}_Q[C],$$

$$\mathrm{Var}_N[\overline{C}] = \frac{1}{\lfloor N\epsilon(N) \rfloor} \sum_{Q=N-\lfloor N\epsilon(N)\rfloor}^{N} \mathrm{Var}_Q[C],$$

(with the choice $\epsilon(N) = 1/\sqrt{\log N}$). Comparing the asymptotic estimates for the expectations and the variances of the smoothed cost obtained in (4.13, 4.14) and the estimates that we just obtained provides the equalities

$$c_2 = 0, \qquad a_2 = 2\sigma'(0), \qquad c_2 = 2\sigma''(0).$$

This proves the Second Part of Theorem 3. It remains to study these derivatives $\sigma'(0)$ and $\sigma''(0)$ in order to obtain the Third Part of Theorem 3. This is the purpose of the following subsection.

4.8. **Evaluation and computation of constants.** We study the two derivatives $\sigma'(0)$ and $\sigma''(0)$. We first consider the particular case of the trivial digit–cost that gives rise to the number of steps $C = P$. After a (useful) parenthesis, where we return for a moment to the continuous case and prove Theorem 1, we study the general case, and finally, we prove that $\sigma''(0)$ is not zero.

Number of steps. Let us begin by the particular case when the cost is the number P of steps. Then the digit–cost c is the trivial one, $c = 1$. Then σ is defined in an implicit way from the dominant eigenvalue $\lambda(s)$ of the plain transfer operator by the relation $\lambda(\sigma(w)) = e^{-w}$. Then

$$\sigma'(w) = -\frac{e^{-w}}{\lambda'(\sigma(w))}, \qquad \sigma''(w) = -\frac{e^{-w}}{\lambda'^3(\sigma(w))}\left(\lambda'^2(\sigma(w)) - e^{-w}\lambda''(\sigma(w))\right),$$

and finally the constants that appear in the main terms of the mean and the variance are

$$\mu = \frac{2}{|\lambda'(1)|}, \qquad \delta^2 = \frac{2}{|\lambda'^3(1)|}\left(\lambda''(1) - \lambda'^2(1)\right).$$

They are closely related to the first derivatives of the function $\Lambda(s) := \log \lambda(s)$ at $s = 1$ via the relations

$$\mu = \frac{2}{|\Lambda'(1)|}, \qquad \delta^2 = \frac{2\Lambda''(1)}{|\Lambda'^3(1)|}. \tag{4.20}$$

Cost of a truncated real trajectory. Before studying the general cost of a rational trajectory, we return, for a moment, to the cost of a truncated real trajectory and we prove Theorem 1. We recall that the moment generating function of the cost $C^{(n)}$ of a truncated trajectory of length n is given by (2.7), and admits an alternative expression as a function of the n-th iterate of the operator $\mathbf{H}_{1,w}$,

$$M_n(w) := \mathrm{E}[\exp(wC^{(n)})] = \int_{\mathcal{I}} \mathbf{H}^n_{1,w}[f](u)du. \tag{4.21}$$

Then Properties *UDE*, *SG*, $An_w(1,0)$, together with Perturbation Theory with respect to w, available when w is near 0, prove that there exists a (complex) neighborhood $\mathcal{W}_4$ of 0, and a real $\theta < 1$ such that, for any $w \in \mathcal{W}_4$, the operator $\mathbf{H}_{1,w}$ decomposes as

$$\mathbf{H}_{1,w} = \lambda(1,w)\,\mathbf{P}_{1,w} + \mathbf{N}_{1,w},$$

where $\mathbf{P}_{1,w}$ is the projector on the dominant eigensubspace, and $\mathbf{N}_{1,w}$ is an operator whose spectral radius equals $r_{1,w}$ which satisfies $|r_{1,w}| \le \theta|\lambda(1,w)|$, for some $\theta < 1$ (independent of w). Then, for any $n \ge 1$, the decomposition

$$\mathbf{H}^n_{1,w}[f](u) = \lambda(1,w)^n\,\mathbf{P}_{1,w}[f](u)du + \mathbf{N}^n_{1,w}[f](u)$$

entails a decomposition for $M_n(w)$

$$M_n(w) = \left(\lambda(1,w)^n \int_{\mathcal{I}} \mathbf{P}_{1,w}[f](u)du\right)(1 + O(\theta^n))$$

with a O–term which is uniform when w belongs to $\mathcal{W}_4$. This is a Quasi–Power decomposition, that fulfills the hypothesis of Quasi-Powers Theorem stated in Section 4.3, with $\beta_n := n$, $\kappa_n = \theta^n$, and

$$U(w) = \log \lambda(1, w), \qquad V(w) = \log\left(\int_{\mathcal{I}} \mathbf{P}_{1,w}[f](u)du\right).$$

In this case

$$U'(w) = \frac{\lambda'_w(1, w)}{\lambda(1, w)}, \qquad U''(w) = \frac{\lambda''_{w^2}(1, w) - \lambda'^2_w(1, w)}{\lambda^2(1, w)}$$

so that the constants that appear in the main terms of the mean and the variance are

$$\widehat{\mu}(c) = \lambda'_w(1, 0), \qquad \widehat{\delta}^2(c) = \lambda''_{w^2}(1, 0) - \lambda'^2_w(1, 0). \tag{4.22}$$

As we explain it in the Introduction, the proof of Theorem 1 only uses properties of the operator $\mathbf{H}_{s,w}$ when $s = 1$ and w is near 0. Then, it does not need precise estimates à la Dolgopyat. We prove Theorem 1 at this moment because it is an (easy) consequence of the Quasi-Powers Theorem and we need the expressions given in (4.22).

Cost of a rational trajectory. Let us now study the general case of a rational trajectory. Taking the derivative of the relation that implicitly defines σ, i.e., $\lambda(\sigma(w), w) = 1$, one obtains

$$\sigma'(w)\lambda'_s(\sigma(w), w) + \lambda'_w(\sigma(w), w) = 0,$$

$$\sigma''(w)\lambda'_s(\sigma(w), w) + \sigma'^2(w)\lambda''_{s^2}(\sigma(w), w) + 2\sigma'(w)\lambda''_{sw}(\sigma(w), w) + \lambda''_{w^2}(\sigma(w), w) = 0.$$

Remark that, when $w = 0$, the point $(\sigma(w), w)$ is $(1, 0)$ and the derivatives with respect to s satisfy

$$\lambda'_s(\sigma(w), w)|_{w=0} = \lambda'(1), \qquad \lambda''_{s^2}(\sigma(w), w)|_{w=0} = \lambda''(1).$$

Finally, the constants that appear in the main terms of the mean and the variance are

$$\mu(c) = \mu \cdot \widehat{\mu}(c) \tag{4.23}$$

$$\delta^2(c) = \widehat{\mu}(c)^2 \cdot \delta^2 + \mu \cdot \widehat{\delta}^2(c) + \mu^2 \nu(c), \tag{4.24}$$

and, thanks to Equations (4.20, 4.22), they are expressed as a function of $\mu, \delta, \widehat{\mu}(c), \widehat{\delta}(c)$. The quantity $\nu(c)$ is equal, by definition, to $-\lambda''_{sw}(1, 0)$; it is strictly positive and plays the rôle of a covariance coefficient between the number of iterations P and the cost C.

Condition $U''(0) \neq 0$. In the first two cases, the variance coefficient $U''(0)$ is closely related to the second derivative of functions related to the dominant eigenvalue $\lambda(s, w)$; in the first case, the function $\Lambda(s) := \log \lambda(s, 0)$ at $s = 1$; in the second case, function $w \mapsto \Lambda(1, w) := \log \lambda(1, w)$ at $w = 0$. The application of the Quasi-Powers Theorem needs that these derivatives are not zero. In fact, the function "dominant eigenvalue" is always log–convex with respect to each variable s or w. Furthermore, in the case when the dynamical systems has an infinite number of branches, and the cost is not constant, it is always strictly log–convex (see [45]).

Computation of constants of the mean. The constants that appear in the mean value admit closed expressions. Taking the derivative at $(s,w) = (1,0)$ of the relation $\mathbf{H}_{s,w}[f_{s,w}] = \lambda(s,w)f_{s,w}$ (with respect to s or to w) gives

$$\frac{d}{ds}\mathbf{H}_{1,0}[f_{1,0}] + \mathbf{H}_{1,0}[\frac{d}{ds}f_{1,0}] = \lambda'_s(1,0)f_{1,0} + \lambda(1,0)\frac{d}{ds}f_{1,0},$$

$$\frac{d}{dw}\mathbf{H}_{1,0}[f_{1,0}] + \mathbf{H}_{1,0}[\frac{d}{dw}f_{1,0}] = \lambda'_w(1,0)f_{1,0} + \lambda(1,0)\frac{d}{dw}f_{1,0},$$

In each case, by integration on $\mathcal{I}$, and using that $\mathbf{H}_{1,0} = \mathbf{H}_1$ is a density transformer, one obtains

$$\frac{2}{\mu} = |\lambda'(1)| = -\int_{\mathcal{I}} \mathbf{H}'_1[f_1](x)dx = \int_{\mathcal{I}} |\log u| \cdot f_1(u)du, \tag{4.25}$$

$$\widehat{\mu}(c) = \lambda'_w(1,0) = \int_{\mathcal{I}} \frac{d}{dw}\mathbf{H}_{1,0}[f_1](x)dx = \sum_{h\in\mathcal{H}} c(h)\cdot \int_{h(\mathcal{I})} f_1(u)du. \tag{4.26}$$

Since there exists a closed form for the invariant density f_1 in each of the three cases of interest (cf Figure 4), constants μ, $\widehat{\mu}(c)$ and $\mu(c)$ can be easily computed.
The situation is quite different for the second derivatives, and constants $\delta, \widehat{\delta}(c), \delta(c)$ do not seem to admit a closed form. However, Lhote has proved that these constants can be computed in polynomial time [32].

About Porter's constant. In the case of the Standard Euclidean Algorithm, the constant μ_1 that intervenes in the subdominant term of the mean of P is related to the so–called Porter's Constant κ [38] by the equality

$$\mu_1 = \kappa - \frac{1}{2} - \frac{1}{2\lambda'(1)}[\frac{\zeta'(2)}{\zeta(2)} - \frac{1}{2}],$$

and Porter's constant satisfies

$$\kappa - \frac{1}{2} = \frac{-1}{\lambda'(1)}[3\log 2 + 4\gamma - 4\frac{\zeta'(2)}{\zeta(2)} + \lambda'(1)].$$

Since $\mu_1 = V'(0)$ involves the derivative $A'(w)$ and thus the derivative of the operator $\mathbf{R}(w)$ defined in (4.2), the following equality relates the two derivatives $\lambda''(1)$ and $d'(1) := \mathbf{P}'_1[1](0)$.

$$2\log 2\ [d'(1) + \zeta(2)\lambda''(1)] = \frac{3}{2}\log 2 + 2\gamma - \frac{\zeta'(2)}{\zeta(2)} - \frac{\zeta(2)}{2\log 2}.$$

This relation between two constants that admit a priori no closed form seems new.

4.9. **Gaussian Asymptotic Distribution of costs on rational trajectories: a special application.** We wish to study now the length of the binary word that encodes the element (u,v) of Ω_N. There are two ways for coding this pair: the first one directly uses the binary encoding $A(u,v)$ of integer pair (u,v), while the second one directly uses the binary encoding $B(u,v)$ of the sequence of digits $(m_1, m_2, \ldots, m_p)$ produced by the Euclidean Algorithm. Many applications use CF expansions in an extensive way: for instance, it is possible to make computations on rational numbers by using directly their CF expansions [5, 50]. It is thus important to encode efficiently the sequence of CF–digits so that this encoding can be directly used in further computations, and to compare the average length of these two coding words. Classical results in Information Theory entail that the mean value of $B(u,v)$ is at least equal to the mean value of $A(u,v)$. This leads us to study the length $B(u,v)$ of the CF-encoding of a pair (u,v), related to some digit–cost b, which will be chosen to be near–optimal with respect to distribution of digits m_i. We now explain how a near–optimal choice of the encoding leads to use

$$B(u,v) := \sum_{i=1}^{P(u,v)} 2\ell(m_i) - 1, \qquad \text{with} \quad \ell(m) := \lceil \log_2 m \rceil + 1.$$

The choice of a near–optimal code will follow the principles of Fano–Shannon encoding that we recall now: Assume that a source produces (independent) symbols m_i with probability p_i and the symbols are sorted with respect to decreasing probabilities. One builds a binary tree according to the recursive principle: Denote by n_0 the first index index n such that $\sum_{i \le n} p_i \ge 1/2$. Then all the symbols m_i for $i \le n_0$ are encoding by a word that begins with $'0'$ while the symbols m_i for $i > n_0$ are encoding by a word that begins with $'1'$. Then, recursively divide the first group and the second group following the same principles.

We first adopt this Fano-Shannon principle to the case of the Standard Euclidean Algorithm, and we consider the approximation

$$p_m \approx \frac{1}{m} - \frac{1}{m+1} = \frac{1}{m(m+1)}.$$

Then all the digits m_i that comprise k binary digits have their code word that begins with $k-1$ times the symbol $'1'$ followed with a $'0'$. Then, we do not any more apply the Fano-Shannon process, and we use now the binary encoding of the digit. Finally, the code word of a digit m whose binary length is k is a sequence of $k-1$ symbols equal to $'1'$, then a zero, then the sequence of the binary digits of m where the most significant 1 is removed. For instance the digit 21 (in decimal expansion) is written as 10101 (in binary expansion), and will be encoded as 1111 0 0101. The process is quite easy to decode. For instance, the word 00011111100001000101 is decoded as $(1,1,1,68,1,3)$. We encounter there a code that was proposed by Elias [17] for encoding integers. We can also adopt the Fano-Shannon principle to the case of the other two Algorithms, even if the existence of the sign ϵ makes the process more involved.

In this way, for the Standard Algorithm, all the digits m of binary length k are encoded by a binary word of length $2k-1$, so that a near–optimal digit–cost is just $b := 2\ell - 1$, where ℓ is the binary length. This is thus an instance of a cost of moderate growth. Then, Theorem 3 applies to this case. Moreover, the constant $\mu(b)$ is easy to compute with (4.26) and we obtain the following:

The length of the Elias encoding of an integer pair of Ω_N via the standard continued fraction expansion follows asymptotically a Gaussian law, with a *mean that is asymptotic to*

$$\widehat{\mu}(b)\log_2 N, \qquad \textit{with} \qquad \widehat{\mu}(b) := \frac{12\log 2}{\pi^2}\left[-\log 2 + 2\log\prod_{k=0}^{\infty}\left(1+\frac{1}{2^k}\right)\right].$$

This encoding is near–optimal since this average value is asymptotic to $2.04868\log_2 N$*, whereas the average code-length relative to the usual encoding is* $2\log_2 N$*.*

5. Local Limit Theorem for the Number of iterations.

This Section is devoted to proving Theorem 4. We now focus on the basic parameter, the number P of steps, of any of the three Euclidean Algorithms, and we are interested in proving a Local Limit Theorem for this cost. We thus study probabilities of the form

$$\Pr_N[P(u,v)=p] = \frac{\sum_{n\le N} c_n^{(p)}}{|\Omega_N|} \qquad \text{with} \qquad c_n^{(p)} := |\{u, P(u,n)=p\}|.$$

5.1. **Outline of the proof for Theorem 4.** We recall that the Dirichlet generating function $S_p(s)$ of coefficients $c_n^{(p)}$ admits an alternative expression which involves the plain transfer operator $\mathbf{H}_s$ (see Equation (2.18))

$$S_p(2s) := \sum_{(u,v)\in\Omega} \frac{1}{v^{2s}}\,|\{P(u,v)=p\}| = \sum_{n\ge 1}\frac{c_n^{(p)}}{n^{2s}} = \mathbf{F}_s\circ\mathbf{H}_s^{p-1}[1](0),$$

and, by Perron's formula, the following holds, for any $D > 1/2$

$$\Psi^{(p)}(T) := \sum_{n\le T} c_n^{(p)}(T-n) = \frac{1}{2i\pi}\int_{D-i\infty}^{D+i\infty} \mathbf{F}_s\circ\mathbf{H}_s^{p-1}[1](0)\frac{T^{2s+1}}{s(2s+1)}ds.$$

Since we look for a local limit law, we consider the case where parametrers p and N are related; More precisely, we let

$$T := N, \qquad p = p(N) = \lfloor \mu\log N + x\delta\sqrt{\log N}\rfloor$$

where μ and δ are the constants that intervene in the dominant asymptotic terms of the mean and the variance of cost P (see (4.20)). Moreover, the worst–case behaviour of P shows that it is sufficient to consider the case where $p(N)$ is $O(\log N)$.
It will be convenient to use the vertical line $\Re(s)=1$ as integration contour. Finally, we deal with integrals of the form

$$(5.1) \qquad I_N := \Psi^{(p(N))}(N) = \frac{1}{2i\pi}\int_{1-i\infty}^{1+i\infty}\mathbf{F}_s\circ\mathbf{H}_s^{p(N)}[1](0)\frac{N^{2s+1}}{s(2s+1)}ds,$$

and we wish to describe their asymptotic behaviour for $N\to\infty$.

There are five main steps in the proof. The first four steps aim evaluating I_N, each of which replaces the integral of the previous step by its "dominant" part. Step(v) returns to the probability to be studied, i.e.,

$$\Pr_N[P(u,v) = \lfloor \mu\log N + x\delta\sqrt{\log N}\rfloor].$$

Step (i). With arguments à la Dolgopyat (given in Lemma 12), replace I_N by the integral I_N^- where the integration domain is only $[1-it_1, 1+it_1]$.

Step (ii). Thanks to dominant spectral properties near the real axis, replace in the integrand of I_N^- the operator $\mathbf{H}_s^{p(N)}$ by its "dominant" part $D(s)\lambda(s)^{p(N)}$ and obtain a new integral denoted by J_N.

Step (iii). With saddle point principles (see [10]), replace in integral J_N the integration domain by a segment $[1-id_n, 1+id_n]$, with d_n of order $\sqrt{\log n/n}$ and $n = \log N$, and obtain a new integral denoted by K_N.

Step (iv). Study K_N and obtain an asymptotic estimate for K_N, and thus for $I_N = \Psi^{(p(N))}(N)$.

Step (v). Transfer this estimate into an estimate for $\Phi^{(p(N))}(N)$, with Part (c) of Lemma 10. This quantity $\Phi^{(p(N))}(N)$ is closely related to the probability to be studied.

5.2. **Proof of Theorem 4.** We describe now each step.

Step (i). Perturbation Theory proves the existence of a real t_1, strictly positive, such that on the segment $[1-it_1, 1+it_1]$, the operator $\mathbf{H}_s$ admits a spectral decomposition of the form

$$\mathbf{H}_s^p = \lambda(s)^p \mathbf{P}_s + \mathbf{N}_s^p, \tag{5.2}$$

where the spectral radius of $\mathbf{N}_s$, i.e., the subdominant spectral radius of $\mathbf{H}_s$ is less than r (for some $r < 1$).

Then, the integral I_N is split into two parts; the first one, I_N^+, corresponds to the part of the line with $|t| \geq t_1$, and the second one, I_N^-, corresponds to the part of the line with $|t| \leq t_1$.

Again, Dolgopyat estimates play a fundamental rôle for dealing with I_N^+. Theorem 2 together with Lemma 1 show the following:

Lemma 12. *For any $t_1 > 0$, there exists $M_2 > 0$, $\gamma_2 < 1$ and ϵ, with $0 < \epsilon < 1$ such that, for any integer p and any s with $s = 1 + it, |t| \geq t_1$, one has*

$$|S_p(2s)| \leq M_2 \cdot \gamma_2^p \cdot \sup(1, |t|^\epsilon).$$

Proof. Fix some $\epsilon, 0 < \epsilon < 1$, and consider the parameters M, t_0, γ of Theorem 2. Then, when $s = 1 + it$, with $|t| \geq t_0$, one has

$$|S_p(2s)| = |\mathbf{F}_s \mathbf{H}_s^{p-1}[1](0)| \leq ||\mathbf{F}_s \mathbf{H}_s^{p-1}||_{1,t} \leq M\gamma^p |t|^\epsilon.$$

On the otherside, on any segment $t_1 \leq |t| \leq t_0$, Lemma 1 shows that there exists $M_1, \gamma_1 < 1$ such that

$$|\mathbf{F}_s \mathbf{H}_s^{p-1}[1](0)| \leq M_1 \gamma_1^p.$$

Finally, the result is obtained with $M_2 := \sup(M, M_1), \gamma_2 := \sup(\gamma, \gamma_1)$. $\square$

This result provides an upper bound for I_N^+,

$$|I_N^+| \leq 2M_2\, N^3\, \gamma_2^{p(N)} \int_{t_1}^{+\infty} \frac{\sup(1, t^\epsilon)}{t(2t+1)} dt.$$

Since ϵ satisfies $0 < \epsilon < 1$, the integral exists. Moreover, if x is $O(\sqrt{\log N})$, then $p(N)$ is $\Theta(\log N)$ and I_N^+ is $N^3\, O(N^{-\alpha_1})$, with some positive α_1, so that $I_N = I_N^- + N^3\, O(N^{-\alpha_1})$

Step (ii). The integral I_N^- is itself split into two parts according to the spectral decomposition (5.2), the part that involves the dominant part of the integrand $\lambda(s)^p \mathbf{P}_s$, and the part that involves the remainder part of the integrand $\mathbf{N}_s^p$, whose norm is less than r^p; this last part gives rise to a term which is $N^3 O(r^{p(N)})$. If x is $O(\sqrt{\log N})$, then $p(N)$ is $\Theta(\log N)$ and this term is $N^3 O(N^{-\alpha_2})$ for some strictly positive α_2. Let $\alpha := \min(\alpha_1, \alpha_2)$. Then, the integral I_N to be studied is

(5.3)

$$I_N = J_N + N^3 O(N^{-\alpha}) \qquad \text{with} \qquad J_N = \frac{1}{2i\pi}\int_{1-it_1}^{1+it_1} \lambda(s)^{p(N)} N^{2s+1} D(s) ds.$$

Here, $D(s)$ denotes a function that does not depend on N,

$$D(s) := \frac{1}{s(2s+1)} \mathbf{F}_s \circ \mathbf{P}_s[1](0). \tag{5.4}$$

Step (iii). We let $n := \log N$. It remains to study the integral J_N which will provide the main term of the asymptotics, and we use the decomposition

$$\lambda(s)^{p(N)} N^{2s+1} = \exp[(\mu\Lambda(s) + (2s+1))n] \cdot \exp[\delta x \Lambda(s)\sqrt{n}], \tag{5.5}$$

that deals with the quite important function $\Lambda(s) := \log \lambda(s)$. It satisfies at $s = 1$ (see Equation (4.20))

$$\Lambda(1) = 0, \qquad \Lambda'(1) = -\frac{2}{\mu}, \qquad \Lambda''(1) = 4\frac{\delta^2}{\mu^3}. \tag{5.6}$$

The first factor of (5.5) involves the function

$$B(s) := \mu\Lambda(s) + 2s + 1 \tag{5.7}$$

which satisfies, thanks to (5.6) the relations $B(1) = 3$, $B'(1) = 0$. Then the point $s = 1$ is a saddle point for $B(s)$, and principles of the Saddle Point Method apply here (see [10]); we consider some sequence d_n that satisfy $nd_n^2 \to \infty$ and $nd_n^3 \to 0$; we let $d_n = a_n/\sqrt{n}$ where a_n tends to ∞ and is $o(n^{1/6})$. Then, the integral K_N

$$K_N = \frac{D(1)}{2i\pi} \int_{1-id_n}^{1+id_n} \lambda(s)^{p(N)} N^{2s+1} ds,$$

will provide the main term of the asymptotics and

$$J_N = K_N + N^3\, O\left(\frac{a_n}{\sqrt{n}} \exp[-K\, a_n^2]\right) \tag{5.8}$$

for some $K > 0$.

Step (iv). With the change of variables $s := 1 + iw/\sqrt{n}$, one obtains for function B defined in (5.7)

$$n[B(s) - B(1)] = n\left[-\frac{w^2}{2n}B''(1) + O\left(\frac{w^3}{n^{3/2}}\right)\right] = -\frac{w^2}{2} 4\frac{\delta^2}{\mu^2} + O\left(\frac{w^3}{\sqrt{n}}\right),$$

and for function Λ,

$$\sqrt{n}\delta x \Lambda(s) = \sqrt{n}\delta x \left[\frac{iw}{\sqrt{n}}\Lambda'(1) + O\left(\frac{w^2}{n}\right)\right] = -2iwx\frac{\delta}{\mu} + x\, O\left(\frac{w^2}{\sqrt{n}}\right).$$

Suppose now that $|x|$ is $O(a_n)$. Then,

$$K_N = \left[1 + O\left(\frac{a_n^3}{\sqrt{n}}\right)\right] \frac{e^{nB(1)}D(1)}{2\pi\sqrt{n}} L_N \tag{5.9}$$

with

$$L_N := \int_{-a_n}^{+a_n} \exp[-2w^2\frac{\delta^2}{\mu^2} - 2iwx\frac{\delta}{\mu}]dw.$$

Now, by "completing" the square

$$2w^2\frac{\delta^2}{\mu^2} + 2iwx\frac{\delta}{\mu} = \frac{z^2}{2} + \frac{x^2}{2}, \qquad \text{with} \qquad z := 2w\frac{\delta}{\mu} + ix,$$

and using the new variable z, the integral L_N can be written as

$$L_N = e^{-x^2/2}\ \frac{\mu}{2\delta}\ \int_{S_n} e^{-z^2/2}dz,$$

where S_n is the horizontal segment

$$S_n := [-\frac{2\delta}{\mu}a_n + ix, +\frac{2\delta}{\mu}a_n + ix]$$

If we now restrict $|x|$ to be less than $a_n(\delta/\mu)$, each end point of the segment S_n has a square whose real part is strictly positive, and

$$L_N = \sqrt{2\pi}\, e^{-x^2/2}\, \frac{\mu}{2\delta}\left[1 + O\left(a_n \exp[-Ka_n^2]\right)\right]. \tag{5.10}$$

We choose now $a_n := \Theta(\log n)^{1/2}$, so that the two remainder terms in (5.9) and (5.10) are of similar order. Then,

$$K_N = \left[1 + O\left(\frac{\log^{3/2} n}{\sqrt{n}}\right)\right] \frac{e^{nB(1)}D(1)e^{-x^2/2}}{\sqrt{2\pi n}}\frac{\mu}{2\delta}.$$

We return now to $N = e^n$. If x satisfies $|x| \le \theta(\log\log N)^{1/2}$ with θ sufficiently small, one has

$$K_N = \left[1 + O\left(\frac{(\log\log N)^{3/2}}{(\log N)^{1/2}}\right)\right] \frac{N^3 D(1)}{\sqrt{\log N}}\frac{e^{-x^2/2}}{\sqrt{2\pi}}\frac{\mu}{2\delta}. \tag{5.11}$$

With (5.11, 5.8, 5.3), it is easy to return first to J_N, second to I_N. Finally,

$$\Psi^{(p(N))}(N) = \left[1 + O\left(\frac{(\log\log N)^{3/2}}{(\log N)^{1/2}}\right)\right] \frac{N^3 D(1)}{\sqrt{\log N}}\frac{e^{-x^2/2}}{\sqrt{2\pi}}\frac{\mu}{2\delta}.$$

Step (v)**.** We wish to transfer this estimate to $\Phi^{(p(N))}(N)$. We make precise the dependence of function $p(T)$ and we let

$$p_x(T) := \lfloor \mu \log T + x\delta\sqrt{\log T}\rfloor.$$

Now, with definition of Φ and ideas of Lemma 10, one has

$$\Phi^{(p(T))}(T) \le \frac{1}{\epsilon}\int_T^{T+\epsilon} \Phi^{(p(T))}(t)dt = \frac{1}{\epsilon}[\Psi^{(p(T))}(T+\epsilon) - \Psi^{(p(T))}(T)].$$

The right member decomposes as a sum of two terms,

(5.12)

$$\frac{1}{\epsilon}[\Psi^{(p(T))}(T+\epsilon) - \Psi^{(p(T+\epsilon))}(T+\epsilon)] + \frac{1}{\epsilon}[\Psi^{(p(T+\epsilon))}(T+\epsilon) - \Psi^{(p(T))}(T)].$$

First, Part (c) of Lemma 10 is applied to the second term of (5.12) with

$$G(T) = \frac{(\log\log T)^{3/2}}{(\log T)^{1/2}}, \qquad \epsilon(T) = G(T)^{1/2}.$$

Consider now the first term of (5.12). With the previous choice of $\epsilon(T)$, we remark that $p_x(T + T\epsilon(T)) = p_y(T)$ with

$$y = x\left[1 + O\left(\frac{\log(1 + G(T)^{1/2})}{\log T}\right)\right] = x\left[1 + O\left(\frac{G(T)^{1/2}}{\log T}\right)\right],$$

so that

$$\frac{1}{TG(T)^{1/2}}[e^{-x^2/2} - e^{-y^2/2}] = xe^{-x^2/2}\, O\left(\frac{1}{T\log T}\right).$$

Since

$$\frac{|x|}{\log T} = O\left(\frac{(\log\log T)^{1/2}}{\log T}\right) = o\left(\frac{(\log\log T)^{3/4}}{(\log T)^{1/4}}\right),$$

the first term of (5.12) is of smaller order than the remainder term of the second term. Finally, we obtain,

$$\Phi^{(p_x(N))}(N) = \left[1 + O\left(\frac{(\log\log N)^{3/4}}{(\log N)^{1/4}}\right)\right] \frac{3N^2 D(1)}{\sqrt{\log N}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \frac{\mu}{2\delta},$$

uniformly for $|x| \le \theta(\log\log N)^{1/2}$ with θ sufficiently small.

Recall that $\Phi^{(p(N))}(N)$ is closely related to the probability to be studied, via the equality

$$\mathrm{Pr}_N[P = \lfloor \mu \log N + x\delta\sqrt{\log N}\rfloor] = \frac{\Phi^{(p(N))}(N)}{|\Omega_N|}.$$

Then, a comparison between (5.4) and (4.12) and a division by $|\Omega_N|$,

$$|\Omega_N| = C_N(0) = 3D(1)\frac{\mu}{2}N^2\left[1 + O\left(\frac{1}{N}\right)^{\alpha}\right],$$

lead to the final result

$$\mathrm{Pr}_N[P = \lfloor \mu \log N + x\delta\sqrt{\log N}\rfloor] = [1 + O(\epsilon_N)]\ \frac{e^{-x^2/2}}{\delta\sqrt{2\pi\log N}}, \tag{5.13}$$

with

$$\epsilon_N = \frac{(\log\log N)^{3/4}}{(\log N)^{1/4}}. \tag{5.14}$$

This ends the proof of Theorem 4. □

Comparing with Hensley's results. We obtain a faster speed of convergence, since ours is $O(1/(\log N)^{(1/4)-\epsilon})$ while the speed of convergence obtained by Hensley is $O(1/(\log N)^{1/24})$.

6. Conclusion and Open Problems.

The paper deals with the distribution of costs related to three Euclidean Algorithms and proves that, for costs of moderate growth, the asymptotic distribution is Gaussian. Moreover, the method which we use —distributional dynamical analysis— appears to provide a general framework, where it seems possible to solve other problems in the field : for instance, what happens for other Euclidean Algorithms of the Fast Class? for Euclidean Algorithms of the Slow Class? for a cost of fast growth? for a cost that involves not only the digits (as it is the case when we are interested in the bit complexity of Euclidean Algorithms)?

Acknowledgements. Hervé Daudé considered the problem during his thesis, and made experiments that show the evidence of the Gaussian limit for the number of steps. With Eda Cesaratto, we worked a lot with the weighted transfer operator, and some of the ideas that we shared at that moment were very useful for this paper. Finally, we have many interesting discussions with Philippe Flajolet about the saddle point method and the notion of smoothed costs.
This work has been supported in part by the IST Programme of the EU under contract number IST-1999-14186 (Alcom-FT).

References

[1] Akhavi, A., Vallée, B. Average bit–complexity of Euclidean Algorithms, Proceedings of ICALP'00, *Lecture Notes in Computer Science* 1853, pp 373- 387, Springer.

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

[3] Baladi, V. *Positive Transfer operators and decay of correlations,* Advanced Series in non linear dynamics, vol 16, World Scientific (2000).

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

[5] Beeler, M., Gosper, R. W., and Schroeppel, R. HAKMEM. Memorandum 239, M.I.T., Artificial Intelligence Laboratory, Feb. 1972.

[6] Bourdon, J. Daireaux, B. Vallée, B. Dynamical analysis of α-Euclidean algorithms, Journal of Algorithms,44 (2002) pp 246–285.

[7] Boyarsky, A. and Gora, P. *Laws of Chaos, Invariant measures and dynamical systems in one dimension*, Probability and its applications, Birkhauser (1997).

[8] 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.

[9] Broise, A. Transformations dilatantes de l'intervalle et théorèmes limites, *Asterisque* 238, pp 5–109, Société Mathématique de France (1996).

[10] de Bruijn, N.G. *Asymptotic methods in Analysis,* North-Holland, 1958

[11] Collet, P. Some ergodic properties of maps of the interval, *Dynamical systems, Proceedings of the first UNESCO CIMPA School on Dynamical and Disordered Systems* (Temuco, Chile, 1991), Hermann (1996).

[12] Curtiss, J. H. A note on the theory of moment generating functions, *Annals of Mathematical Statistics*, 13 (1942), pp. 430–433.

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

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

[15] Dolgopyat, D. On decay of correlations in Anosov flows, Ann. of Math. vol 147, pp 357–390 (1998)

[16] Dolgopyat, D. Limit Theorems for partially hyperbolic systems, To appear in Transactions of AMS

[17] Elias, P. Universal codeword sets and representations of the integers, *IEEE Transactions on Information Theory.* Vol IT-21 , No 2, March 1975, pp 194-203.

[18] Ellison, W. and Ellison, F. *Prime Numbers,* Hermann, Paris 1985.

[19] 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.

[20] Flajolet, P. Singular Combinatorics. In *Proceedings of the International Congress of Mathematicians* vol III, Li Tatsien Ed., World Scientific (2002), pp. 561–571.

[21] Flajolet, P. and Sedgewick, R. Analytic Combinatorics, Book in preparation, see also the *Rapports de Recherche INRIA* 1888, 2026, 2376, 2956.

[22] 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.

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

[24] Hwang, H-K. Théorèmes limite pour les structures combinatoires et les fonctions arithmétiques, PhD thesis, Ecole Polytechnique, Dec. 1994.

[25] H.-K. Hwang Large deviations for combinatorial distributions: I. Central limit theorems, The Annals of Applied Probability vol 6, pp 297–319, (1996)

[26] H.-K. Hwang, On convergence rates in the central limit theorems for combinatorial structures, European Journal of Combinatorics, vol 19, pp 329–343 (1998)

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

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

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

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

[31] Lasota, A. and Mackey, M. *Chaos, Fractals and Noise; Stochastic Aspects of Dynamics,* Applied Mathematical Science 97, Springer (1994).

[32] Lhote, L. Modélisation et approximation de sources complexes, Mémoire de DEA, Université of Caen, 2002.

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

[34] 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.

[35] Pollicott, M. Lectures on decay of correlations for flows, Unpublished lecture notes (1997)

[36] Pollicott, M. On the mixing of Axiom A attracting flows and a conjecture of Ruelle, Ergodic Theory Dynam. Systems, vol 19 pp 535–548 (1999)

[37] Pollicott, M. and Sharp, R. Exponential error terms for growth functions on negatively curved surfaces, Amer. J. Math., vol 120, pp 1019–1042, (1998).

[38] Porter, J.W. On a theorem of Heilbronn, Mathematika 22 (1975) pp 20–28.

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

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

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

[42] Schweiger, F. *Ergodic Theory of fibred systems and Metric Number Theory,* Oxford Science Publications, Clarendon Press, Oxford (1995)

[43] STEIN, E. *Harmonic analysis: real variable methods, orthogonality and Oscillatory integrals,* Princeton university Press, Princeton (1993)

[44] TENENBAUM, G. *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.

[45] VALLÉE, B. Dynamical sources in information theory: fundamental intervals and word prefixes, *Algorithmica* (2001) vol. 29, pp 262–306.

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

[47] VALLÉE, B. Algorithms for computing signs of 2×2 determinants: dynamics and average–case analysis, Congrès ESA'97 (5th Annual European Symposium on Algorithms) (Graz, Septembre 97), *Lecture Notes in Computer Science* 1284, pp 486–499.

[48] VALLÉE, B. Dynamical Analysis of a Class of Euclidean Algorithms, to appear in *Theoretical Computer Science* (2003).

[49] VALLÉE, B. Digits and Continuants in Euclidean Algorithms. Ergodic Versus Tauberian Theorems, *Journal de Théorie des Nombres de Bordeaux* 12 (2000) pp 531-570.

[50] VUILLEMIN, J. Exact real computer arithmetic with continued fractions. *IEEE Transactions on Computers 39*, 8 (Aug. 1990), 1087–1105.

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

VIVIANE BALADI : CNRS UMR 7586, INSTITUT MATHÉMATIQUES DE JUSSIEU, 75251 PARIS, (FRANCE), BRIGITTE VALLÉE: CNRS UMR 6072, GREYC, UNIVERSITÉ DE CAEN, F-14032 CAEN (FRANCE).