EUCLIDEAN ALGORITHMS ARE GAUSSIAN

VIVIANE BALADI AND BRIGITTE VALLÉE

Abstract. We obtain a Central Limit Theorem for a general class of additive parameters (costs, observables) associated to three standard Euclidean algorithms, with optimal speed of convergence. We also provide very precise asymptotic estimates and error terms for the mean and variance of such parameters. For costs that are lattice (including the number of steps), we go further and establish a Local Limit Theorem, with optimal speed of convergence. We view an algorithm as a dynamical system restricted to rational inputs, and combine tools imported from dynamics, such as transfer operators, with various other techniques: Dirichlet series, Perron's formula, quasi-powers theorems, and the saddle-point method. Such dynamical analyses had previously been used to perform the average-case analysis of algorithms. For the present (dynamical) analysis in distribution, we require estimates on transfer operators when a parameter varies along vertical lines in the complex plane. To prove them, we adapt techniques introduced recently by Dolgopyat in the context of continuous-time dynamics [16].

1. Introduction

According to Knuth [30, 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." Indeed, Euclid's algorithm is currently a basic building block of computer algebra systems and multi-precision arithmetic libraries, and, in many such applications, most of the time is spent in computing gcd's. However, the Euclidean algorithm has not yet been completely analyzed, and it is the purpose of this paper to provide such an analysis.

We shall state informally our Central Limit Theorem (CLT) and Local Limit Theorem (LLT) later in this section, but first we discuss Euclidean algorithms and cost functions (equivalently: additive parameters), recalling previously known facts about their average-case and distributional analyses. Our results have been announced in [5].

Continued fraction expansions of real numbers. Every $x \in]0,1]$ admits a finite or infinite (CF)-continued fraction expansion of the form

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_n + \dots}}}} . \tag{1.1}$$

Date: October 2003.

Ordinary continued fraction expansions can be viewed as trajectories of a one-dimensional dynamical system, the Gauss map $T : [0,1] \to [0,1]$,

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

(Here, $\lfloor x \rfloor$ is the integer part of x.) For an irrational x, the trajectory $\mathcal{T}(x) = (x, T(x), T^2(x), \ldots, T^n(x), \ldots)$ never meets 0 and is encoded by the infinite sequence of *digits* $(m_1(x), m_2(x), m_3(x), \ldots, m_n(x), \ldots)$, defined by $m_i(x) := m(T^{i-1}(x))$, with $m(x) := \lfloor \frac{1}{x} \rfloor$. It is usual to consider the truncated trajectory $\mathcal{T}_n(x) := (x, T(x), \ldots, T^n(x))$, and let n tend to ∞. If $x \neq 0$ is rational, the trajectory $\mathcal{T}(x)$ reaches 0 in a finite number of steps, and this number, $P(x)$, is called the depth of x. We set $P(x) = \infty$ for irrational x.

Associate a nonnegative real value $c(m)$ to each possible digit $m \geq 1$. We may also regard c as a function on the set $\mathcal{H}$ of inverse branches of T. We set for each x and each n

$$C_n(x) := \sum_{i=1}^{\min(n,\, P(x))} c(m_i(x)) \,. \tag{1.2}$$

We shall refer to c as a *digit-cost* and to C_n as the associated *total cost* (sometimes, just *cost*), since a variety of costs, in the usual sense of computational complexity, are usually expressible in this way.

The total cost (1.2) is a Birkhoff sum, i.e., a sum over iterates of the dynamics T. Hence, the functions $c(m)$ play the rôle of "observables" in physics, and it is of interest to characterize their probabilistic behavior. Here, T admits a unique invariant density $f_1(x) = (1/\log 2)(1+x)^{-1}$, the Gauss density. When the integrability condition,

$$\widehat{\mu}(c) := \sum_{h \in \mathcal{H}} c(h) \cdot \int_{h([0,1])} f_1(x) dx = \int_0^1 c(m(y))\, f_1(y) dy < \infty \,, \tag{1.3}$$

is satisfied, then the Birkhoff ergodic theorem implies that for Lebesgue almost every x in $[0,1]$ the average $C_n(x)/n$ converges to $\widehat{\mu}(c)$ as $n \to \infty$. We discuss next how *transfer operators* give refinements of this fact.

The density transformer, also known as the Perron-Frobenius operator,

$$\mathbf{H}_1[f](x) = \sum_{h \in \mathcal{H}} |h'(x)| \cdot f \circ h(x) \tag{1.4}$$

was introduced early in the study of continued fractions (see e.g. Lévy [33], Khinchin [29], Kuzmin [31], Wirsing [50], Babenko [3], and Mayer [34]). The density transformer is a special case of a *transfer operator.* Very general transfer operators were introduced by Ruelle, in connection with his thermodynamic formalism (see e.g. [40]). We shall see that "weighted" transfer operators appear naturally in the probabilistic analysis of dynamics.

Fix a reference probability measure on $[0,1]$, absolutely continuous with a smooth density f, and denote by $\mathbb{E}[\cdot]$ the corresponding expectation. To establish probabilistic results on total costs along truncated trajectories as $n \to \infty$, it is standard to use the sequence of *moment generating functions,* i.e., the sequence of expectations $\mathbb{E}[\exp(wC_n)]$, for w complex. In probabilistic i.i.d. situations, $\mathbb{E}[\exp(wC_n)]$ is the n-th power of some expectation. In our setting, a *quasi-powers* approximation (with remainder term) can be obtained for $\mathbb{E}[\exp(wC_n)]$ after expressing it in terms of the nth iterate $\mathbf{H}_{1,w}^n$ of a transfer operator. The transfer operator in question is the following perturbation of the density transformer:

$$\mathbf{H}_{1,w}[f] = \sum_{h \in h} \exp[wc(h)] \cdot |h'| \cdot f \circ h \,. \tag{1.5}$$

The density transformer $\mathbf{H}_1 = \mathbf{H}_{1,0}$ (acting on $\mathcal{C}^1$ functions) has a dominant eigenvalue $\lambda = 1$, and a *spectral gap:* the rest of the spectrum lies in a disk of radius < 1. For a non constant digit-cost satisfying a moderate growth condition (2.5), elementary *perturbation theory* [28] implies that $\mathbf{H}_{1,w}$ inherits the spectral gap when w is near 0. This gives the above-mentioned quasi-powers expansion and, together with convexity of the logarithm of the dominant eigenvalue, leads to a proof that the asymptotic distribution of the total cost is Gaussian, with (optimal) speed of convergence $O(1/\sqrt{n})$. This Central Limit Theorem [stated below more precisely as **Theorem 1**] is quite well-known. See for instance [12] or [9] for interval maps, and [1] for a more abstract framework and references to the pioneering paper of Nagaev. It is convenient to base the proof below on a compact and versatile statement, the "Quasi-Powers Theorem" of Hwang [25, 26, 27] [**Theorem 0** below], which encapsulates the consequences of the Lévy continuity theorem and the Berry-Esseen inequality.

Continued fractions of rational numbers: Average-case analysis of Euclidean algorithms. There are variants of the standard continued fraction algorithm induced by variations of the standard division procedure. See [48] for many examples and a classification into the "Fast" and the "Slow" Class. In this paper, we study three algorithms of the Fast Class, the *standard, centered,* and *odd* algorithms, specified at the beginning of Section 2. An execution of a Euclidean algorithm on the input (u, v) formed with two integers u, v such that $u/v = x$ gives rise to a rational trajectory $\mathcal{T}(x)$ which ends at zero in $P(x)$ steps, and the total cost of the trajectory is then

$$C(x) := \sum_{i=1}^{P(x)} c(m_i(x)) \,. \tag{1.6}$$

The reference parameter is no longer the truncation degree n, but the size $N := \max(u, v)$ of the input (u, v). The reference probability measure $\mathbb{P}_N$ is now the uniform discrete measure on the (finite) set of inputs of size $\le N$.

As it is often the case, the discrete problem is more difficult than its continuous counterpart. However, the average-case complexity of Euclidean algorithms is already well-understood, as we explain next. For the number of steps $P(u, v)$, which corresponds to the trivial digit-cost $c \equiv 1$, the standard Euclidean algorithm was first analyzed in the average-case around 1969, independently by Heilbronn [22] and Dixon [15]. The centered algorithm was studied by Rieger [38].

Consider now a general digit-cost c of moderate growth and the associated total cost C of the rational trajectories. The expectation $\mathbb{E}_N[C]$ is described by the partial sums of the coefficients of a generating function $S(s)$ (a common tool in the average-case study of algorithms [19, 20]) where the parameter s "marks" the size N of inputs. As it is usual in number theory, the generating functions $S(s)$ are Dirichlet series. Recently, Vallée [48] has related $S(2s)$ to the *quasi-inverse* $(I - \mathbf{H}_s)^{-1}$ of another perturbation $\mathbf{H}_s$ of the density transformer, together with its weighted version $\mathbf{H}_s^{(c)}$:

$$\mathbf{H}_s[f] = \sum_{h \in \mathcal{H}} |h'|^s \cdot f \circ h \, , \qquad \mathbf{H}_s^{(c)} := \sum_{h \in \mathcal{H}} c(h) \cdot |h'|^s \cdot f \circ h. \tag{1.7}$$

Then, spectral information on $\mathbf{H}_s$ may be used to show that $(I - \mathbf{H}_s)^{-1}$ is analytic in the half-plane $\{\Re s > 1\}$, and analytic on $\Re s = 1$ except for a simple pole at $s = 1$. Under these conditions, one can extract asymptotically the coefficients of $S(s)$ by means of Delange's Tauberian theorems [14, 43]. (Just like in the usual prime number theorem, or in weighted dynamical prime number theorems, see e.g. Sections 6–7 of [36].) For costs of moderate growth and Euclidean algorithms in the Fast Class, this dynamical approach gives [48] that the mean value $\mathbb{E}_N[C]$ of the total cost of the rational trajectory satisfies $\mathbb{E}_N[C] \sim \widehat{\mu}(c) \cdot \mu \log N$. Here, $\widehat{\mu}(c)$ is the asymptotic mean value (1.3) of truncated real trajectories, and μ equals $2/|\lambda'(1)|$, where $\lambda(s)$ is the dominating eigenvalue of $\mathbf{H}_s$.

Euclidean algorithms: Distributional analysis and dynamical methods. Main results. We have seen that, with respect to any cost of moderate growth, rational trajectories behave in 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 the total 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.

Concerning the standard Euclidean algorithm and the number of steps (i.e., the constant cost $c \equiv 1$), Hensley [24] has obtained a Central Limit Theorem, and a Local Limit Theorem with speed of convergence $O((\log N)^{-1/24})$. In the present work, we apply dynamical methods for the first time to the distributional analysis of discrete algorithms; in this way, we improve Hensley's result while extending it to a large class of cost functionals and to several algorithms.

Our strategy consists in describing the moment generating function $\mathbb{E}_N[\exp(wC)]$ as a quasi-power. It turns out that $\mathbb{E}_N[\exp(wC)]$ is related to the partial sums of the coefficients in a bivariate series $S(s, w)$. This series is of Dirichlet type with respect to the variable s, while the extra parameter w "marks" the cost c, and we require uniform estimates with respect to w. Tauberian theorems are now insufficient, since they do not provide remainder terms: We need a more precise "extractor" of coefficients, and the Perron formula is well-suited to this purpose.

We establish below (2.17) a simple relation between $S(2s,w)$ and the quasi-inverse of a two-variable transfer operator $\mathbf{H}_{s,w}$ defined by

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

Note that this operator is a simultaneous extension of the three operators $\mathbf{H}_1, \mathbf{H}_{1,w}, \mathbf{H}_s$ defined in (1.4), (1.7), and (1.5), while $\mathbf{H}_s^{(c)}$ is just its derivative with respect to w at $w = 0$. In order to apply Perron's formula with a convenient integration contour, it is thus useful to know that, in a half plane "perforated" at $s = 1$, of the form $\{\Re s \ge 1-\epsilon\,, |s-1| > \epsilon/2\}$ for small $\epsilon > 0$, a certain norm of the quasi-inverse satisfies

$$\|(I - \mathbf{H}_{s,w})^{-1}\|_s \le \max(1, |\Im s|^{\xi})\,, \tag{1.8}$$

with $\xi < 1$, uniformly in w. Then, spectral properties of $\mathbf{H}_{s,w}$ inherited from $\mathbf{H}_1$ give the desired quasi-power expansion for $\overline{\mathbb{E}}_N[\exp(wC)]$, for w close to 0, and some $\overline{\mathbb{P}}_N$ close to $\mathbb{P}_N$. Note that Hensley in [24] has used a transfer operator $\mathbf{H}_{s,0}$, but in an appreciably different way: Hensley obtains distributional results on rational trajectories upon aproximating discrete measures on rationals by continuous measures. In particular, his approach avoids parameters s of large imaginary parts.

Adapting powerful methods due to Dolgopyat [16], we show [**Theorem 2** and Lemma 6 below] that the quasi-inverse satisfies the estimates (1.8) for large $|\Im s|$. Dolgopyat was interested in the decay of correlations for hyperbolic flows satisfying some uniform nonintegrability condition (*UNI*). Later on, Pollicott and Sharp used Dolgopyat's bounds together with Perron's formula to find error terms in asymptotic estimates for geodesic flows on surfaces of variable negative curvature; see e.g. [37], where only univariate Dirichlet series with positive cofficients appear. To the best of our knowledge, the present paper is the first instance where these powerful tools are extended to dynamical systems with infinitely many branches and applied to distributional analyses in discrete combinatorics.

Let us now state informally our two main results about the three algorithms:

Theorem 3. Consider a nonzero digit-cost c of moderate growth. We show the following Central Limit Theorem: the asymptotic distribution of the total cost $C(u/v)$ of an execution of the algorithm on the rational input u/v, uniformly randomly drawn from $\{(u,v)\,, u \le v \le N\}$, is asymptotically Gaussian, with best possible speed of convergence, of order $O(1/\sqrt{\log N})$. We give expansions for the expectation and variance, which are asymptotically proportional to $\log N$. The constants $\mu(c)$ and $\delta^2(c)$ in the main terms of the expectation and the variance are expressed in function of the partial derivatives at $(s,w) = (1,0)$ of the dominant eigenvalue of $\mathbf{H}_{s,w}$, and alternatively in terms of $\mu = \mu(1)$, $\delta^2 = \delta^2(1)$, and the constants $\widehat{\mu}(c)$, $\widehat{\delta}(c)$ from Theorem 1. In particular $\mu(c)$ admits a closed form.

Theorem 4. A digit-cost is *lattice* if it is non zero and there exists $L \in \mathbb{R}_*^+$ so that c/L is integer-valued. The largest such L is then called the *span* of c. For instance, any nonzero constant c is lattice. For lattice costs of moderate growth, we obtain a Local Limit Theorem with optimal speed of convergence $O(1/\sqrt{\log N})$. This time, we use estimates for $\overline{\mathbb{E}}_N[\exp(i\tau C)]$, where τ varies in a compact set of the real line. They lead, with the saddle-point method, to a very natural and concise proof.

Three special instances of our results for lattice costs are of major interest.
(i) *Digit-cost* $c \equiv 1$. For each of our three algorithms the number of steps is asymptotically Gaussian, with mean $\mu \log N$ (where μ admits a closed form) and variance $\delta^2 \log N$, in the sense of the CLT and the LLT with speed $O(1/\sqrt{\log N})$.
(ii) *Digit-cost* $c = c_m$ = *the characteristic function of a digit* m. The total number of occurrences of a fixed digit m in a rational trajectory for our three algorithms is asymptotically Gaussian (CLT, LLT), with mean $\mu \cdot \widehat{\mu}(c_m) \log N$. The constant $\widehat{\mu}(c_m)$ is explicit for the three algorithms. In the case of the standard Euclidean algorithm we recover (see [48])

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

(iii) *Digit-cost* c = *the binary length* ℓ *of the digit.* The binary encoding of a rational trajectory is asymptotically Gaussian (CLT, LLT), with mean-value $\mu \cdot \widehat{\mu}(\ell) \log N$. The constant $\widehat{\mu}(\ell)$, explicit for the three algorithms, is a variant of the Khinchine constant [29]. For the standard Euclidean algorithm, it is equal to (see [48])

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

Plan. In Section 2, after describing our three Euclidean algorithms and their underlying dynamical systems, we state and prove Theorem 1 (using Theorem 0) and state our main results, Theorems 2 (our version of Dolgopyat's bounds), 3 (CLT), and 4 (LLT). The first part of Section 3 is devoted to proving Theorem 2 and checking that its assumptions hold for our three algorithms. In the last two subsections of Section 3, we obtain quasiperiodicity results. With Theorem 2, they entail expected properties for the Dirichlet series. In Sections 4 and 5, we present the proofs of Theorems 3 and 4, respectively.

2. Dynamical methods and statement of results.

After a description of the three Euclidean algorithms to be studied and their associated dynamical systems (§2.1), we introduce the weighted transfer operator in §2.2. Section 2.3 explains the fundamental rôle this operator plays in the distributional analysis of truncated real trajectories. Theorem 1 is stated there, and proved using Hwang's Quasi-Power result (Theorem 0). Next, we turn to rational inputs and Euclidean algorithms. We introduce in §2.4 Dirichlet series of moment generating functions, which we relate to the quasi-inverse of the weighted transfer operator. We briefly explain in §2.5 how to apply the Perron formula to extract coefficients of Dirichlet series, stating also estimates à la Dolgopyat (Theorem 2) useful for the applicability of this formula. Finally, in §2.6 we state our CLT (Theorem 3) and our LLT for lattice costs (Theorem 4).

2.1. **Euclidean algorithms and their associated interval maps.** Three Euclidean algorithms are to be analyzed; each of them is related to a Euclidean division. Let $v \geq u \geq 1$ be integers. The classical division, corresponding to the standard Euclidean algorithm $\mathcal{G}$, $v = mu + r$ produces an integer $m \geq 1$ and an integer remainder r such that $0 \leq r < u$. The centered division (centered algorithm $\mathcal{K}$) requires $v \geq 2u$ and takes the form $v = mu + s$, with integer $s \in [-u/2, +u/2[$. Letting $s = \epsilon r$, with $\epsilon = \pm 1$ (and $\epsilon = +1$ if $s = 0$), it produces an integer remainder r such that $0 \leq r \leq u/2$, and an integer $m \geq 2$. The odd division (odd algorithm $\mathcal{O}$) produces an odd quotient: it is of the form $v = mu + s$ with m odd and integer $s \in [-u, +u[$. Letting $s = \epsilon r$, with $\epsilon = \pm 1$ (and $\epsilon = +1$ if $s = 0$), it produces an integer remainder r with $0 \leq r \leq u$, and an odd integer $m \geq 1$. In the three cases, the divisions are defined by pairs $q = (m, \epsilon)$, which are called the digits. (See Figure 1.)
Instead of the integer pair (u, v), we may consider the rational u/v, since the pair (du, dv) produces the same sequence of digits as (u, v), up to multiplying all remainders r by d. Then, the division expressing the pair (u, v) as a function of (r, u) is replaced by a linear fractional transformation (LFT) h that expresses the rational u/v as a function of r/u. For each algorithm, the rational u/v belongs to the interval $\mathcal{I}'$ defined in Fig 1.

To summarize, on the input (u, v), each algorithm performs a sequence of admissible Euclidean divisions, of the form $v = mu + \epsilon r$ with $r/u \in \mathcal{I}' \cup \{0\}$, and $(m, \epsilon) \in \mathcal{D}_1$. On the input u/v, it performs a sequence of LFT's from a *generic* set $\mathcal{H}$ (depending on the algorithm) whose elements $h_{[q]}$ are indexed by the admissible pairs $q = (m, \epsilon)$ of $\mathcal{D}_1$ and are of the form $h_{[m,\epsilon]}(x) = 1/(m + \epsilon x)$. The LFTs appearing in the final step belong to a subset $\mathcal{F} \subset \mathcal{H}$ and are indexed by the admissible pairs $q = (m, \epsilon)$ of $\mathcal{D}_1 \cap \mathcal{D}_2$. (Fig. 1).

Each algorithm applied to a rational u/v builds a specific continued fraction

$$\frac{u}{v} = \cfrac{1}{m_1 + \cfrac{\epsilon_1}{m_2 + \cfrac{\epsilon_2}{\ddots + \cfrac{\epsilon_{P-1}}{m_P}}}}, \tag{2.1}$$

of depth P, decomposing also u/v as

$$(u/v) = h_1 \circ h_2 \circ \ldots \circ h_P(0) = h(0), \tag{2.2}$$

with $h_i \in \mathcal{H}, 1 \leq i \leq P-1$, and $h_P \in \mathcal{F}$. We are interested in various costs related to an execution of the algorithm. The most basic one is the number of steps $P(u, v)$. In general, given a digit-cost function c on the set $\mathcal{H}$, we consider additive cost functions C of the form

$$C(u, v) := \sum_{i=1}^{P(u,v)} c(h_i). \tag{2.3}$$

We next see how to associate to each algorithm a *dynamical system of the interval,* $T : \mathcal{I} \to \mathcal{I}$. The interval $\mathcal{I}$ is defined in Figure 1. The map T extends the map defined on rationals by $T(u/v) = r/u$, where r is the

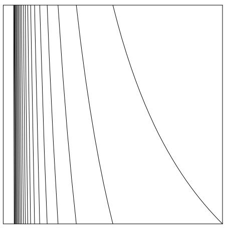
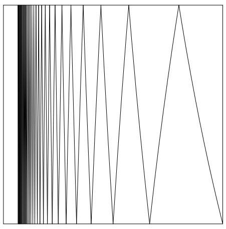
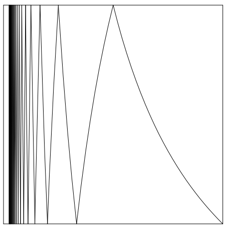
Algorithm	$\mathcal{G}$ (standard)	$\mathcal{K}$ (centered)	$\mathcal{O}$ (odd)
Intervals	$\mathcal{I} = [0,1]$, $\mathcal{I}' =]0,1[$	$\mathcal{I} = [0,1/2]$, $\mathcal{I}' =]0,1/2]$	$\mathcal{I} = [0,1]$, $\mathcal{I}' =]0,1]$
Generic conditions *$\mathcal{D}_1$ on pairs (m,ϵ)*	$m \geq 1$, $\epsilon = +1$	$m \geq 2$, $\epsilon = \pm 1$ if $m = 2$ then $\epsilon = +1$	$m \geq 1$ odd, $\epsilon = \pm 1$ if $m = 1$ then $\epsilon = +1$
Final cond. $\mathcal{D}_2$	$m \geq 2$	$\epsilon = +1$	$\epsilon = +1$
Graph of *$T(x) = \lvert \frac{1}{x} - A(\frac{1}{x}) \rvert$*			
Function $A(y)$	Integer part of y	Nearest integer to y i.e. m s.t. $y - m \in [-1/2, +1/2[$	Nearest odd integer to y i.e. m odd s.t. $y - m \in [-1, +1[$
Contraction ratio	$\rho = 1/\phi^2$	$\rho = 1/(\sqrt{2}+1)^2$	$\rho = 1/\phi^2$
Invariant density f_1	$\dfrac{1}{\log 2}\dfrac{1}{1+x}$	$\dfrac{1}{\log \phi}[\dfrac{1}{\phi + x} + \dfrac{1}{\phi^2 - x}]$	$\dfrac{1}{\phi - 1 + x} + \dfrac{1}{\phi^2 - x}$
Entropy	$\dfrac{\pi^2}{6 \log 2}$	$\dfrac{\pi^2}{6 \log \phi}$	$\dfrac{\pi^2}{9 \log \phi}$

FIGURE 1. The three Euclidean systems ($\phi = (1+\sqrt{5})/2$).

remainder of the Euclidean division on (u,v). We get

$$T(x) := \left| \frac{1}{x} - A\left(\frac{1}{x}\right) \right| , \quad x \neq 0 , \quad T(0) = 0 ,$$

A depends on the algorithm and is defined in Figure 1. It is easy to see that the maps associated to the three algorithms belong to the following class:

Definition. [Piecewise complete maps of the interval] *A map $T : \mathcal{I} \to \mathcal{I}$ is piecewise complete if there exist a (finite or countable) set $\mathcal{Q}$, whose elements are called digits, and a partition $\{\mathcal{I}_q\}_{q \in \mathcal{Q}}$ (modulo a countable set) of the interval $\mathcal{I}$ into open subintervals $\mathcal{I}_q$ such that the restriction of T to $\mathcal{I}_q$ extends to a a bijective mapping of class $\mathcal{C}^2$ from the closure of $\mathcal{I}_q$ to $\mathcal{I}$.*

We may consider the set $\mathcal{H} = \{h_{[q]}\}$ of branches of the inverse function T^{-1} of a general piecewise complete map T, naturally indexed by the set $\mathcal{Q}$. The set of the inverse branches of the iterate T^k is $\mathcal{H}^k$; its elements are of the form $h_{[q_1]} \circ h_{[q_2]} \circ \cdots \circ h_{[q_k]}$ where k is called the *depth* of the branch. Setting $\mathcal{H}^0 = \{\mathrm{Id}\}$, the set $\mathcal{H}^\star := \cup_{k \geq 0} \mathcal{H}^k$ is the semi-group generated by $\mathcal{H}$. Each interval $h(\mathcal{I})$ for h of depth k is called a *fundamental interval of depth k.*

2.2. **Transfer operators of interval maps.**

The maps $(\mathcal{I}, T)$ associated to the three Euclidean dynamical systems belong in fact to a subclass of piecewise complete mappings :

Definition [Good class]. *A piecewise complete interval map* $(\mathcal{I}, T)$ *belongs to the good class if:*

(i) *T is piecewise uniformly expanding, i.e., 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 infimum of such* ρ *is called the contraction ratio, and satisfies*

$$\rho = \limsup_{n\to\infty} \left(\max\{|h'(x)|; h \in \mathcal{H}^n, x \in \mathcal{I}\}\right)^{1/n} . \tag{2.4}$$

(ii) *There is* $\widehat{K} > 0$*, called the distorsion constant, so that every inverse branch* h *of* T *satisfies* $|h''(x)| \le \widehat{K}|h'(x)|$ *for all* $x \in \mathcal{I}$*.*

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

(iv) *The map* T *is not piecewise affine and it is not conjugated with a piecewise affine map with all slopes of equal absolute value.*

(Note that maps in the good class are automatically topologically mixing.) The distortion condition in (ii) follows from the $\mathcal{C}^2$ assumption when there are finitely many branches, and is just Renyi's condition otherwise.

To check that maps associated to our algorithms are in the good class, for $\sigma_0 = 1/2$, use $|h'_{[m,\epsilon]}| = O(m^{-2})$. (See also Figure 1, and [41] for proofs).

If $\mathcal{I}$ is endowed with an initial probability density g_0 with respect to Lebesgue measure, T acts on it and transforms it into a new density g_1. The operator $\mathbf{H}$ such that $g_1 = \mathbf{H}[g_0]$ is called the density transformer, or the Perron-Frobenius operator (acting now on L^1 functions, soon we shall restrict its domain). An application of the change of variable formula gives

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

It is useful to deal with a more general operator, the transfer operator $\mathbf{H}_s$ which depends on a complex parameter s:

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

(Note that $\mathbf{H}_1 = \mathbf{H}$.) If $\sigma := \Re s > \sigma_0$, then $\mathbf{H}_s$ acts boundedly on the Banach space $\mathcal{C}^1(\mathcal{I})$ of $\mathcal{C}^1$ functions on $\mathcal{I}$ endowed with the norm

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

We consider nonnegative cost functions satisfying the following condition:

Definition [Condition $\mathcal{MG}$ –Moderate growth–]. *Let* $\mathcal{H}$ *be the inverse branches of a map in the good class. A digit-cost* $c : \mathcal{H} \to \mathbb{R}^+$ *is of moderate growth if it is non identically zero and if the series*

$$\sum_{h\in\mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s \tag{2.5}$$

converges when $(\Re s, \Re w)$ *belongs to a real neighborhood of* $(1, 0)$*, or, equivalently to* $\Sigma_0 \times W_0$ *with* $\Sigma_0 =]\widehat{\sigma}_0, +\infty]$*, for* $\sigma_0 \le \widehat{\sigma}_0 < 1$*, and* $W_0 =] - \infty, \nu_0[$ *for* $\nu_0 > 0$*.*

Definition [Class $\mathcal{GMG}$] *A triple* $(\mathcal{I}, T, c)$ *formed with an interval map* $T : \mathcal{I} \to \mathcal{I}$ *of the good class, and a digit-cost* c *of moderate growth will be said to be of* $\mathcal{GMG}$*-type.*

We extend the digit-cost to a cost function, also denoted c, on $\mathcal{H}^\star$ by

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

We can now define the weighted composition operator which depends on two (complex) parameters s and w,

$$\mathbf{H}_{s,w}[f](x) := \sum_{h \in \mathcal{H}} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x) . \tag{2.6}$$

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

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

and the quasi-inverse $(I - \mathbf{H}_{s,w})^{-1}$ can be written (formally) as

$$(I - \mathbf{H}_{s,w})^{-1}[f](x) := \sum_{h \in \mathcal{H}^\star} \exp[wc(h)] \cdot |h'(x)|^s \cdot f \circ h(x) . \tag{2.7}$$

We recall next some well-known spectral properties of the transfer operator $\mathbf{H}_{s,w}$. Endow the Banach space $\mathcal{C}^1(\mathcal{I})$ with the norm $\| \cdot \|_{1,1}$. It is known that for $(\Re s, \Re w) \in \Sigma_0 \times W_0$ the operator $\mathbf{H}_{s,w}$ is bounded but not compact acting on $\mathcal{C}^1(\mathcal{I})$ [see e.g. [4, 9]]; however, it is quasi-compact. We recall the definition of *quasi-compactness* for a bounded operator $\mathbf{L}$ on a Banach space: Denote by $\mathrm{Sp}\, \mathbf{L}$ the spectrum of $\mathbf{L}$, by $R(\mathbf{L})$ its spectral radius, and by $R_e(\mathbf{L})$ its *essential spectral radius,* i.e., the smallest $r \geq 0$ such that any $\lambda \in \mathrm{Sp}(\mathbf{L})$ with modulus $|\lambda| > r$ is an isolated eigenvalue of finite multiplicity. An operator $\mathbf{L}$ is *quasi-compact* if $R_e(\mathbf{L}) < R(\mathbf{L})$ holds.
We denote the partial derivatives of first and second order of a function $F(s,w)$ at (a,b) by $F'_w(a,b)$, $F'_s(a,b)$, $F''_{w^2}(a,b)$, $F''_{s^2}(a,b)$, $F''_{ws}(a,b)$.

Proposition 0 [Classical spectral properties of transfer operators]. *Let* $\mathbf{H}_{s,w}$ *be the transfer operator (2.6) associated to a* $\mathcal{GMG}$ *triple* $(\mathcal{I}, T, c)$ *with contraction constant* ρ*. Denote by* $R(s,w)$ *its spectral radius and* $R_e(s,w)$ *its essential spectral radius. Let* Σ_0*,* W_0 *be the real sets from (2.5). When* $w = 0$*, we omit the second index in the operator and its associated objects.*

(1) [Quasi-compactness.] *Let* $\rho < \widehat{\rho} < 1$*. If* $(\sigma = \Re s, \nu = \Re w) \in \Sigma_0 \times W_0$*, then* $\mathbf{H}_{s,w}$ *acts boundedly on* $\mathcal{C}^1(\mathcal{I})$*. Then* $R(s,w) \leq R(\Re s, \Re w)$ *and* $R_e(s,w) \leq \widehat{\rho} \cdot R(\Re s, \Re w)$*, in particular* $\mathbf{H}_{s,w}$ *is (uniformly) quasi-compact for real* (s,w)*.*

(2) [Unique dominant eigenvalue.] *For real* $(\sigma, \nu) \in \Sigma_0 \times W_0$*,* $\mathbf{H}_{\sigma,\nu}$ *has a unique eigenvalue* $\lambda(\sigma,\nu)$ *of maximal modulus, which is real and simple, the dominant eigenvalue. The associated eigenfunction* $f_{\sigma,\nu}$ *is strictly positive, and the associated eigenvector* $\widehat{\mu}_{\sigma,\nu}$ *of the adjoint operator* $\mathbf{H}^*_{\sigma,\nu}$ *is a positive Radon measure. With the normalization conditions,* $\widehat{\mu}_{\sigma,\nu}[1] = 1$ *and* $\widehat{\mu}_{\sigma,\nu}[f_{\sigma,\nu}] = 1$*, the measure* $\mu_{\sigma,\nu} := f_{\sigma,\nu} \widehat{\mu}_{\sigma,\nu}$ *is a probability measure. In particular,* $\widehat{\mu}_1$ *is Lebesgue measure, with* $\lambda(1) = 1$*.*

(3) [Spectral gap.] *For real parameters* $(\sigma, \nu) \in \Sigma_0 \times W_0$, *there is a spectral gap, i.e., the subdominant spectral radius* $r_{\sigma,\nu} \geq R_e(\sigma, \nu)$ *defined by* $r_{\sigma,\nu} := \sup\{|\lambda|; \lambda \in \mathrm{Sp}(\mathbf{H}_{\sigma,\nu}), \lambda \neq \lambda(\sigma, \nu)\}$, *satisfies* $r_{\sigma,\nu} < \lambda(\sigma, \nu)$.

(4) [Analyticity in compact sets.] *The operator* $\mathbf{H}_{s,w}$ *depends analytically on* (s, w) *for* $(\Re s, \Re w) \in \Sigma_0 \times W_0$. *Thus,* $\lambda(\sigma, \nu)^{\pm 1}$, $f^{\pm 1}_{\sigma,\nu}$, *and* $f'_{\sigma,\nu}$ *depend analytically on* $(\sigma, \nu) \in \Sigma_0 \times W_0$, *and are uniformly bounded in any compact subset.*

(5) [Analyticity in a neighborhood of $(1, 0)$.] *If* (s, w) *is complex near* $(1, 0)$ *then* $\lambda(s, w)^{\pm 1}$, $f^{\pm 1}_{s,w}$, *and* $f'_{s,w}$ *are well-defined and analytic; moreover, for any* θ, *with* $r_1 < \theta < 1$, *one has* $r_{1,w}/|\lambda(1, w)| \leq \theta$.

(6) [Derivatives of the pressure.] *For* $(\sigma, \nu) \in \Sigma_0 \times W_0$, *define the pressure function* $\Lambda(\sigma, \nu) = \log \lambda(\sigma, \nu)$.
(6.a) $\Lambda'(1)$ *is the opposite of the Kolmogorov entropy of the dynamical system* (T, μ_1). *Also,* $\Lambda'_w(1, 0)$ *is the* μ_1*–average of the cost:*

$$\Lambda'(1) = -\int_{\mathcal{I}} \log |T'(x)|\, f_1(x)\, dx < 0\,, \qquad \Lambda'_w(1, 0) = \sum_{h \in \mathcal{H}} c(h) \int_{h(\mathcal{I})} f_1(x)\, dx.$$

(6.b) *The pressure is strictly convex in* s, *i.e.,* $\Lambda''_{s^2}(\sigma, \nu) > 0$. *Also* $\Lambda(\sigma + q\nu, \nu)$ *is strictly convex as a function of* ν *for all fixed* $q \neq 0$ *and fixed* σ. *If* c *is not constant, the pressure is strictly convex in* w, *i.e.,* $\Lambda''_{w^2}(\sigma, \nu) > 0$.

(7) [Function $w \mapsto \sigma(w)$.] *There is a complex neighborhood* $\mathcal{W}$ *of* 0 *and a unique function* $\sigma : \mathcal{W} \to \mathbb{C}$ *such that* $\lambda(\sigma(w), w) = 1$, *this function is analytic, and* $\sigma(0) = 1$.

Proof. We refer to [4, 9, 47] except for claim (7) (which follows from $\lambda'(1) \neq 0$ and the implicit function theorem) and for claim (6):
(6.a) Taking the derivatives at (σ, ν) of $\mathbf{H}_{s,w}[f_{s,w}] = \lambda(s, w) f_{s,w}$, integrating on $\mathcal{I}$, and using that $\mathbf{H}^*_{\sigma,\nu}$ preserves $\widehat{\mu}_{\sigma,\nu}$, gives the expressions as integrals. To finish, apply Rohlin's formula.
(6.b) Convexity of the pressure is an old theme, see, e.g., [40] and also [36], Chapter 4, Prop. 10 in [44], Prop. 3.8 in [13]. We follow here the last work where it is proven that $\Lambda(s)$ is strictly convex as soon as the second part of Condition (iv) in the definition of the good class holds. Except if c is constant, the same is true for $w \mapsto \Lambda(s, w)$. The proof on lines $L : w \mapsto \Lambda(s + qw, w)$ is similar. If L is not strictly convex, there exists A so that $c(h) + q \log |h'(x)| = A$ for all h, and all x. This implies that T has affine branches, which is not compatible with the first part of (iv). □

We have already remarked that the three Euclidean dynamical systems belong to the good class. Note that, for an inverse branch of depth k, of the form $h_{[q_1]} \circ \cdots \circ h_{[q_k]}$, the interval $h(\mathcal{I})$ gathers all real x for which the k first digits of the CF expansion are $(q_1, q_2, \ldots, q_k)$. A sufficient condition for the cost c to be of moderate growth is $c(m, \epsilon) := c(h_{[m,\epsilon]}) = O(\log m)$. Furthermore, each inverse branch of any depth is a linear fractional transformation $h(x) = (ax + b)/(cx + d)$, with a, b, c, d coprime integers, with determinant $ad - bc = \pm 1$, and denominator $D[h]$ related to $|h'|$ through:

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

Therefore, the transfer operator can be alternatively defined by

$$\mathbf{H}_{s,w}[f](x) = \sum_{h\in\mathcal{H}} \exp[wc(h)] \frac{1}{D[h](x)^{2s}}\ f(h(x))\,. \tag{2.8}$$

The above reformulation will be useful in §2.4.
We associate to each of the three algorithms and its final set $\mathcal{F} \subset \mathcal{H}$, a transfer operator
(2.9)

$$\mathbf{F}_{s,w}[f](x) := \sum_{h\in\mathcal{F}} \exp[wc(h)] \frac{1}{D[h](x)^{2s}}\ f(h(x)) = \mathbf{H}_{s,w}[f \cdot 1_{\cup_{h\in\mathcal{F}} h(\mathcal{I})}](x).$$

It is easy to see that $\mathbf{F}_{s,w}$ acts boundedly on $\mathcal{C}^1(\mathcal{I})$ for $(\Re s, \Re w) \in \Sigma_0 \times W_0$, and to generalize the relevant statements of Proposition 0 to this operator.

2.3. **Transfer operators and real trajectories.** We consider one of our three algorithms, or more generally any triple of $\mathcal{GMG}$ type, with a non-constant cost c. The interval $\mathcal{I}$ is endowed with a probability measure with smooth density f and the set of endpoints of fundamental intervals (rational points for Euclidean algorithms) can be neglected. We are interested in the asymptotic behavior of the distribution of $C_n(x) := \sum_{i=1}^n c(h_i)$ when the truncation degree n tends to ∞. As already mentioned, a very convenient tool is the Lévy moment generating function of the cost, $\mathbb{E}[\exp(wC_n)] = \sum_{h\in\mathcal{H}^n} \exp[wc(h)] \cdot \int_{h(\mathcal{I})} f(y)\,dy$. The change of variables $y = h(u)$ gives
(2.10)

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

The above relation is fundamental for analysing costs on truncated real trajectories, as we explain next.
By Proposition 0, for any θ with $r_1 < \theta < 1$, there is a small complex neighborhood $\mathcal{W}$ of 0, so that, for $w \in \mathcal{W}$, the operator $\mathbf{H}_{1,w}$ splits as $\mathbf{H}_{1,w} = \lambda(1,w)\,\mathbf{P}_{1,w} + \mathbf{N}_{1,w}$, where $\mathbf{P}_{1,w}$ is the projector for the eigenvalue $\lambda(1,w)$ and $R(\mathbf{N}_{1,w}) \le \theta\lambda(1,w)$. Therefore,

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

which entails

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

with a uniform O-term for $w \in \mathcal{W}$. In other words the moment generating function behaves as a "quasi-power," and we may apply the following result:

Theorem 0. [Hwang's Quasi-Power Theorem] [25, 26, 27] *Assume that the moment generating functions for a sequence of functions $\widehat{C}_N$ on probability spaces $(\widehat{\Omega}_N, \widehat{\mathbb{P}}_N)$ are analytic in a complex neighborhood $\mathcal{W}$ of zero, where*

$$\mathbb{E}_N[\exp(w\widehat{C}_N)] = \exp[\beta_N U(w) + V(w)]\left(1 + O(\kappa_N^{-1})\right)\ , \tag{2.11}$$

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

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

$$\mathbb{V}_N[\widehat{C}_N] = \beta_N U''(0) + V''(0) + O(\kappa_N^{-1})\,.$$

Furthermore, the distribution of $\widehat{C}_N$ *on* $\widehat{\Omega}_N$ *is asymptotically Gaussian, with speed of convergence* $O(\kappa_N^{-1} + \beta_N^{-1/2})$.

For each fixed $k \geq 3$*, there is a polynomial* P_k *of degree at most* k*, with coefficients depending on the derivatives of order at most* k *at* 0 *of* U *and* V*, so that the moment of order* k *satisfies*

$$\mathbb{E}_N[\widehat{C}_N^k] = P_k(\beta_N) \ + O\left(\frac{\beta_N^{k-1}}{\kappa_N}\right)\,, \tag{2.12}$$

with a O*–term uniform in* k*.*

Proof. [Sketch.] This statement encapsulates a classical calculation analogous to the proof of the central limit theorem by characteristic functions.The speed of convergence then results from the Berry-Esseen inequalities. the moment estimates are consequences of the derivability of analytic functions. ☐

For our application, we set $\widehat{\Omega}_n = (\mathcal{I}, f\,dx)$ for all n, $\widehat{C}_n = C_n$, $\beta_n = n$, $\kappa_n = \theta^{-n}$. The function U is the pressure function $w \mapsto \Lambda(1,w)$, and $V(w) = \log\left(\int_{\mathcal{I}} \mathbf{P}_{1,w}[f](u)\,du\right)$. Since c is not constant, the function $\Lambda(1,w)$ is strictly convex at $w=0$, (see Proposition 0) and $U''(0) \neq 0$. Thus, using the formula for $\Lambda'_w(1,0)$ in Proposition 0, Theorem 0 entails the following Gaussian asymptotic distribution result, which applies in particular to our three Euclidean algorithms.

Theorem 1. *For a triple* $(\mathcal{I}, T, c)$ *of* $\mathcal{GMG}$ *type with non-constant* c *and any probability* $\mathbb{P}$ *on* $\mathcal{I}$ *with a* $\mathcal{C}^1$ *density, there are* $\widehat{\mu}(c) > 0$ *and* $\widehat{\delta}(c) > 0$ *so that for any* n*, and any* $Y \in \mathbb{R}$

$$\mathbb{P}\left[x \ \Big| \ \frac{C_n(x) - \widehat{\mu}(c)n}{\widehat{\delta}(c)\sqrt{n}} \leq Y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Y} e^{-y^2/2}\,dy + O\left(\frac{1}{\sqrt{n}}\right)\,.$$

Furthermore, (recalling that r_1 *is the subdominant spectral radius of the density transformer* $\mathbf{H}$*), for any* θ *which satisfies* $r_1 < \theta < 1$*, one has :*

$$\mathbb{E}\,[C_n] = \widehat{\mu}(c)\cdot n + \widehat{\eta}(c) + O(\theta^n)\,, \quad \mathbb{V}\,[C_n] = \widehat{\delta}^2(c)\cdot n + \widehat{\delta}_1(c) + O(\theta^n)\,,$$

with $\widehat{\mu}(c) = \Lambda'_w(1,0) = \lambda'_w(1,0)$*,* $\widehat{\delta}^2(c) = \Lambda''_{w^2}(1,0) = \lambda''_{w^2}(1,0) - \lambda'^2_w(1,0)$*. Finally,* $\widehat{\mu}(c)$ *involves the invariant density* f_1*,*

$$\widehat{\mu}(c) = \sum_{q\in\mathcal{Q}} c(q)\int_{\mathcal{I}_q} f_1(x)\,dx\,.$$

Note that f_1 is explicitly given in Figure 1 for $\mathcal{G}$, $\mathcal{O}$, and $\mathcal{K}$, so that $\widehat{\mu}(c)$ is computable in these cases.

2.4. **Dirichlet generating functions and transfer operators.** We restrict now our study, for each of our three algorithms, to (nonzero) rational inputs $x = (u/v) \in \mathcal{I}'$ and a cost c of moderate growth. The intervals $\mathcal{I}'$ are defined in Figure 1. We consider the sets

$$\widetilde{\Omega} := \{(u, v) \in \mathbb{N}_\star^2 , \frac{u}{v} \in \mathcal{I}'\} , \qquad \Omega := \{(u, v) \in \widetilde{\Omega} \,; \gcd (u, v) = 1\} ,$$

and we endow the sets

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

with uniform probabilities $\widetilde{\mathbb{P}}_N$ and $\mathbb{P}_N$, respectively. For the moment we only consider Ω_N.

To study the distribution of the total cost $C(u, v)$ (2.3) associated to some digit-cost c and restricted to Ω_N we use its *moment generating function:*

$$\mathbb{E}_N[\exp(wC)] := \frac{\Phi_w(N)}{\Phi_0(N)} , \tag{2.13}$$

where $\Phi_w(N) = \Phi_{c,w}(N)$ is the cumulative value of $\exp[wC]$ on Ω_N,

$$\Phi_w(N) := \sum_{(u,v)\in\Omega_N} \exp[wC(u, v)] , \quad \Phi_0(N) = |\Omega_N| . \tag{2.14}$$

Extending the principles defined in [45, 47, 48], we replace the sequence of moment generating functions by a single bivariate *Dirichlet series,* henceforth called the *Dirichlet-moment generating function:*

$$S(s, w) := \sum_{(u,v)\in\Omega} \frac{1}{v^s} \exp[wC(u, v)] = \sum_{n\ge1} \frac{c_n(w)}{n^s} , \tag{2.15}$$

where $c_n(w) := \sum_{(u,v)\in\Omega_n ,v=n} \exp[wC(u, v)]$.

Since the partial sum of the coefficients of the series $S(s, w)$ satisfies

$$\sum_{n\le N} c_n(w) = \Phi_w(N) , \tag{2.16}$$

to analyze the moment generating function $\mathbb{E}_N[\exp(wC)]$ of the cost C on Ω_N, it suffices to estimate the functions $\Phi_w(N)$ (asymptotically in $N \to \infty$, and uniformly in w in a complex neighborhood of 0).

As we previously did for 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 a input $(u, v) \in \Omega$, performing $P(u, v)$ steps uniquely decomposes the rational

$$\frac{u}{v} = h_1 \circ h_2 \circ \ldots \circ h_P(0) = h(0),$$

with $h_i \in \mathcal{H}, 1 \le i \le P - 1$ and $h_P \in \mathcal{F}$. Thus, each Euclidean algorithm defines a bijection between the sets Ω and $\mathcal{H}^\star \times \mathcal{F}$. In view of (2.8, 2.9) and (2.15), the relations

$$v = D[h](0), \qquad C(u, v) = c(h) = \sum_{i=1}^{P(u,v)} c(h_i)$$

provide the desired expression for the Dirichlet moment generating function $S(s,w)$ in terms of the transfer operators $\mathbf{H}_{s,w}$ and $\mathbf{F}_{s,w}$:

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

Returning to $\widetilde{\Omega}$, we remark that each element (u',v') of $\widetilde{\Omega}$ can be written in a unique way as (du, dv) with $d = \gcd(u,v)$ and $(u,v) \in \Omega$, and, as already observed, an execution of the algorithm on (u',v') leads to the same decomposition (2.2), the two costs $C(u',v')$ and $C(u,v)$ being equal. We may define a Dirichlet moment generating function $\widetilde{S}(s,w)$, and we get

$$\widetilde{S}(2s,w) := \sum_{(u,v)\in\widetilde{\Omega}} \frac{1}{v^{2s}} \exp[wC(u,v)] = \zeta(2s)\,\mathbf{F}_{s,w} \circ (I - \mathbf{H}_{s,w})^{-1}[1](0)\,. \tag{2.18}$$

Using well-known properties of the Riemann zeta function $\zeta(s)$, all our results for $\widetilde{\Omega}_N$ will follow from those on Ω_N.

In view of (2.13, 2.15, 2.16), the relations (2.17), and (2.18) connecting the Dirichlet moment generating function with the transfer operator are the analogues for rational trajectories of the relation (2.10) for the truncated real trajectories. In the case of rational trajectories, we have to work with the quasi-inverse and extract the coefficients of Dirichlet series: This is why the discrete problem is more difficult to solve than the continuous problem.

2.5. **Perron's formula and Dolgopyat's estimates.** We wish to evaluate the sum $\Phi_w(N)$ of the first N coefficients of the Dirichlet series $S(2s,w)$. Our first main tool towards this goal is the Perron formula. The *Perron formula* of order two (see e.g. [17]) applied term by term to a Dirichlet series $F(s) = \sum_{n\ge1} a_n n^{-s}$ and a vertical line $\Re s = D > 0$ inside the domain of convergence of F says that

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

Applying Perron's formula to the Dirichlet series $S(2s,w)$, we find

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

Thus, Perron's formula gives us information on $\Psi_w(N)$, which is just a Cesàro sum of the $\Phi_w(Q)$:

$$\Psi_w(N) = \sum_{Q\le N}\sum_{n\le Q} c_n(w) = \sum_{Q\le N}\Phi_w(Q)\,.$$

Proposition 0 and Relation (2.17) show that $s \mapsto S(2s,w)$ has a possible pole at $s = \sigma(w)$, where $\sigma(w)$ is the unique complex number near 1 for which $\lambda(\sigma(w),w) = 1$. In particular, the integral in (2.20) is well-defined only if $D > \sup_w(\Re\sigma(w))$ (note that $\sup_w(\Re\sigma(w)) > 1$ can be made arbitrarily close to 1 by taking w close enough to zero). To combine the Perron formula with Cauchy's residue theorem, we wish to modify the integration contour $\Re s = D$ into a contour containing $\sigma(w)$ as a unique pole of $S(2s,w)$. This is possible and leads to a quasi-power expansion if there is $\alpha > 0$, such that

(i) $S(2s, w)$ admits $s = \sigma(w)$ as a unique pole in the strip $|\Re s - 1|| \le \alpha$.
(ii) In a "perforated" half-plane $\{\Re s \ge 1 - \alpha_0\,, |s-1| > \alpha_0/2\}$, for small $\alpha > \alpha_0 > 0$, the estimates $|S(2s, w)| \le \max(1, (\Im s)^\xi)$ for $0 < \xi < 1$, hold, uniformly in w close to 0.

Note that bounds of the type *(ii)* are extremely difficult to obtain for general Dirichlet series. By the above discussion, such bounds would follow from similar estimates on the quasi-inverse of the operator, which are closely related to those obtained by Dolgopyat [16]. In the spirit of Chernov [11], Dolgopyat [16] introduced several "uniform nonintegrability" (*UNI*) conditions, which allowed him to control oscillatory integrals associated to iterates of transfer operators $\mathbf{H}_s$ for $s = \sigma + it$, with fixed σ close to 1, and prove exponential decay of correlations for some flows. We shall give a new formulation of Dolgopyat's strongest such *UNI* condition, which appeared implicitly in Section 5 of his paper [16] and turns out to be satisfied by our three algorithms. This condition is stated as an assumption on the derivatives of the inverse branches of the dynamical system in Section 3.2. In Section 3, we shall prove the following theorem, which is the central functional analytic result of the paper. In the statement, we use the following family of equivalent norms on $\mathcal{C}^1(\mathcal{I})$:

$$\|f\|_{1,t} := \sup|f| + \frac{\sup|f'|}{|t|}\,, \quad t \ne 0\,. \tag{2.21}$$

Theorem 2 [Dolgopyat-type estimates]. *Let $(\mathcal{I}, T, c)$ be a triple of $\mathcal{GMG}$-type, with contraction ratio $\rho < 1$, and such that Condition* UNI *from §3.2 holds. Let $\mathbf{H}_{s,w}$ be its weighted transfer operator (2.6) acting on $\mathcal{C}^1(\mathcal{I})$.*
For any ξ, with $0 < \xi < 1/5$, there is a (real) neighborhood $\Sigma_1 =]1-\alpha, 1+\alpha[$ of 1 (which depends only on $(\mathcal{I}, T)$ and not on c), a (real) neighborhood W_1 of 0, and there is $M > 0$ such that, for all $s = \sigma + it$, $w = \nu + i\tau$ with $(\sigma, \nu) \in \Sigma_1 \times W_1$ and $|t| \ge 1/\rho^2$,

$$\|(I - \mathbf{H}_{s,w})^{-1}\|_{1,t} \le M \cdot |t|^\xi\,. \tag{2.22}$$

Note that we shall have to modify Dolgopyat's arguments since we must consider dynamical systems which possess an infinite number of branches (see in particular Lemma 1), and we work with bivariate weighted transfer operators $\mathbf{H}_{s,w}$ involving a cost function.

2.6. **Statement of the Central and Local Limit Theorems.**

We shall see in §4.1 that Perron's Formula (2.20) combined with the fundamental relation (2.17), together with the bounds à la Dolgopyat (Theorem 2) provide a quasi-powers estimate for the Cesàro sum $\Psi_w(N)$. It does not seem easy to transfer this information on $\Psi_w(N)$ to estimates on $\Phi_w(N)$, because the coefficients are complex. The way we overcome this is by first proving (Lemma 11 in Section 4.2) quasi-power estimates for the moment generating function of some "smoothed" version of the cost C, for which the transfer is possible by standard methods. We are then able to apply Theorem 0 to the smoothed model, and show that the two models are close enough in distribution so that the following holds:

Theorem 3. [Central Limit Theorem for rational trajectories.] *For a Euclidean algorithm amongst $\mathcal{G}$, $\mathcal{K}$, $\mathcal{O}$, there is $\gamma > 0$, so that, for any cost c of moderate growth, letting $\Lambda(s)$ be the function from Proposition 0:*
(a) The distribution of the total cost C on Ω_N is asymptotically Gaussian, with speed of convergence $O(1/\sqrt{\log N})$, i.e., there exist two constants $\mu(c) > 0$ and $\delta(c) > 0$ such that, for any N, and any $y \in \mathbb{R}$

$$\mathbb{P}_N\left[(u,v); \frac{C(u,v) - \mu(c)\log N}{\delta(c)\sqrt{\log N}} \le y\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-x^2/2}\,dx + O\left(\frac{1}{\sqrt{\log N}}\right).$$

(b) The mean and the variance satisfy $\mathbb{E}_N[C] = \mu(c)\log N + \eta(c) + O(N^{-\gamma})$, and $\mathbb{V}_N[C] = \delta^2(c)\log N + \delta_1(c) + O(N^{-\gamma})$.
Generally, for each $k \ge 1$, there is a polynomial P_k of degree k so that

$$\mathbb{E}_N[C^k] = P_k(\log N) + O\left(\frac{(\log N)^{2k}}{N^\gamma}\right),$$

with a O-term uniform in k.
(c) In the special case $c \equiv 1$, denoting $\mu := \mu(1)$, $\delta^2 := \delta^2(1)$, we have

$$\mu = \frac{2}{|\Lambda'(1)|} > 0, \quad \delta^2 = \frac{2|\Lambda''(1)|}{|\Lambda'(1)^3|} > 0.$$

In the general case,

$$\mu(c) = \mu\cdot\widehat{\mu}(c), \qquad \delta^2(c) = \widehat{\mu}^2(c)\cdot\delta^2 + \mu\cdot\widehat{\delta}^2(c) + \mu^2\widehat{\mu}(c)\cdot\chi(c) > 0,$$

where $\widehat{\mu}(c) > 0$ and $\widehat{\delta}^2(c) \ge 0$ are given in Theorem 1, and $\chi(c) = \Lambda''_{sw}(1,0)$. Claims (a), (b), and (c) also hold for $\widetilde{\mathbb{P}}_N$ on $\widetilde{\Omega}_N$.

Note that $2/\mu$ is the Kolmogorov entropy for $(T, f_1 dx)$. Also, γ does not depend on the cost. The constant $\chi(c)$ can be viewed as a covariance coefficient between the number of steps P and the cost c. Since there exists a closed form for f_1 in the three cases of interest (cf. Figure 1), the constants μ, and thus $\mu(c)$ can be easily computed (see remark after Theorem 1 and [48]). The constants $\widehat{\delta}(c)$, δ, $\delta(c)$ do not seem to admit a closed form. However, Lhote has proved that they can be computed in polynomial time [32].

In Section 5, restricting to lattice costs, we obtain bounds for $\mathbb{E}_N[\exp(i\tau C)]$ with $\tau \in [-\pi, \pi]$, and combine them with saddle-point estimates to get:

Theorem 4. [Local Limit Theorem for lattice costs.] *For any algorithm among $\mathcal{G}$, $\mathcal{K}$, $\mathcal{O}$, and any lattice cost c of span L and of moderate growth, letting $\mu(c) > 0$ and $\delta^2(c) > 0$ be the constants from Theorem 3, the following holds,*

$$\mathbb{P}_N\left[(u,v); -\frac{L}{2} < C(u,v) - \mu(c)\log N - \delta(c)x\sqrt{\log N} \le \frac{L}{2}\right] = \frac{e^{-x^2/2}}{\delta(c)\sqrt{2\pi\log N}} + O\left(\frac{1}{\log N}\right),$$

with a O uniform for $x \in \mathbb{R}$. The same holds for $\widetilde{\mathbb{P}}_N$ in $\widetilde{\Omega}_N$.

3. PROPERTY *UNI* AND DOLGOPYAT ESTIMATES.

In this section we shall first prove Theorem 2 and check (§3.5) that its additional assumption *(UNI)* holds for our three Euclidean algorithms. Then, we show that Theorem 2 together with aperiodicity results imply useful estimates on $S(s,w)$.
To prove Theorem 2, we use ideas due to Dolgopyat [16]; however, we have to adapt them to our context, i.e., bivariate weighted transfer operators $\mathbf{H}_{s,w}$ associated to triples $(\mathcal{I},T,c)$ of $\mathcal{GMG}$-type with infinitely many branches, as explained after the statement of Theorem 2 in Section 2. One of the main ideas of Dolgopyat was 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, he made two transfers of estimates: first from this L^2-bound into a bound for the sup-norm, next from this sup-norm-bound into the desired bound for the $(1,t)$-norm. Following his strategy, we establish preliminary results in Lemmata 1, 2 and 3. (Note that Lemma 1 is new.) §3.2 is central: our version of the *UNI* Condition is stated and shown to entail the desired exponential estimate for the L^2-norm of the operator (Lemmata 4 and 5). The two transfers are operated in §3.3: they lead to Theorem 2. We then check in Lemma 6 (§3.4) that its additional assumption *(UNI)* holds for our three Euclidean algorithms. In §3.5, we show an aperiodicity result on vertical s-lines and w-lines (Lemma 7 and Proposition 1). Finally, in §3.6, we return to the Dirichlet series $S(s,w)$ and obtain further estimates in the compact neighborhood of the real axis which is not covered by Theorem 2 (Lemmata 8 and 9).

3.1. Preparatory material. Transfer operators $\mathbf{H}_{s,w}$ and their normalizations $\widetilde{\mathbf{H}}_{s,w}$.

Triples $(\mathcal{I},T,c)$ of $\mathcal{GMG}$ type and their associated transfer operators $\mathbf{H}_{s,w}$, acting on $\mathcal{C}^1(\mathcal{I})$, for $(\Re s,\Re w)\in\Sigma_0\times W_0$ were introduced in § 2.2. We summarized in Proposition 0 well-known spectral properties that we shall need, in particular the existence of a dominant eigenvalue $\lambda(\sigma,\nu)$, [with positive eigenfunction $f_{\sigma,\nu}$] or $\lambda(s,w)$ for suitable real $(\sigma,\nu)\in\Sigma_0\times W_0$ [from (2.5)] and complex (s,w) close to $(1,0)$. Recall that $\mathcal{H}^n$ is the set of inverse branches of T^n, $\mathcal{H}^\star$ the set of all inverse branches of any depth, $\rho<1$ the contraction constant and $\widehat{K}$ the distorsion constant.
Defining $K=\widehat{K}/(1-\rho)$, it is not difficult to check that,

$$|h''(x)|\le K|h'(x)|\,,\quad \forall x\in\mathcal{I}\,,h\in\mathcal{H}^\star\,.$$

The above bounded distortion property will play an important rôle. Note for further use that, since the ratio $h''(x)/h'(x)$ is the derivative of $\log|h'(x)|$, setting $L:=e^K$, we get

$$\frac{1}{L}\le\frac{|h'(x)|}{|h'(y)|}\le L\qquad\text{for all } x,y\in\mathcal{I}\text{, for all } h\in\mathcal{H}^\star\,. \tag{3.1}$$

It will be convenient to work with the *normalized operators* $\widetilde{\mathbf{H}}_{s,w}$ defined by

$$\widetilde{\mathbf{H}}_{s,w}[f]=\frac{1}{\lambda(\sigma,\nu)f_{\sigma,\nu}}\mathbf{H}_{s,w}[f_{\sigma,\nu}\cdot f]\,,\quad s=\sigma+it,\ w=\nu+i\tau\,. \tag{3.2}$$

By construction, for $(\sigma,\nu)\in\Sigma_0\times W_0$, the operator $\widetilde{\mathbf{H}}_{\sigma,\nu}$ acting on $\mathcal{C}^1(\mathcal{I})$ has spectral radius equal to 1, and fixes the constant function $\equiv 1$. Also,

$\mathrm{Sp}\,\mathbf{H}_{\sigma+it,\nu} = \lambda(\sigma,\nu)\mathrm{Sp}\,\widetilde{\mathbf{H}}_{\sigma+it,\nu}$. Remark next the inequality $||\widetilde{\mathbf{H}}_{s,w}[f]||_0 \le ||f||_0\,\widetilde{\mathbf{H}}_{\sigma,\nu}[1] = ||f||_0$, which implies the useful bound

$$||\widetilde{\mathbf{H}}_{s,w}||_0 \le 1\,. \tag{3.3}$$

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

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, $A(x) \le kB(x)$. It is synonymous with $A(x) = O(B(x))$ with an absolute O-term. The symbol $\mathcal{W}$ denotes a complex neighborhood of 0 for the variable w. If $J \subset \mathcal{I}$ is a union of intervals, we denote by $|J|$ its Lebesgue measure.*

Relating $\widetilde{\mathbf{H}}_{\sigma,\nu}$ and $\widetilde{\mathbf{H}}_{1,0}$. In order to exploit properties of Lebesgue measure, which is fixed (only) by the dual of $\mathbf{H}_{1,0}$, Dolgopyat uses the following property (see e.g. last lines of p. 367 in [16]): When $(\mathcal{I},T)$ has finitely many branches, there is $A_\sigma \to 1$ as $\sigma \to 1$ so that for positive $f \in \mathcal{C}^1(\mathcal{I})$

$$\widetilde{\mathbf{H}}_{\sigma,0}[f](x) \le A_\sigma\,\widetilde{\mathbf{H}}_{1,0}[f](x)\,. \tag{3.4}$$

The above inequality is not true in general when there are infinitely many branches (it fails for the Gauss map). The purpose of the following lemma is to extend (3.4) to the case of infinitely many branches and bivariate operators, comparing $\widetilde{\mathbf{H}}^n_{\sigma,\nu}$, and $\mu_{\sigma,\nu}$, to their analogues for $(\sigma,\nu) = (1,0)$:

Lemma 1. *For $(\sigma,\nu) \in \Sigma_0 \times W_0$, denote*

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

Let $\mathcal{L}$ be a compact subset of $\Sigma_0 \times W_0$. For $\mathcal{J} \subset \mathcal{H}^k$, denote by $J = \cup_{h\in\mathcal{J}} h(\mathcal{I})$. Then, for $(\sigma,\nu) \in \mathcal{L}$,

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

furthermore, for any $f \in \mathcal{C}^1(\mathcal{I})$, for any integer $k \ge 1$,

$$||\widetilde{\mathbf{H}}^k_{\sigma,\nu}[f]||^2_0 \ll A^{2k}_{\sigma,\nu}\;||\widetilde{\mathbf{H}}^k_{1,0}[|f|^2]||_0\,. \tag{3.5}$$

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

The function $A_{\sigma,\nu}$ depends continuously on (σ,ν) and $A_{1,0} = 1$.

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, by the bounded distortion property (3.1), the ratios (two by two) of the three quantities $a(h)$, $b(h)$, $c(h)$,

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

admit upper and lower bounds that do not depend on h, and are uniform for $(\sigma,\nu) \in \mathcal{L}$. 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 $\mathcal{J} \subset \mathcal{H}^k$, and applying the Cauchy-Schwarz inequality, one gets

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

Then, dominant spectral properties, together with bounded distortion, entail 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 first claim.

Consider now $f \in \mathcal{C}^1(\mathcal{I})$. 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 $\mathcal{L}$, and, by the 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 $\mathbf{H}^k_{1,0}[|f|^2](x)$, which 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 $\lambda(2\sigma-1, 2\nu)^k$. □

Lasota–Yorke bounds. The following lemma describes how $\widetilde{\mathbf{H}}_{s,w}$ acts with respect to the quasi-norm $||.||_1$ when s varies over a vertical line:

Lemma 2. *For every compact subset $\mathcal{L}$ of $\Sigma_0 \times W_0$, there is $C > 0$, so that for all (s,w) with $(\Re s, \Re w) \in \mathcal{L}$, and all $f \in \mathcal{C}^1(\mathcal{I})$,*

$$||\widetilde{\mathbf{H}}^n_{s,w} f||_1 \le C\,(|s|\,||f||_0 + \rho^n\,||f||_1) , \quad \forall n \ge 1 .$$

Proof. The quantity $\widetilde{\mathbf{H}}^n_{s,w}[f]$ can be written as a sum over $h \in \mathcal{H}^n$ of terms

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

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

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

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

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

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

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

We can ensure

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

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

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

The final result follows from (3.3). □

First use of the $(1,t)$-norm. In the bound from Lemma 2 for the derivative of $\widetilde{\mathbf{H}}^n_{s,w}f(x)$, there appear two terms, one which contains a factor $|s|$, the other a decreasing exponential in n. In order to suppress the effect of the factor $|s|$, Dolgopyat uses the family of equivalent norms

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

which appear in the statement of Theorem 2. With this norm and Lemma 2, together with (3.3), we obtain the first (easy) result:

Lemma 3. *For any $t_1>0$, for every compact subset $\mathcal{L}$ of $\Sigma_0\times W_0$, there is $M_0>0$ so that for all $n\ge 1$, all (s,w) for which $(\Re s,\Re w)\in\mathcal{L}$ and $|\Im s|\ge t_1$ we have $||\widetilde{\mathbf{H}}^n_{s,w}||_{1,\Im s}\le M_0$.*

3.2. ***UNI*** Condition and L^2-estimates.

Assuming *UNI,* Dolgopyat first proves that there is $\gamma<1$ so that

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

for all large t and $n_0=O(\log|t|)$. In this subsection, we extend this result to the bivariate operator $\mathbf{H}_{s,w}$, when the number of branches of T is possibly infinite.

Writing $s=\sigma+it$, $w=\nu+i\tau$, $|\widetilde{\mathbf{H}}^n_{s,w}[f](x)|^2$ can be expressed as:

$$\frac{1}{\lambda(\sigma,\nu)^{2n}}\sum_{(h,k)\in\mathcal{H}^n\times\mathcal{H}^n}\exp[wc(h)+\overline{w}c(k)]\cdot\exp[it\Psi_{h,k}(x)]\cdot R_{h,k}(x), \tag{3.6}$$

with

$$\Psi_{h,k}(x):=\log\frac{|h'(x)|}{|k'(x)|}, \tag{3.7}$$

$$R_{h,k}(x)=|h'(x)|^\sigma|k'(x)|^\sigma\frac{1}{f^2_{\sigma,\nu}(x)}(f\cdot f_{\sigma,\nu})\circ h(x)\cdot(\overline{f}\cdot f_{\sigma,\nu})\circ k(x). \tag{3.8}$$

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

$$r_{h,k}(x) = e^{i\omega}|h'(x)|^\sigma |k'(x)|^\sigma \frac{1}{f^2_{\sigma,\nu}(x)}(g \cdot f_{\sigma,\nu}) \circ h(x)(\ell \cdot f_{\sigma,\nu}) \circ k(x)\,, \tag{3.9}$$

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

The functions $\Psi_{h,k}$ play an important rôle here: the sum (3.6) will be split into two parts, according to their properties. The first sum will gather the pairs for which the derivative $|\Psi'_{h,k}(x)|$ has a "small" lower bound, and condition *UNI* will precisely require that there are not "too many" such pairs (h,k), providing a convenient bound for the corresponding integral I^-_n (Lemma 4). The second sum will gather the other pairs (h,k), for which the derivative $|\Psi'_{h,k}(x)|$ has a "large" lower bound. In this case, the Van Der Corput Lemma on oscillatory integrals will be applicable (Lemma 5), giving a bound for the corresponding integral I^+_n.

Let us introduce some notations needed for our formulation of *UNI.* For two inverse branches h et k of same depth, we introduce a "distance:"

$$\Delta(h,k) = \inf_{x \in \mathcal{I}} |\Psi'_{h,k}(x)| = \inf_{x \in \mathcal{I}} \left| \frac{h''}{h'}(x) - \frac{k''}{k'}(x) \right| . \tag{3.10}$$

For h in $\mathcal{H}^n$, and $\eta > 0$, we denote

$$J(h,\eta) := \bigcup_{k \in \mathcal{H}^n, \Delta(h,k) \le \eta} k(\mathcal{I})\,. \tag{3.11}$$

Property *UNI(a)* expresses that the Lebesgue measure of $J(h,\delta)$ is $\ll \delta$ when δ is scaled similarly to the maximal length of fundamental intervals of depth n. For any $\widehat{\rho} > \rho$, this length is $O(\widehat{\rho}^n)$ (up to absolute constants) and plays the role of a reference scale. This is a reformulation of the *UNI* condition implicit in Dolgopyat's Section 5 [16], which we (finally) state:

***UNI* Condition.** *A dynamical system of the good class, with a contraction ratio ρ, fulfills the UNI Condition if each inverse branch of T extends to a $\mathcal{C}^3$ function and*
(a) For any a $(0 < a < 1)$ we have $|J(h, \rho^{an})| \ll \rho^{an}\,, \forall n\,, \forall h \in \mathcal{H}^n\,.$
(b) $Q := \sup\{|\Psi''_{h,k}(x)|; n \ge 1\,, h,k \in \mathcal{H}^n, x \in \mathcal{I}\} < \infty.$

Note that *UNI* does not involve the cost: this is because c is constant on the monotonicity intervals of T.

Remark. Condition (b) follows from the existence of $\widetilde{Q} < \infty$ so that

$$|h'''(x)| \le \widetilde{Q}|h'(x)|\,, \forall n \ge 1\,, \forall h \in \mathcal{H}^n\,. \tag{3.12}$$

It suffices to check (3.12) for $n = 1$ (similarly as for the distorsion condition). (Note that this condition is always satisfied if there are finitely many $\mathcal{C}^3$ inverse branches).

Study of the L^2-norm: the close pairs.

Lemma 4. *Recall $A_{\sigma,\nu}$ from Lemma 1. Suppose that Condition UNI(a) holds. For any compact subset $\mathcal{L}$ of $\Sigma_0 \times W_0$, for all $(\sigma = \Re s, \nu = \Re w) \in \mathcal{L}$,* *for all n, for all a, with $0 < a < 1$, the integral I_n^- of the sum (3.6) restricted to pairs $(h,k) \in \mathcal{H}^n \times \mathcal{H}^n$ for which $\Delta(h,k) \le \rho^{an}$ satisfies*

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

Proof. Up to a positive constant that only depends on (σ,ν) (and is uniform on the compact subset $\mathcal{L}$), $|I_n^-|$ 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 the bounded distortion property, for all pairs $(h,k) \in \mathcal{H}^\star \times \mathcal{H}^\star$, up to multiplicative absolute constants, one has

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

Then, as in the beginning of the proof of Lemma 1, with the same bounded distortion property and using that $\mu_{\sigma,\nu}$ is an invariant probability for the normalized operator, the ratios (two by two) of the four quantities $a(h)$, $b(h)$, $c(h)$, $d(h)$,

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

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

admit upper and lower bounds which do not depend on h and are uniform when (σ,ν) varies in a compact set. 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 1, $\mu_{\sigma,\nu}[J] \le C_{\sigma,\nu} A_{\sigma,\nu}^n \; |J|^{1/2}$, which holds for any subset J that is a union of fundamental intervals of depth n, is applied to $J(h,\rho^{an})$. *UNI (a)* 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 relative to pairs (h,k) which were not treated by Lemma 4:

Lemma 5. *Suppose that Condition* UNI*(b) holds. Letting $\lceil x \rceil$ denote the smallest integer $\ge x$, set*

$$n_0 = n_0(t) = \left\lceil \frac{1}{|\log\rho|} \log|t| \right\rceil . \tag{3.13}$$

Then, for any compact subset $\mathcal{L}$ of $\Sigma_0 \times W_0$, for any $(\sigma = \Re s, \nu = \Re w) \in \mathcal{L}$, and $|t| = |\Im s| \geq 1/\rho^2$, for any $0 < a < 1/2$, the integral $I^+_{n_0}$ of the sum (3.6) for $n = n_0$, restricted to $(h,k) \in \mathcal{H}^{n_0} \times \mathcal{H}^{n_0}$ with $\Delta(h,k) \geq \rho^{an_0}$, satisfies

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

Proof. We start with a general integer n and bound

$$|I^+_n| \leq \frac{1}{\lambda(\sigma,\nu)^{2n}} \sum_{\substack{(h,k)\in\mathcal{H}^n\times\mathcal{H}^n \\ \Delta(h,k)\geq\rho^{an}}} \exp[\nu(c(h)+c(k))] \cdot \left|\widehat{I}(h,k)\right| \, , \tag{3.14}$$

where the integral $\widehat{I}(h,k) := \int_{\mathcal{I}} \exp[it\Psi_{h,k}(x)] \, R_{h,k}(x)dx$ involves $\Psi_{h,k}$, $R_{h,k}$ defined in (3.7), (3.8), and decomposes into four integrals of the form

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

with $r_{h,k}$ defined in (3.9). We shall apply the following lemma to each oscillatory integral $I(h,k)$:

Van der Corput Lemma (See e.g. [42]). *For each interval $\mathcal{I}$ and every $Q > 0$, there is $C(Q)$, so that for all $t \in \mathbb{R}$, $\Psi \in \mathcal{C}^2(\mathcal{I})$ with $|\Psi''(x)| \leq Q$, $|\Psi'(x)| \geq \Delta$ with $|t|^{-1} \leq \Delta \leq 1$, and $r \in \mathcal{C}^1(\mathcal{I})$ with $||r||_0 \leq R$, $||r||_{1,1} \leq RD$, the integral $I(t) = \int_{\mathcal{I}} \exp[it\Psi(x)] \, r(x) \, dx$ satisfies*

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

Consider (t, n) with $1/|t| \leq \rho^{an}$. Setting

$$M(h,k) := \sup_{x\in\mathcal{I}} |h'(x)|^\sigma |k'(x)|^\sigma \frac{1}{f^2_{\sigma,\nu}(x)} f_{\sigma,\nu} \circ h(x) f_{\sigma,\nu} \circ k(x) \, , \tag{3.15}$$

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

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

The arguments used in the proof of Lemma 2 for the function r_h apply to the function $r_{h,k}$, and

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

Then, by Property *UNI(b),* the Van Der Corput Lemma can be applied to each integral $I(h,k)$, which thus satisfies

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

Now, we choose $n = n_0$ as in (3.13). Since $a < 1/2$, and $\sqrt{|t|} \geq 1/\rho$ we have $n_0 \geq 2$ and $\rho^{-an_0} \leq \rho^{-(n_0-1)} \leq |t|$, so that the Van der Corput Lemma may be applied. The previous inequality becomes

$$|I(h,k)| \ll M(h,k) \; ||g||_{1,t}||\ell||_{1,t} \; \rho^{(1-2a)n_0} \, . \tag{3.16}$$

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

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

Now, take x_0 in $\mathcal{I}$. Then, from the bounded distortion property (3.1), and the definition of $M(h,k)$ in (3.15), we get

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

and therefore
(3.17)

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

From (3.16, 3.17, 3.14), we finally obtain $|I^+_{n_0}| \ll \rho^{(1-2a)n_0}\ ||f||^2_{1,t}$. □

Study of the L^2-norm: the final result. Consider the integer n_0 from (3.13) of Lemma 5 (for $|t| \geq 1/\rho^2$) and some a with $(2/5) < a < (1/2)$. Then, since $a/2 > (1-2a) > 0$, there exists a (real) neighborhood of $(\sigma,\nu) = (1,0)$ on which

$$A_{\sigma,\nu}\cdot\rho^{a/2} \leq \rho^{1-2a} \qquad \text{for any} \quad (\sigma,\nu)\in W\times\Sigma. \tag{3.18}$$

Then, from Lemmata 4 and 5,

$$\int_{\mathcal{I}}|\widetilde{\mathbf{H}}^{n_0}_{s,w}[f](x)|^2dx \ll \rho^{(1-2a)n_0}||f||^2_{1,t}\,. \tag{3.19}$$

3.3. **End of proof of Theorem 2.** We operate now the transfers between various norms.

From the L^2-norm to the sup-norm. Since the normalized density transformer $\widetilde{\mathbf{H}}_1$ is quasi-compact with respect to the $(1,1)$-norm, and fixes the constant function 1, it satisfies

$$||\widetilde{\mathbf{H}}^k_1[|g|^2]||_0 = \left(\int_{\mathcal{I}}|g|^2(x)\,dx\right) + O(r_1^k)||g^2||_{1,1}\,, \tag{3.20}$$

where r_1 is the subdominant spectral radius of $\mathbf{H}_1$.

Consider an iterate $\widetilde{\mathbf{H}}^n_{s,w}$ with $n \geq n_0$. 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 (3.5) from Lemma 1 and (3.20) with $k := n - n_0$, together with the bound (3.19) for the L^2-norm and finally Lemma 2 to evaluate $||g^2||_{1,1}$, one obtains

$$||\widetilde{\mathbf{H}}^n_{s,w}[f]||^2_0 \leq A^{2(n-n_0)}_{\sigma,w}\left[\rho^{(1-2a)n_0} + r_1^{n-n_0}\,|t|\right]||f||^2_{1,t}\,.$$

We now choose $n = n_1$ as a function of t so that the two terms $\rho^{(1-2a)n_0}$ and $r_1^{n-n_0}|t|$ are almost equal (with $n_0(t)$ defined in (3.13)):

$$n_1 = (1+\eta)n_0 \qquad \text{with} \qquad \eta := 2(1-a)\frac{\log\rho}{\log r_1} > 0\,. \tag{3.21}$$

Choose now d such that $0 < \eta(5a-2) < d < 1-2a < 1/5$ (which is possible if a is of the form $a = 2/5+\epsilon$, with a small $\epsilon > 0$). Recalling (3.18) where a first neighborhood was defined, and considering a (real) neighborhood $\Sigma\times W$ of $(1,0)$ for which

$$\sup\left[\lambda(\sigma,\nu)^{1+\eta}, A^{\eta}_{\sigma,\nu}\right] < \rho^{-\eta(5a/2-1)} < \rho^{-d/2}\,, \tag{3.22}$$

we finally obtain, for $n_1(t)$ and η defined in (3.21)

$$||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_0 \ll \rho^{n_1 b}\,||f||_{1,t}, \qquad \text{with} \qquad b := \frac{1-2a-d}{1+\eta}\,. \tag{3.23}$$

From the sup-norm to the $||.||_{1,t}$-norm. Applying Lemma 2 twice and using (3.23) yields the inequality

$$\begin{aligned} ||\widetilde{\mathbf{H}}^{2n_1}_{s,w}[f]||_1 &\ll |s|\,||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_0 + \rho^{n_1}\,||\widetilde{\mathbf{H}}^{n_1}_{s,w}[f]||_1 \\ &\ll |s|\,\rho^{n_1 b}||f||_{1,t} + \rho^{n_1}|t|\left(\frac{|s|}{|t|}||f||_0 + \rho^{n_1}\frac{||f||_1}{|t|}\right) \\ &\ll |t|\rho^{n_1 b}||f||_{1,t}\,, \end{aligned} \tag{3.24}$$

which finally entails for $n_2 = 2n_1$ (and $n_1(t)$ as above)

$$||\widetilde{\mathbf{H}}^{n_2}_{s,w}||_{1,t} \ll \rho^{n_2 b/2}\,. \tag{3.25}$$

The last step in Theorem 2. For fixed t with $|t| > 1/\rho^2$, any integer n can be written $n = kn_2 + \ell$ with $\ell < n_2(t)$. Then (3.25) and Lemma 3 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 $bn_2/2 = bn_1 = (1-2a-d)n_0$, with n_0 defined in (3.13), we finally obtain

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

$$\text{with}\;\; \xi := 1-2a-d, \qquad b := \frac{\xi}{1+\eta}, \qquad \gamma := \rho^{b/2} < 1, \qquad M_1 = \frac{M_0}{\rho^{\xi}}.$$

Then ξ is any value between 0 and 1/5. Therefore, returning to the operator $\mathbf{H}_{s,w}$, we have shown

$$||\mathbf{H}^{n}_{s,w}||_{1,t} \le M_1\cdot\gamma^n\cdot|t|^{\xi}\cdot\lambda(\sigma,\nu)^n\,, \quad \forall n\,,\forall|t|\ge 1/\rho^2\,. \tag{3.26}$$

Finally, for any $(\sigma,\nu)\in\Sigma\times W$ as in (3.22), one has

$$\gamma\lambda(\sigma,\nu) \le \rho^{\frac{\xi}{2(1+\eta)}}\cdot\rho^{-\frac{\xi}{4(1+\eta)}} = \rho^{\frac{\xi}{4(1+\eta)}} = \widehat{\gamma} < 1\,.$$

This proves Theorem 2 with $M := M_1/(1-\widehat{\gamma})$. □

3.4. ***UNI* Condition and Euclidean dynamical systems.** To apply Theorem 2 to our three algorithms, we prove that they satisfy *UNI.*

For two LFT's h_1 and h_2, with $h_i(x) = (a_ix+b_i)/(c_ix+d_i)$, we have

$$\Psi'_{h_1,h_2}(x) = \left|\frac{h''_1}{h'_1}(x) - \frac{h''_2}{h'_2}(x)\right| = \frac{|c_1d_2 - c_2d_1|}{|(c_1x+d_1)(c_2x+d_2)|},$$

so that the distortion property gives

$$\Delta(h_1,h_2) := \left|\frac{c_1}{d_1} - \frac{c_2}{d_2}\right| \cdot \inf_{x\in\mathcal{I}} \left|\frac{h'_1(x)h'_2(x)}{h'_1(0)h'_2(0)}\right|^{1/2} \geq \frac{1}{L}\left|\frac{c_1}{d_1} - \frac{c_2}{d_2}\right|. \tag{3.27}$$

Hence Δ only depends on the difference of the quotients c_i/d_i of the denominators of the LFT's. We shall next show that this difference is the "honest" (ordinary) distance between the rationals $h_1^*(0)$ and $h_2^*(0)$, where h^* is the *mirror LFT* of h, defined by:

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

This mirror operation appears in [41] where Schweiger relates it to the natural extension, and in [7], where the authors use the geometric notion of "folded" and "unfolded."

Clearly, the mirror map is an involution satisfying the morphism property $(h\circ k)^* = k^*\circ h^*$. It is not difficult to see that if $h \in \mathcal{H}^p$ is a LFT from a Euclidean dynamical system, associated to the sequence $\epsilon_0, (m_1,\epsilon_1), (m_2,\epsilon_2)\dots(m_p,\epsilon_p)$, the mirror LFT h^* corresponds to $\epsilon_p, (m_p,\epsilon_{p-1})\dots(m_{p-1},\epsilon_{p-2})\dots(m_1,\epsilon_0)$, i.e., the decomposition involves the same "digits" as h, but in the inverse order.

By (3.27), the distance $\Delta(h_1,h_2)$ between h_1 and h_2 indeed satisfies

$$\Delta(h_1,h_2) \geq \frac{1}{L}\left|h_1^*(0) - h_2^*(0)\right|. \tag{3.28}$$

It is not difficult to check for each of our three Euclidean algorithms that the set $\{h^*, h\in\mathcal{H}\}$ is the set of inverse branches of a dynamical system $(\mathcal{I}^*, T^*)$, the *dual dynamical system,* which belongs to the good class, with the same contraction ratio $\rho^* = \rho$ and a distortion constant L^*, as we now explain: For the Classical Euclidean algorithm $\mathcal{G}$, since all the ϵ are equal to 1 note that $(\mathcal{I}^*, T^*) = (\mathcal{I}, T)$. For the two others, the equality $\rho = \rho^*$ follows from the definition (2.4) of ρ, the distortion property for the dual system, and the fact that, in both systems the worse branch, for which the contraction ratio is attained on a fixed point, is a LFT h with $h = h^*$. The three dual dynamical systems are described in Figure 2.

Our next goal is to check *UNI* for our three algorithms. Note first that checking *UNI (b)* amounts to verifying (3.12), which is easy for the three algorithms. For *UNI(a)*:

Lemma 6. *A Euclidean dynamical system of the good class which admits a dual system that belongs to the good lass satisfies* UNI (a). *In particular, the dynamical systems associated to the Euclidean algorithms* $\mathcal{G}, \mathcal{K}$, *and* $\mathcal{O}$ *satisfy* UNI.

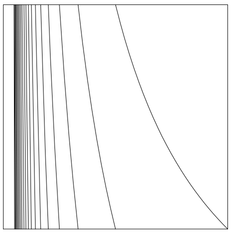
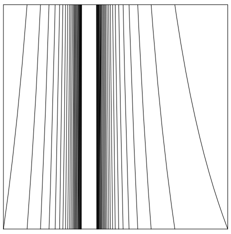
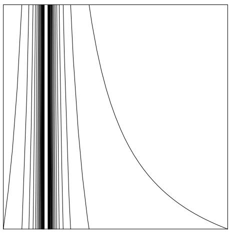
Algorithm	$\mathcal{G}^*$	$\mathcal{K}^*$	$\mathcal{O}^*$
Interval	$[0,1]$	$[\phi-2,\phi-1]$	$[\phi-2,\phi]$
Generic conditions	$m \geq 1\,,\epsilon=+1$	$m \geq 2\,,\epsilon=\pm 1$ if $m=2$ then $\epsilon=+1$	$m \geq 1$ odd, $\epsilon=\pm 1$ if $m=1$ then $\epsilon=+1$
Graphs			

Figure 2. The three dual Euclidean dynamical systems: Standard, Centered, Odd. ($\phi = \frac{1+\sqrt{5}}{2}$)

Proof. Fix $0 < a < 1$ and $h \in \mathcal{H}^n$. Denote by $J^*(h,\eta)$ the union of the intervals $k^*(\mathcal{I}^*)$ for LFT's $k \in \mathcal{H}^n$ satisfying $\Delta(h,k) \leq \eta$. First, we estimate $|J^*(h,\eta)|$, second, we relate $|J^*(h,\eta)|$ and $|J(h,\eta)|$.
By (3.28) if $\Delta(h,k) \leq \eta$ then $|h^*(0) - k^*(0)| \leq L\eta$. Since the length of a fundamental interval of depth n in the dual system is at most ρ^{an} (up to an absolute constant) we get

$$|J^*(h,\eta)| \ll 2L\eta + 2\rho^{an}\,. \tag{3.29}$$

The fundamental intervals $k(\mathcal{I})$ and $k^*(\mathcal{I}^*)$ (of depth n) have almost the same length: Indeed, by the bounded distortion property, and since $(k^*)'(0) = k'(0)$ for all $k \in \mathcal{H}^\star$,

$$\frac{1}{(LL^*)^{1/2}} \leq \frac{|k(\mathcal{I})|}{|k^*(\mathcal{I}^*)|} = \frac{|k(\mathcal{I})|}{|k'(0)|}\frac{|(k^*)'(0)|}{|k^*(\mathcal{I}^*)|} \leq (LL^*)^{1/2}\,. \tag{3.30}$$

Thus, since the intervals $k^*(\mathcal{I}^*)$ in $J^*(h,\eta)$ are disjoint, if $\eta \leq \rho^{an}$, using (3.30, 3.29)

$$|J(h,\eta)| \leq (LL^*)^{1/2}|J^*(h,\eta)| \ll (LL^*)^{1/2}(2L\eta + \rho^{an}) \ll \rho^{an}\,. \quad \square \tag{3.31}$$

3.5. **Aperiodicity and Condition *UNI*.** Theorem 2 provides estimates for the $(1,t)$-norm of $(I - \mathbf{H}_{s,w})^{-1}$ when $|t| = |\Im s|$ is sufficiently large. We now consider the case when $|t|$ is not large, and we explain how the *UNI* Condition intervenes in this context.

We first recall a classical aperiodicity result

Lemma 7. [Aperiodicity.] *Let* $\mathbf{H}_{s,w}$ *be the transfer operator associated to* $(\mathcal{I}, T, c)$ *of* $\mathcal{GMG}$ *type, denote by* $R(s,w)$ *its spectral radius, and consider a point* $(t_0, \tau_0) \neq (0,0)$. *The following are equivalent:*

(a) *There exists* $(\sigma, \nu) \in \Sigma_0 \times W_0$ *for which*

$$R(\sigma,\nu) \in \operatorname{Sp} \mathbf{H}_{\sigma+it_0, \nu+i\tau_0} .$$

(b) *There exists* $f \in \mathcal{C}^1(\mathcal{I})$, $|f| = 1$, *so that, for all* n, *for all* $h \in \mathcal{H}^n$

$$|h'(x)|^{it_0} \cdot \exp[i\tau_0 c(h)] \cdot f \circ h(x) = f(x) , \qquad \forall x \in \mathcal{I} . \tag{3.32}$$

Proof. This is a classical argument : see Prop. 6.2 in [36] or Prop. 9 in [44]. Since $\operatorname{Sp} \mathbf{H}_{\sigma+it_0,\nu+i\tau} = \lambda(\sigma,\nu) \operatorname{Sp} \widetilde{\mathbf{H}}_{\sigma+it,\nu}$ we may replace $R(\sigma+it, \nu+i\tau)$ and $R(\sigma,\nu)$ by the spectral radii of the corresponding normalized operators. By Proposition 0, the spectral radius of $\widetilde{\mathbf{H}}_{\sigma+it,\nu+i\tau}$ is at most $1 = R(\widetilde{\mathbf{H}}_{\sigma,\nu})$, while its essential spectral radius is at most $\widehat{\rho} < 1$. If (a) holds, then $\widetilde{\mathbf{H}}_{\sigma+it_0,\nu+i\tau_0}$ has an eigenvalue $\lambda = 1$ with an eigenfunction $f \in \mathcal{C}^1(I)$ with $\max_I |f| = 1$. Suppose that this maximum is attained at $x_0 \in \mathcal{I}$. Then the equality $\widetilde{\mathbf{H}}^n_{\sigma+it_0,\nu+i\tau_0}[f](x_0) = f(x_0)$ can be written as $\sum_{h\in\mathcal{H}^n} a_h b_h = 1$ with

$$a_h := \frac{1}{f_{\sigma,\nu}(x_0)} \exp[\nu c(h)] \cdot f_{\sigma,\nu} \circ h(x_0) \cdot |h'(x_0)|^{\sigma}$$

$$b_h := \frac{1}{f(x_0)} f \circ h(x_0) \exp[i\tau_0 c(h)] |h'(x_0)|^{it_0}$$

Normalization implies that $\sum_{h\in\mathcal{H}^n} a_h = 1$ while each factor b_h has modulus at most 1. Therefore, each b_h equals 1, and, for every $n \geq 1$,

$$f \circ h(x_0) = f(x_0) \exp[-i\tau_0 c(h)] |h'(x_0)|^{-it_0} , \ \forall h \in \mathcal{H}^n .$$

In particular, by density, $|f|$ is the constant function 1 and the above remarks also hold for any $x \in \mathcal{I}$. Then, for all $x \in I$, all $n \geq 1$, and all $h \in \mathcal{H}$,

$$\frac{f(x)}{f \circ h(x)} = \exp[i\tau_0 c(h)] |h'(x)|^{it_0} . \tag{3.33}$$

Following the arguments backwards, we get the other implication. □

This aperiodicity result, together with the *UNI* condition, provides some useful informations about the spectrum of $\mathbf{H}_{s,w}$. We recall the definition of lattice costs: a cost c is said to be lattice if it is not zero, and there exists $L > 0$ for which c/L is integer. The largest such L is the span of the cost.

Proposition 1. *Consider a* $\mathcal{GMG}$ *system satisfying* UNI(a) *and a point* $(t_0, \tau_0) \neq (0,0)$. *Then the following are equivalent;*

(a) *There exists* $(\sigma, \nu) \in \Sigma_0 \times W_0$ *for which*

$$R(\sigma,\nu) \in \operatorname{Sp} \mathbf{H}_{\sigma+it_0, \nu+i\tau_0} .$$

(b) $t_0 = 0$, *the cost* c *is a lattice cost with span* L, *and* $\tau_0 = 2k\pi/L$ *with* $k \in \mathbb{Z}_\star$.

In particular, for a $\mathcal{GMG}$ *system satisfying* UNI(a), *for any* $t \neq 0$, *1 does not belong to* $\operatorname{Sp} \mathbf{H}_{1+it,0}$. *Furthermore, if* c *is lattice with span* L, *for all* t, *and all* τ *not multiple of* $2\pi/L$, 1 *does not belong to* $\operatorname{Sp} \mathbf{H}_{1+it,\tau}$.

Proof. We first see that condition (b) of Lemma 7 for $t_0 \neq 0$ is incompatible with *UNI(a)*. If (b) holds, then taking the logarithm, differentiating and putting $\widehat{f} := \log f \in \mathcal{C}^1(\mathcal{I})$, we get

$$\Psi'_{h,k}(x) = \frac{i}{t_0}[h'(x)\widehat{f}' \circ h(x) - k'(x)\widehat{f}' \circ k(x)]\,, \forall n\,, \quad \forall h\,, k \in \mathcal{H}^n\,.$$

Then, for $\rho < \widehat{\rho} < 1$, the distance $\Delta(h,k)$ satisfies $\Delta(h,k) \ll \widehat{\rho}^n$ for any h, $k \in \mathcal{H}^n$, which contradicts *UNI(a)*.
Finally, assume *UNI (a)* and Condition (a) and $t_0 = 0$. Then relations of Condition (b) taken at the fixed points x_h of h imply that $\exp[i\tau_0 c(h)] = 1$ for all $h \in \mathcal{H}$. This is only possible if c is lattice of span L and τ_0 is a multiple of $2\pi/L$. □

3.6. **Final bounds for the Dirichlet series $S(s,w)$.** With Relation (2.17) which relates the Dirichlet series $S(2s,w)$ to the quasi-inverse of the operator $\mathbf{H}_{s,w}$, we obtain now the expected properties for the Dirichlet series $S(s,w)$. The first result is relative to the case when w is near 0, and will be useful in Section 4, while the second result is relative to the case when w belongs to a compact vertical segment, and will be useful in Section 5.

Lemma 8. *Consider one of the three algorithms $\mathcal{G}, \mathcal{K}, \mathcal{O}$, and a cost c of moderate growth. For any ξ with $0 < \xi < 1/5$, there is $\alpha_0 > 0$ and, for all $\widehat{\alpha}_0$, with $0 < \widehat{\alpha}_0 < \alpha_0$, there are a (complex) neighborhood $\mathcal{W}'$ of 0 and a constant M'' such that for any $w \in \mathcal{W}'$, the following holds:*
(i) $\Re\sigma(w) > 1 - (\alpha_0 - \widehat{\alpha}_0)$.
(ii) The (meromorphic) functions $s \mapsto S(2s,w)$, $s \mapsto \widetilde{S}(2s,w)$ have a single pole at $s = \sigma(w)$ in the strip $|\Re s - 1| \le \alpha_0$, and this pole is simple.
(iii) $\max\bigg(|S(s,w)|, |\widetilde{S}(s,w)|\bigg) \le M'' \max(1, |t|^\xi)\,, \forall s,\ \Re s = 1 \pm \alpha_0$.

Proof. Let $\mathcal{W}$ be a complex neighborhood of 0 in which $\sigma(w)$ from Proposition 0(7) is well-defined. Each vertical line $\Re s = \sigma$ is split into three:
Near the real axis. For (s,w) in a (complex) neighborhood $\mathcal{A}$ of $(1,0)$, Proposition 0(5) gives a decomposition $\mathbf{H}_{s,w} = \lambda(s,w)\,\mathbf{P}_{s,w} + \mathbf{N}_{s,w}$ where $\mathbf{P}_{s,w}$ is rank-one and the spectral radius of $\mathbf{N}_{s,w}$ is $\le \theta$. It is easy to see that the $(1,1)$-norm of $(I - \mathbf{N}_{s,w})^{-1}$ is bounded by some $\widehat{M_2}$ on $\mathcal{A}$. Since $\sigma(0) = 1$, taking a smaller neighborhood $\mathcal{W}_2 \subset \mathcal{W}$ of 0, there are $\alpha_2 > 0$ and $t_2 > 0$ so that the following set is contained in $\mathcal{A}$

$$\mathcal{A}_2 := \{(s,w); w \in \mathcal{W}_2, |\Re s - 1| \le \alpha_2,\ |\Im s| \le t_2\}\,.$$

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

$$(I - \mathbf{H}_{s,w})^{-1} = \frac{\lambda(s,w)}{1 - \lambda(s,w)}\,\mathbf{P}_{s,w} + (I - \mathbf{N}_{s,w})^{-1}\,. \tag{3.34}$$

It has as only singularities in $\mathcal{A}_2$ a simple pole at each point $(s = \sigma(w), w)$, with residue the nonzero operator

$$\mathbf{R}(w) := \frac{-1}{\lambda'_s(\sigma(w),w)}\,\mathbf{P}_{\sigma(w),w}\,. \tag{3.35}$$

Next, note that since $\lambda'_s(1,0) \neq 0$ we have $\ell := \inf_{\mathcal{A}_2}|\frac{\lambda(s,w)-1}{s-\sigma(w)}| > 0$. Fix α with $0 < \alpha \le \alpha_2$. Up to taking a smaller $\mathcal{W}_2$, we have $|\Re\sigma(w) - 1| \le \ell\alpha/2$

for $w \in \mathcal{W}_2$. Thus for $\Re s = 1 \pm \alpha$, $|t| \le t_2$, and $w \in \mathcal{W}_2$, the dominant eigenvalue satisfies

$$|\lambda(s,w) - 1| \ge \ell |s - \sigma(w)| \ge \ell |\Re s - \Re \sigma(w)| \ge \frac{\ell \alpha}{2},$$

and, by (3.34) we have $||(I - \mathbf{H}_{s,w})^{-1}||_{1,1} \le M_2/\alpha$.

Compact region. Suppose that $t \neq 0$. We first prove that $d(1, \operatorname{Sp} \mathbf{H}_{1+it,0}) > 0$. Fixing $\widehat{\rho} > \rho$, the spectrum $\operatorname{Sp} \mathbf{H}_{1+it,0}$ decomposes into two parts $\mathcal{S}^- = \operatorname{Sp} \mathbf{H}_{1+it,0} \cap \{|\lambda] \le \widehat{\rho}\}$ and $\mathcal{S}^+ = \operatorname{Sp} \mathbf{H}_{1+it,0} \cap \{|\lambda] > \widehat{\rho}\}$. Proposition 0(1) implies that $\mathcal{S}^+$ is a finite set of eigenvalues of finite multiplicity and Proposition 1 proves that 1 does not belong to $\mathcal{S}^+$. Then $d(1, \mathcal{S}^+) > 0$. On the other hand, $d(1, \mathcal{S}^-) \ge 1 - \rho$, and finally $d(1, \operatorname{Sp} \mathbf{H}_{1+it,0}) > 0$. Then, by perturbation theory of finite parts of the spectrum, there exist $\alpha_3 > 0$, $\beta > 0$ and a complex neighborhood $\mathcal{W}_3$ of 0 such that the distance between 1 and the spectrum of $\mathbf{H}_{s,w}$ is at least β on the compact set

$$\mathcal{A}_3 := \{(s,w);\ w \in \mathcal{W}_3, |\Re s - 1| \le \alpha_3,\ t_2 \le |t| \le t_0\},$$

where $t_0 = 1/\rho^2$ from Theorem 2. Thus $(s,w) \mapsto (I - \mathbf{H}_{s,w})^{-1}$ is analytic on the compact set $\mathcal{A}_3$ and its $(1,1)$-norm is bounded by some M_3 there.

Domain $|\Im s| \ge 1/\rho^2$. Consider $\Sigma_1 \times W_1$ from Theorem 2, for our fixed $\xi > 0$. There exist $\alpha_1 \in]0, \sigma(0) - \sigma_0]$ and a complex neighborhood $\mathcal{W}_1$ of $w = 0$ such that any pair in $\{(s,w);\ w \in \mathcal{W}_1, |\Re s - 1| \le \alpha_1\}$ has its real part (σ, ν) in $\Sigma_1 \times W_1$.

Choose first $\alpha_0 := \min(\alpha_2, \alpha_3, \alpha_1)$, next $\mathcal{W}_4$ so that $\Re \sigma(w) > 1 - (\alpha_0 - \widehat{\alpha}_0)$ on $\mathcal{W}_4$, and finally $\mathcal{W}' := \bigcap_{j=1}^4 \mathcal{W}_j$. Taking $M' := \max(M_2/\alpha_0, M_3, M_1)$, with M_1 as in Theorem 2, we obtain the claim for $S(s,w)$. For $\widetilde{S}(s,w)$, we apply (2.18) and use that $\zeta(s)$ is bounded on vertical strips near $\Re s = 2$. ☐

Lemma 9. *Consider one of the three algorithms $\mathcal{G}, \mathcal{K}, \mathcal{O}$, and a lattice cost c of moderate growth with span L. Consider some ξ with $0 < \xi < 1/5$, and $\upsilon > 0$. Then, there are $\gamma_1 > 0$, and a constant Q_3 such that, for any $\tau \in \mathbb{R}$, $|\tau| \ \in [\upsilon, \pi/L]$,*

(i) $s \mapsto S(s, i\tau)$ and $s \mapsto \widetilde{S}(s, i\tau)$ are analytic in the strip $|\Re s - 1| \le \gamma_1$.

(ii) $\max\Big(|S(s,i\tau)|, |\widetilde{S}(s,i\tau)|\Big) \le Q_3 \max(1, |t|^\xi), \forall s, \Re s = 1 \pm \gamma_1.$

Proof. Fix $0 < \xi < 1/5$. By *UNI*, Theorem 2 gives $\alpha > 0$ and Q_1 so that for $|\Im s| \ge 1/\rho^2$, $|\Re s - 1| \le \alpha$, and arbitrary real τ,

$$||(I - \mathbf{H}_{s,i\tau})^{-1}||_{1,t} \le Q_1 \cdot |\Im s|^\xi, \tag{3.36}$$

Suppose that $|t| \le 1/\rho^2$ and $|\tau| \in [\upsilon, \pi/L]$ We first prove, as in Lemma 8, that $d(1, \operatorname{Sp} \mathbf{H}_{1+it,i\tau}) > 0$. Proposition 0(1) implies that the spectrum $\operatorname{Sp} \mathbf{H}_{1+it,i\tau}$ decomposes into two parts $\mathcal{S}^- = \operatorname{Sp} \mathbf{H}_{1+it,i\tau} \cap \{|\lambda] \le \widehat{\rho}\}$ and $\mathcal{S}^+ = \operatorname{Sp} \mathbf{H}_{1+it,i\tau} \cap \{|\lambda] > \widehat{\rho}\}$, where $\mathcal{S}^+$ is a finite set of eigenvalues of finite multiplicity. Proposition 1 gives that 1 does not belong to $\mathcal{S}^+$. Then $d(1, \mathcal{S}^+) > 0$. On the other hand, $d(1, \mathcal{S}^-) \ge 1 - \rho$, and finally $d(1, \operatorname{Sp} \mathbf{H}_{1+it,i\tau}) > 0$. Then, by perturbation theory of finite parts of the spectrum, there are $0 < \gamma_1 \le \alpha$ and $\beta > 0$ such that the distance between 1 and the spectrum of $\mathbf{H}_{\sigma+it,i\tau}$ is at least β on the compact set

$|\sigma - 1| \le \gamma_1$, $|t| \le 1/\rho^2$, $|\tau| \in [\upsilon, \pi/L]$. Then $s \mapsto (I - \mathbf{H}_{s,i\tau})^{-1}$ is analytic there, and there is Q_2 so that

$$||(I - \mathbf{H}_{1\pm\gamma_1+it,i\tau})^{-1}||_{1,1} \le Q_2 \,, \forall |\tau| \in [\upsilon, \pi/L] \,, \forall |t| \le 1/\rho^2 \,. \quad \square \tag{3.37}$$

4. Limit Gaussian distribution for costs of moderate growth.

In this section we prove our Central Limit Theorem, Theorem 3. We first explain our use of Perron's formula (§4.1). We next introduce in §4.2 a smoothed model endowed with probability $\overline{\mathbb{P}}_N$. For this smoothed model, Lemma 10 allows us to deduce from bounds on $\overline{\Psi}_w$ the bounds on $\overline{\Phi}_w$ which entail quasi-power estimates. Theorem 0 gives asymptotic normality for the smoothed model, with asymptotic estimates for its expectation and variance (§4.3, in particular Lemma 12). A comparison of the uniform and smoothed distributions of (Lemma 14 in §4.4) finally yields Theorem 3.

4.1. **Using Perron's formula.** Choose $0 < \xi < 1/5$. Lemma 8 specifies $\alpha_0 \in]0, 1/2]$ and $\mathcal{W}'$. We fix $w \in \mathcal{W}'$, the Dirichlet series $S(s, w)$ being viewed as functions of s. Consider the strip $\mathcal{S}(w)$ limited by the two vertical lines $\Re s = 1 - \alpha_0$ and $\Re s = 1 + \alpha_0$. By Lemma 8, this strip contains $s = \sigma(w) \in \mathbb{R}$ as unique (simple) pole of $S(2s, w)$. In the rectangle $\mathcal{U}(w)$ defined by the strip $\mathcal{S}(w)$ and the two horizontal lines $\Im s = \pm U$, Cauchy's residue theorem provides

$$\frac{1}{2i\pi} \int_{\mathcal{U}(w)} S(2s, w) \frac{T^{2s+1}}{s(2s+1)} \, ds = \frac{E(w)}{\sigma(w)(2\sigma(w)+1)} T^{2\sigma(w)+1} \,, \tag{4.1}$$

where $E(w) = \mathbf{F}_{\sigma(w),w} \circ \mathbf{R}(w)[1](0)$, with $\mathbf{R}(w)$ the residue operator from (3.35). Note in particular that

$$E(0) = -\frac{1}{\lambda'(1)} \mathbf{F}_1 \mathbf{P}_1[1](0) = -\frac{1}{\lambda'(1)} \mathbf{F}_1[f_1](0) \neq 0 \,. \tag{4.2}$$

We now let U tend to ∞. By Lemma 8, the integral on the leftmost vertical line $\Re s = 1 - \alpha_0$ exists and satisfies

$$\int_{1-\alpha_0-i\infty}^{1-\alpha_0+i\infty} S(2s, w) \frac{T^{2s+1}}{s(2s+1)} \, ds = O\left(T^{3-2\alpha_0}\right)$$

(with a O-term which is uniform for $w \in \mathcal{W}'$ and $T \to \infty$), while the integrals on the horizontal lines of $\mathcal{U}(w)$ tend to zero for $U \to \infty$. Finally, Perron's formula (2.20) with $D = 1 + \alpha_0$ gives the contribution from the rightmost vertical side, so that

$$\Psi_w(T) = \frac{E(w)}{\sigma(w)(2\sigma(w)+1)} T^{2\sigma(w)+1} \left[1 + O\left(T^{-2\widehat{\alpha}_0}\right)\right] \,, \tag{4.3}$$

with $E(0) \neq 0$ and a uniform O-term with respect to $w \in \mathcal{W}'$, as $T \to \infty$.

4.2. **Smoothed costs and transfer of estimates.** To exploit the estimates (4.3) on the Cesàro sums $\Psi_w(N)$, we introduce an auxiliary model, the *smoothed model.*

Associate to some nonnegative function $T \mapsto \epsilon(T)$, with $\epsilon(T) \le 1$, the probabilistic models $(\overline{\Omega}_N(\epsilon), \overline{\mathbb{P}}_N(\epsilon))$ as follows: For any integer N, set $\overline{\Omega}_N(\epsilon) = \Omega_N$; next, choose uniformly an integer Q between $N - \lfloor N\epsilon(N) \rfloor$ and N, and draw uniformly an element (u, v) of Ω_Q. Slightly abusing language, we refer to the function C in the model $(\overline{\Omega}_N(\epsilon), \overline{\mathbb{P}}_N(\epsilon))$ as the "smoothed cost." The cumulative value of $\exp[wC]$ for $\overline{\mathbb{P}}_N(\epsilon))$ is

$$\overline{\Phi}_w(N) := \frac{1}{\lfloor N\epsilon(N)\rfloor} \sum_{Q=N-\lfloor N\epsilon(N)\rfloor}^{N} \sum_{n\le Q} c_n(w)\,, \tag{4.4}$$

so that the moment generating function of the smoothed cost is just

$$\overline{\mathbb{E}}_N[\exp(wC)] = \frac{\overline{\Phi}_w(N)}{\overline{\Phi}_0(N)}\,. \tag{4.5}$$

Note that $\overline{\Phi}_w$ can be expressed as a function of Ψ_w, via

$$\overline{\Phi}_w(N) = \frac{1}{\lfloor N\epsilon(N)\rfloor} \left[\Psi_w(N) - \Psi_w(N - \lfloor N\epsilon(N)\rfloor)\right]\,. \tag{4.6}$$

Now, to transfer the bound (4.3) for Ψ_w into a bound for $\overline{\Phi}_w$, we shall appeal to a result that is often used in number theory contexts.

Lemma 10. *Let $\mathcal{W}$ be a complex neighborhood of 0, and let $c_n(w)$ be a sequence of complex-valued functions on $\mathcal{W}$. Assume that $\Psi_w(T) := \sum_{n\le T} c_n(w)(T-n)$ satisfies*

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

with a O-error term which is uniform for $w\in\mathcal{W}$, where $F_w(T) = B(w)T^{a(w)}$ and $B(w)$, $a(w)$ are bounded holomorphic functions such that $\Re a(w) > 1$, $B(w)\neq 0$ on $\mathcal{W}$. Assume further that $G(T)$ tends to 0 for $T\to\infty$ and is of moderate variation, i.e., there exists K so that $|G(cT)|\le K|G(T)|$ for any c with $1/2\le c\le 2$.
Then, if $G(T)^{-1} = O(T)$ for $T\to\infty$, we have

$$\frac{1}{\lfloor TG(T)^{1/2}\rfloor}\left[\Psi_w(T) - \Psi_w\big(T - \lfloor TG(T)^{1/2}\rfloor\big)\right] = F'_w(T)\left[1 + O\big(G(T)^{1/2}\big)\right] \tag{4.7}$$

where the O-term is uniform with respect to $w\in\mathcal{W}$.

Proof. We first show (without using $G(T)^{-1} = O(T)$) that for $T\to\infty$

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

(without the integer parts) with a uniform O-term for $w\in\mathcal{W}$. Consider some sequence $\epsilon(T)$ which tends to 0. The estimate of $\Psi_w(T)$ and the assumption on G entail

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

Then our assumptions on $F_w(T)$ and $a(w)$ imply

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

with a uniform Θ [recall $A(T) = \Theta(B(T))$ as $T \to \infty$ means that there is an absolute constant $C > 0$ so that $A(T) \leq CB(T)$ and $A(T) \geq CB(T)$ as $T \to \infty$]. Therefore, we obtain (4.8) by taking

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

To finish, remark that the difference between (4.8) divided by $G(T)^{1/2}$ and (4.7) can be split into two terms of order

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

We now apply Lemma 10 to the smoothed costs. From (4.3), (4.6), upon setting $G(T) = T^{-2\widehat{\alpha}_0}$ (recall $0 < \widehat{\alpha}_0 \leq 1/2$), we find

$$\overline{\Phi}_w(N) = \frac{E(w)}{\sigma(w)} N^{2\sigma(w)} \left[1 + O\left(N^{-\widehat{\alpha}_0}\right)\right] ,$$

where $E(0) \neq 0$ and the O-term is uniform (as N tends to ∞) when w varies in a sufficiently small neighborhood of 0. For $w = 0$, one has

(4.9)
$$\overline{\Phi}_0(N) = E(0)\, N^2 \left[1 + O\left(N^{-\widehat{\alpha}_0}\right)\right], \qquad \text{with} \quad E(0) = -\frac{\mathbf{F}_1\mathbf{P}_1[1](0)}{\lambda'(1)} .$$

Finally, by (4.5), we obtain:

Lemma 11. [Quasi-powers for smoothed cost] *Let $0 < \gamma_0 < \alpha_0$ with α_0 from Lemma 8. The moment generating function of the smoothed cost corresponding to $\epsilon(N) = N^{-\gamma_0}$ satisfies*

$$\overline{\mathbb{E}}_N[\exp(wC)] = \frac{E(w)}{E(0)\sigma(w)} N^{2(\sigma(w)-\sigma(0))} \left[1 + O\left(N^{-\gamma_0}\right)\right] , \tag{4.10}$$

with $E(w)$ from (4.1) and a uniform O-term when $N \to \infty$ and w is near 0.

4.3. **Asymptotic Gaussian law for the smoothed cost.** In view of Lemma 11 which provides a quasi-power expression for the moment generating function of the smoothed cost, we may apply Theorem 0 to $\widehat{C}_N = C|_{\Omega_N}$ and $\widehat{\Omega}_N = \Omega_N$, $\widehat{\mathbb{P}}_N = \overline{\mathbb{P}}_N$ with $\beta_N := \log N$, $\kappa_N = N^{-\gamma_0}$, and

$$U(w) = 2(\sigma(w) - \sigma(0)), \qquad V(w) = \log \frac{E(w)}{E(0)\sigma(w)} . \tag{4.11}$$

We see from the above that function $\sigma(w)$ (which solves $\lambda(\sigma(w), w) = 1$) plays a central rôle. The next lemma expresses in particular that $U''(0) = 2\sigma''(0) > 0$, so that Theorem 0 can be applied:

Lemma 12. [Computation of constants.] *In the $\mathcal{GMG}$ setting we put $U(w) = 2(\sigma(w) - \sigma(0))$. Then, for the constant cost $c \equiv 1$ (recalling $\Lambda(s) = \log\lambda(s)$ from Proposition 0), one has*

$$\mu := U'(0) = \frac{2}{|\lambda'(1)|} = \frac{2}{|\Lambda'(1)|} , \qquad \delta^2 := U''(0) = \frac{2\Lambda''(1)}{|\Lambda'(1)|^3} > 0 . \tag{4.12}$$

More generally, recalling $\widehat{\mu}(c)$ *and* $\widehat{\delta}^2(c)$ *from Theorem 1, and setting* $\chi(c) = \Lambda''_{sw}(1,0)$*, we have*

$$\begin{aligned} \mu(c) &:= U'(0) = \mu \cdot \widehat{\mu}(c)\,, \\ \delta^2(c) &:= U''(0) = \widehat{\mu}^2(c) \cdot \delta^2 + \mu \cdot \widehat{\delta}^2(c) + \mu^2 \widehat{\mu}(c) \cdot \chi(c) > 0\,. \end{aligned}$$

In particular, $U(w)$ *(or equivalently* $\sigma(w)$*) is strictly convex at* 0.

Proof. Let us begin by the case when $c \equiv 1$. Then σ is defined by $\Lambda(\sigma(w)) = -w$. Therefore

$$\sigma'(w) = -\frac{1}{\Lambda'(\sigma(w))}, \qquad \sigma''(w) = \frac{\sigma'(w)\Lambda''(\sigma(w))}{\Lambda'^2(\sigma(w))},$$

and, recalling from Proposition 0 that $\Lambda'(1) < 0$, we get

$$2\sigma'(0) = \frac{2}{|\Lambda'(1)|}\,, \qquad 2\sigma''(0) = \frac{2\Lambda''(1)}{|\Lambda'(1)|^3}\,.$$

Let us now study the general case. Taking the derivatives of the relation $\Lambda(\sigma(w), w) = 0$, one obtains

$$\begin{aligned} 0 &= \sigma'(w)\Lambda'_s(\sigma(w), w) + \Lambda'_w(\sigma(w), w)\,, \\ 0 &= \sigma''(w)\Lambda'_s(\sigma(w), w) + \sigma'^2(w)\Lambda''_{s^2}(\sigma(w), w) \\ &\qquad\qquad + 2\sigma'(w)\Lambda''_{sw}(\sigma(w), w) + \Lambda''_{w^2}(\sigma(w), w)\,. \end{aligned}$$

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

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

Thus, setting $L(w) := \Lambda(1 + \sigma'(0)w, w)$, we find $\sigma''(0) = \frac{1}{|\Lambda'(1)|} L''(0)$. Since Proposition 0 (6.b) implies that $L''(0) > 0$, we get the strict positivity of $\delta^2(c)$, as claimed. Finally, $U'(0)$ and $U''(0)$ are

$$\mu(c) = 2\sigma'(0) = \frac{-2\Lambda'_w(1,0)}{\Lambda'(1)}\,, \tag{4.13}$$

$$\delta^2(c) = 2\sigma''(0) = \frac{2\Lambda'^2_w(1,0)\Lambda''(1)}{|\Lambda'(1)|^3} + \frac{4\Lambda'_w(1,0)\Lambda''_{sw}(1,0)}{|\Lambda'(1)|^2} + \frac{2\Lambda''_{w^2}(1,0)}{|\Lambda'(1)|}\,, \tag{4.14}$$

and, using (4.12) as well as Theorem 1, they may be expressed as functions of μ, δ, $\widehat{\mu}(c)$, $\widehat{\delta}(c)$, and $\chi(c)$. Note that $\chi(c) = \widehat{\delta}(c) = 0$ for constant c. ☐

By Lemma 12, Theorem 0 applies and provides the following result:

Lemma 13. *Let* $0 < \gamma_0 < \alpha_0$*, with* α_0 *the constant in Lemma 8. The smoothed cost* C *associated to* $\epsilon(N) = N^{-\gamma_0}$ *has an asymptotically Gaussian distribution, with speed of convergence* $O(1/\sqrt{\log N})$*. Moreover,*

$$\overline{\mathbb{E}}_N[C] = U'(0)\log N + V'(0) + O(N^{-\gamma_0})\,, \tag{4.15}$$

$$\overline{\mathbb{V}}_N[C] = U''(0)\log N + V''(0) + O(N^{-\gamma_0})\,, \tag{4.16}$$

with V *defined from the residue function* $E(w)$ *in (4.1) through (4.11), and* $U(w)$*,* $U'(0)$*, and* $U''(0) > 0$ *as in Lemma 12.*

Moreover, for each fixed $k \geq 3$, *there is a polynomial* P_k *of degree exactly* k *with coefficients depending on the derivatives of order at most* k *at* 0 *of* U *and* V, *so that the moment of order* k *satisfies*

$$\overline{\mathbb{E}}_N[C^k] = P_k(\log N) \ + O\left(\frac{(\log N)^{k-1}}{N^{\gamma_0}}\right) , \tag{4.17}$$

with a O*-term uniform in* k.

4.4. **Proof of Theorem 3.** The next lemma will allow us to deduce Theorem 3 from Lemmata 12 and 13:

Lemma 14. *The distance between the distributions* $\mathbb{P}_N$ *and* $\overline{\mathbb{P}}_N(\epsilon)$ *on* Ω_N *is* $O(\epsilon(N))$.

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

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

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

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

Then, the estimates $|\Omega_q| \sim (6/\pi^2)q^2$ entail

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

and give, for $(u,v) \in \mathcal{O}_N$,

$$\left|\,\overline{\mathbb{P}}_N(u,v) - \mathbb{P}_N(u,v)\right| \sim \left(\frac{N}{N'} - 1\right)\mathbb{P}_N(u,v) = \mathbb{P}_N(u,v)O(\epsilon(N))) .$$

Finally, for any $A \subset \Omega_N$, the difference $|\,\mathbb{P}_N(A) - \overline{\mathbb{P}}_N(A)|$ is less than

$$\left|\,\mathbb{P}_N(A\cap\mathcal{O}_N) - \overline{\mathbb{P}}_N(A\cap\mathcal{O}_N)\right| + \left|\,\mathbb{P}_N(A\cap\mathcal{E}_N) - \overline{\mathbb{P}}_N(A\cap\mathcal{E}_N)\right| = O(\epsilon(N)) ,$$

and the proof is completed. □

Proof of Theorem 3. For $\epsilon(N) = N^{-\gamma_0}$, Lemmata 13 and 14 imply that the asymptotic distribution of C is Gaussian, with a speed of convergence in $O(1/\sqrt{\log N})$. This proves Theorem 3(a) for any $\gamma \leq \gamma_0 < \alpha_0$.

To get claim (b) of Theorem 3, we use the fact that for each of our three algorithms, the maximal number of steps $P(u,v)$ on Ω_N is bounded by $K_0 \log N$. This is a classical result due to the fact that the three Euclidean divisions we consider are of the form $v = mu + \epsilon r$ with $r \leq v/2$. For $(u,v) \in \Omega_N$, all the quotients m_i are at most N, and the moderate growth assumption then implies that

$$\sup\{C^k(u,v); (u,v) \in \Omega_N\} \leq K(\log N)^{2k}$$

for some K and all N, and finally, Lemma 14 gives

$$\left|\mathbb{E}_N[C^k] - \overline{\mathbb{E}}_N[C^k]\right| = \ O\left((\log N)^{2k}\, N^{-\gamma_0}\right) .$$

Applying Lemma 12, we get the expression for $\mathbb{E}_N[C^k]$ and for the variance in Theorem 3(b), ending the proof of Theorem 3. □

5. Local Limit Theorem for lattice costs.

This section is devoted to proving Theorem 4. Consider an algorithm with associated dynamics in the good class and satisfying *UNI.* Let c be a lattice cost, of span $L > 0$, and of moderate growth. By Lemma 14, it suffices to consider the smoothed model $\overline{\mathbb{P}}_N$ associated to $\epsilon(N) = N^{-\gamma}$, for γ satisfying $0 < \gamma < \alpha_0$ (α_0 given by Lemma 8) which is to be fixed later. By Lemma 11, we have a quasi-power expression for $\overline{\mathbb{E}}_N[\exp(wC)]$ for small $|w|$.

We first consider costs c with span $L = 1$, so that

$$\overline{\mathbb{E}}_N[e^{i\tau C}] = \sum_{\ell \geq 0} \overline{\mathbb{P}}_N[C = \ell]\ e^{i\tau\ell}\,.$$

Our starting point is the relation

$$\overline{\mathbb{P}}_N[C(u,v) = \ell] = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\tau\ell}\cdot \overline{\mathbb{E}}_N[e^{i\tau C}]\,d\tau\,.$$

Since we are looking for an LLT result, the convenient scale of the problem is $n := \log N$, and we set

$$\ell = p_x(n) := \mu(c)n + \delta(c)x\sqrt{n}, \qquad q_x(n) := \lfloor p_x(n)\rceil\,.$$

(Here $\lfloor\cdot\rceil$ denotes the nearest integer function.) We consider

$$I_n := 2\pi\sqrt{\log N}\cdot\overline{\mathbb{P}}_N[C(u,v) = q_x(\log N)]$$

$$= \sqrt{n}\int_{-\pi}^{+\pi} \exp[-i\tau q_x(n)]\cdot\overline{\mathbb{E}}_N[e^{i\tau C}]\,d\tau\,. \tag{5.1}$$

Our strategy is to decompose the integration interval $[-\pi,+\pi]$ into a neighborhood of zero, $[-\upsilon,\upsilon]$, and its complement $|\tau| \in (\upsilon,\pi]$. This gives rise to two integrals $I_n^{(0)}$ and $I_n^{(1)}$.

5.1. **Estimate of the integral $I_n^{(0)}$.**

We begin with α_0 and a neighborhood $\mathcal{W}'$ as provided by Lemma 8. We assume that $[-\upsilon,+\upsilon]$ is contained in $\mathcal{W}'$ and study

$$I_n^{(0)} := \sqrt{n}\int_{-\upsilon}^{+\upsilon} \exp[-i\tau q_x(n)]\cdot\overline{\mathbb{E}}_N[e^{i\tau C}]\,d\tau\,.$$

Lemma 11 grants us a quasi-power expression for $\overline{\mathbb{E}}_N[\exp(i\tau C)]$. Recalling that $\mu(c) = 2\sigma'(0)$, we see that the integrand can be written as a product $g_n f_n$, where the first factor is

$$g_n(\tau) := \frac{E(i\tau)}{E(0)\sigma(i\tau)}\cdot\left(1 + O(e^{-n\gamma})\right)\cdot\exp[i\tau(p_x(n) - q_x(n))]\,;$$

and the second factor is

$$f_n(\tau) := \exp[2n(\sigma(i\tau) - 1 - i\tau\sigma'(0))]\cdot\exp[-i\tau x\delta(c)\sqrt{n}]\,.$$

The function $z \mapsto \sigma(z) - 1 - z\sigma'(0)$ has a saddle point at $z = 0$, and the main useful properties of $g_n(\tau), f_n(\tau)$ are

$$|g_n(\tau) - 1| \ll |\tau| + \exp[-n\gamma], \qquad |f_n(\tau)| \leq \exp[-n\tau^2\delta_0]\,, \tag{5.2}$$

for $|\tau| \le \upsilon$, where υ is taken sufficiently small, so that

$$\delta_0 := \inf\{|\Re\sigma''(i\tau)|; |\tau| \le \upsilon\} > 0\,. \tag{5.3}$$

We first prove that the dominant part of $I_n^{(0)}$ is

$$J_n^{(0)} := \sqrt{n}\int_{-\tau_n}^{\tau_n} \exp[-i\tau q_x(n)] \cdot \overline{\mathbb{E}}_N[e^{i\tau C}]\, d\tau, \qquad \text{with } \tau_n := \left(\frac{\log n}{\delta_0 n}\right)^{1/2}.$$

Properties (5.2) entail that

$$\sup\{|g_n(\tau) f_n(\tau)|;\ \tau_n \le |\tau| \le \upsilon\} \ll \exp[-n\tau_n^2\delta_0] \ll \frac{1}{n}\,,$$

and this implies that $|I_n^{(0)} - J_n^{(0)}|$ is $O(1/\sqrt{n})$.

We now define the following quantities,

$$\begin{aligned}
J_n^{(1)} &:= \sqrt{n}\int_{-\tau_n}^{\tau_n} \exp[-i\tau x\delta(c)\sqrt{n}] \cdot \exp[2n(\sigma(i\tau) - 1 - i\tau\sigma'(0))]\, d\tau\,,\\
J_n^{(2)} &:= \int_{-\tau_n\sqrt{n}}^{\tau_n\sqrt{n}} \exp[-i\vartheta\delta(c)x - \frac{1}{2}\vartheta^2\delta^2(c)]\, d\vartheta\,,\\
J_n^{(3)} &:= \int_{-\infty}^{\infty} \exp[-i\vartheta\delta(c)x - \frac{1}{2}\vartheta^2\delta^2(c)] d\vartheta = \frac{\sqrt{2\pi}}{\delta(c)} e^{-x^2/2}\,,
\end{aligned}$$

and will show that they satisfy $|J_n^{(i)} - J_n^{(i+1)}| = O(1/\sqrt{n})$ for $1 \le i \le 3$. Let us check this line by line.

From $J_n^{(0)}$ *to* $J_n^{(1)}$. Properties (5.2) imply that

$$\frac{1}{\sqrt{n}}|J_n^{(1)} - J_n^{(0)}| \ll \int_{-\tau_n}^{\tau_n} |\tau| \exp[-n\tau^2\delta_0] d\tau + e^{-n\gamma}\int_{-\tau_n}^{\tau_n} \exp[-n\tau^2\delta_0] d\tau\,,$$

and the two terms are $O(1/n)$.

From $J_n^{(1)}$ *to* $J_n^{(2)}$. There is first a change of variable $\vartheta = \tau\sqrt{n}$. Remark that

$$\vartheta^3/\sqrt{n} = \tau^3 n = O\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)$$

tends to 0 for $n \to \infty$. Furthermore, since $\delta^2(c) = 2\sigma''(0)$,

$$2n\left[\sigma(i\,\frac{\vartheta}{\sqrt{n}}) - \sigma(0) - i\,\frac{\vartheta}{\sqrt{n}}\sigma'(0)\right] = -\vartheta^2\frac{\delta^2(c)}{2} + O\left(\frac{\vartheta^3}{\sqrt{n}}\right)\,,$$

so that the difference between the two integrals satisfies

$$|J_n^{(1)} - J_n^{(2)}| \ll \frac{1}{\sqrt{n}}\int_{-\tau_n\sqrt{n}}^{\tau_n\sqrt{n}} \vartheta^3 \exp[-\vartheta^2\frac{\delta^2(c)}{2}] d\vartheta = O\left(\frac{1}{\sqrt{n}}\right)\,.$$

From $J_n^{(2)}$ *to* $J_n^{(3)}$. The difference between the two integrals satisfies

$$|J_n^{(2)} - J_n^{(3)}| \le \int_{|\tau| \ge \tau_n\sqrt{n}} \exp[-\vartheta^2\frac{\delta^2(c)}{2}] d\vartheta = O\left(\frac{1}{\sqrt{n}}\right)\,.$$

The last equality for $J_n^{(3)}$ is obtained by “completing” the square.

Finally, the first part $I_n^{(0)}$ of the integral I_n satisfies

$$I_n^{(0)} = \sqrt{2\pi}\,\frac{e^{-x^2/2}}{\delta(c)} + O\left(\frac{1}{\sqrt{n}}\right). \tag{5.4}$$

5.2. **Estimate of the integral when τ is not close to 0.** We now consider the integral (5.1) outside of the neighborhood of zero.

Lemma 15. *Assume that c is a lattice cost of moderate growth with span L. For each $\upsilon > 0$, there are $Q < \infty$ and $\gamma_0 > 0$ so that, for the smoothed cost corresponding to $\epsilon(N^{-\gamma})$ with small enough γ, we have*

$$|\overline{\mathbb{E}}_N[\exp(i\tau C)]| \le QN^{-\gamma_0}\,, \forall |\tau| \in [\upsilon, \pi/L]\,.$$

Proof. We apply Lemma 9 with Perron's formula, keeping the notations of Lemma 9. Consider, as in Section 4.1, Perron's formula (2.20) with $D = 1+\gamma_1$,

$$\Psi_{i\tau}(T) = \int_{\Re s = 1+\gamma_1} S(2s, i\tau)\frac{T^{2s+1}}{s(2s+1)}\,ds$$

In the rectangle $\mathcal{U}$ defined by the strip $\Re s = 1 \pm \gamma_1$ and the horizontal lines $\Im s = \pm U$, the function $s \mapsto S(2s, i\tau)$ is analytic (from Lemma 9) and Cauchy's residue theorem gives

$$\int_{\mathcal{U}} S(2s, w)\frac{T^{2s+1}}{s(2s+1)}\,ds = 0.$$

Furthermore, using once more Lemma 9,

$$\int_{\Re s = 1-\gamma_1} S(2s, w)\frac{T^{2s+1}}{s(2s+1)}\,ds = O\left(T^{3-2\gamma_1}\right)\,.$$

We let U tend to ∞. The integrals on the horizontal lines of $\mathcal{U}$ tend to zero, Perron's formula gives the right side, and finally

$$\Psi_{i\tau}(T) = O(T^{3-2\gamma_1})\,. \tag{5.5}$$

Now, if $\gamma < 2\gamma_1$, setting $\gamma_0 = 2\gamma_1 - \gamma$, the smoothed quantity $\overline{\Phi}_{i\tau}(N)$ satisfies

$$\begin{aligned}\overline{\Phi}_{i\tau}(N) &= \frac{1}{\lfloor N^{1-\gamma}\rfloor}\left(\Psi_{i\tau}(N) - \Psi_{i\tau}(N - \lfloor N^{1-\gamma}\rfloor\right) \\ &= O(N^{\gamma-1})O(N^{3-2\gamma_1}) = O(N^{2-\gamma_0})\,. \end{aligned} \tag{5.6}$$

Furthermore, with (4.9), $\overline{\Phi}_0(N) = \Theta(N^2)$. Therefore

$$\overline{\mathbb{E}}_N[\exp(i\tau C)] = \frac{\overline{\Phi}_w(N)}{\overline{\Phi}_0(N)} = O(N^{-\gamma_0})\,. \quad \square$$

We deduce from Lemma 15 that $I_n^{(1)}$ is $O(e^{-n\gamma})$.

5.3. **End of proof of Theorem 4.** Finally, we take α_0 and a neighborhood $\mathcal{W}'$ as granted by Lemma 8. We assume that $[-\upsilon, +\upsilon]$ is contained in $\mathcal{W}'$ and is sufficiently small to entail validity of (5.3). Then, υ is fixed and we choose γ so that γ is less than $2\gamma_1$ (from Lemma 9) and less than α_0 (from Lemma 8).
Now, with (5.4) and Lemma 11, Theorem 4 is proven in the case of span $L = 1$. Remark that if c has span L, then $d := c/L$ has span 1, with $\mu(c) = L\mu(d)$ and $\delta(c) = L\delta(d)$. This proves Theorem 4 in the general case. □

5.4. **The nonlattice case.** We say that a nonlattice digit-cost is *shifted lattice* if there exist $L \in \mathbb{R}_*^+$ and a number L_0, with L_0/L irrational, so that $(c - L_0)/L$ is integer-valued. The number L_0 is called the *shift.* The span of c is the largest possible L. Then, the digit cost c can be written as the sum of two costs $c = L_0 + d$, where d is lattice with span L, and the total cost C equals $c = L_0 P + D$. It is then possible to apply the previous result to D. Furthermore, since the two Λ functions relative to c and d are related by $\Lambda_c(s,w) = wL_0 + \Lambda_d(s,w)$, one has

$$\mu(d) = \mu(c) - \mu L_0, \qquad \delta^2(d) = \delta^2(c) + L_0^2\delta^2 - L_0[\chi(c)\mu^2 + 2\frac{\mu(c)}{\mu}\delta^2]\,,$$

and, finally:
For the algorithms $\mathcal{G}$, $\mathcal{K}$, $\mathcal{O}$ and for any shifted lattice cost c of span L and shift L_0, which is of moderate growth, one has, for any $x \in \mathbb{R}$

$$\mathbb{P}_N\left[(u,v); -\frac{L}{2} < C(u,v) - L_0 \cdot P(u,v) - \mu(d)\log N - \delta(d)x\sqrt{\log N} \le \frac{L}{2}\right]$$
$$= \frac{e^{-x^2/2}}{\delta(d)\sqrt{2\pi\log N}} + O\left(\frac{1}{\log N}\right),$$

with a O-term uniform in x.

The case when $c \not\equiv 0$ is neither lattice nor shifted lattice is essentially different, the noncompact situation requiring Dolgopyat-type estimates in w. The generic nonlattice situation will be considered in a forthcoming paper.

6. CONCLUSION.

This article has presented a unified approach to a large body of results relative to Euclid's algorithm and its major (fast) variants. We recapitulate here some of its consequences, relations to the vast existing literature on the subject, as well as extensions and open questions. It should be stressed that most of the improvements can eventually be traced to the existence of pole-free strips for Dirichlet series of cost parameters, a fact that precisely devolves from our extension of Dolgopyat's estimates to continued fraction systems.

First our methods lead to extremely precise estimates of the moments of costs, and in particular the number of steps, a much studied subject. Our estimates, when specialized to the mean number of steps, yield

$$\mathbb{E}_N[C] = \mu\log n + \eta + O(N^{-\gamma}); \tag{6.1}$$

see above Theorem 3, Parts (*b*) and (*c*). In the case of the standard algorithm, this covers the original estimates of Dixon and Heilbronn in 1969–1970 (the main term), as well as Porter's 1975 improvement (the second term, η), while providing the right shape of the error term ($O(N^{-\gamma})$), for which Porter further showed that one could take $\gamma = \frac{1}{6} - \epsilon$ in the case of the standard algorithm. We refer the reader to the accounts by Knuth [30] and Finch [18] for more material on this classical topic. Our formula (6.1) also extends Rieger's analyses (first published around 1980, see [38, 39]) of the centered algorithm and the odd algorithm. Note that, in our perspective, the second-order constant η comes out as a spectral quantity. It is an open problem to derive an explicit form starting from our expressions (e.g., such a form involving $\zeta'(2)$ is known under the name of Porter's constant in the standard case).

In sharp contrast to the mean value case, variances do not seem to be amenable to elementary methods. The first-order estimate has been given by Hensley (1994) in the paper [24] that first brought functional analysis to the field. Our formula,

$$\mathbb{V}_N[C] = \delta^2 \log n + \delta_1 + O(N^{-\gamma}), \tag{6.2}$$

stated in Theorem 3 can be viewed as a precise form of Hensley's estimate relative to the standard algorithm, one that also extends to the odd and centered algorithms. (Incidentally, the quantity δ, called Hensley's constant [18], is not recognized to be related to classical constants of analysis, though it is now known to be polynomial-time computable thanks to a recent study of Lhote [32]; the nature of δ_1 is even more obscure.) Note that the complex-analytic properties of the moment generating functions provided by the functional-analytic methods furnish similarly strong estimates for moments of arbitrarily high order (our Theorem 3), a fact which also appears to be new.

Regarding distributional results, several points are worthy of note. Dixon in his 1970 paper had already obtained exponential tail bounds and these were further improved, albeit in a still fairly qualitative way by Hensley in his 1994 study (see his Theorem 1 in [24]). Our approach gives much more precise information on the distribution, as we now explain.

For simplicity, let us specialize once more the discussion to the number of steps. First, regarding the central region of the Local Limit Theorem (Theorem 4), the nature of the error terms obtained ($O(N^{-\gamma_0})$) and the fact that the saddle point method lends itself to the derivation of full asymptotic expansions (see, e.g., Henrici's book [23]) entail the existence of a *full asymptotic expansion* associated with the Gaussian approximation of Theorem 4, namely, for a computable numeric sequence $\{c_j\}$,

$$\mathbb{P}_N[\cdot] = \frac{e^{-x^2/2}}{\delta\sqrt{2\pi \log n}} \left(1 + \sum_{k=1}^{r} \frac{c_r}{(\log N)^{r/2}} + O\left(\log^{-(r+1)/2} N\right)\right) \tag{6.3}$$

(the argument of $\mathbb{P}_N$ being that of Theorem 4). This expansion is valid for x in any compact set of $\mathbb{R}$.

Regarding large deviations, Lemma 11 (quasi-powers for smoothed costs) implies that $\overline{\mathbb{E}}_N[\exp(wC)]$ is of the form of a quasi-power, provided w

stays in a small enough fixed neighbourhood of 0. By a standard argument (originally due to Cramér and adapted to the quasi-powers framework by Hwang [25, 26, 27]) and Lemma 14, this implies the existence of a large deviation rate function, that is, of a function $I(y)$ such that, for the right tail corresponding to $y \geq 0$, there holds

$$\lim_{N\to\infty} \frac{1}{\log N} \log \mathbb{P}_N[C > (1+y)\mu \log N] = -I(y), \tag{6.4}$$

with $I(y)$ defined on an interval $[0, \delta]$ for some $\delta > 0$. In other words, the probability of exceeding the mean by $y\mu \log N$ is exponentially small in the scale of the problem, being of the form $\exp[-I(y)\log N] = N^{-I(y)}$. A simplified version of (6.4) is then: the probability of observing a value at z or more standard deviations from the mean is bounded by a quantity that decays like $e^{-C_1 z^2}$ (Hensley's notations [24]). Analogous properties naturally hold for the left tail. Cramer's technique of "shifting the mean" and use of the resulting "shifted" quasi-powers lead in fact to very precise quantitative versions of (6.4) in the style of Hwang—thereby providing optimal forms of tail inequalities *à la* Dixon–Hensley.

Similar properties hold for other costs measures, like the ones detailed in the introduction. Of particular interest is the statistics of the number of digits assuming a particular value, for which estimates parallel to (6.1), (6.2), (6.3), and (6.4) are also seen to hold. For instance, the frequency of occurrence of digit m in the expansion of a rational number has mean

$$\sim \log_2\left(1 + \frac{1}{m(m+2)}\right),$$

and it exhibits Gaussian fluctuations. It is again unknown to us, which of the involved constant (beyond the mean value factor) may be related to classical constants of analysis. The spectral forms of our Lemma 12 may at least provide a starting point for such investigations.

A major challenge is to derive distributional information that are so to speak "superlocal". By this, we mean the problem of estimating the behaviour of the number of steps and other cost measures over fractions of denominator *exactly* equal to N, i.e., rationals belonging to $\Omega_N \setminus \Omega_{N-1}$. In view of what is known in the average-case [18, 30], we expect arithmetical properties of N to come into the picture. The analytical difficulty in this case stems from the fact that a further level of "unsmoothing" (how to go from Ω_N to $\Omega_N \setminus \Omega_{N-1}$?) would be required.

Further works. Our dynamical approach provides a general framework, where it seems possible to answer other questions about distributional analysis. For instance, all the questions that we solve for rational trajectories can be asked for the periodic trajectories. A periodic trajectory is produced by a quadratic number, and the reference parameter is related to the length of the geodesics on the modular surface. In a forthcoming paper, and following the same principles as in [36, 37, 46], we prove that the distribution of costs on periodic trajectories can be studied in a similar way as here, replacing the quasi–inverse $(I - \mathbf{H}_{s,w})^{-1}$ by $\det(I - \mathbf{H}_{s,w})$.

Open problems. We also ask various questions about Euclidean algorithms: for instance, what happens for other Euclidean algorithms of the Fast Class (in particular for the Binary algorithm [8, 45])? The extension of our results to cost functions that are still "small" but take into account the boolean cost (also known as the bit complexity) of each arithmetic operation is on our agenda. Note that an average-case analysis is already known to be possible via operator techniques [2, 48]. On another register, the extension to "large" costs is likely to lead us to the realm of stable laws: see for instance Gouezel and Vardi's works [21, 49] for occurrences of these laws in continued fraction related matters.

Acknowledgements. During his thesis, Hervé Daudé made experiments providing evidence for the Gaussian limit property of the number of steps. Joint work with Eda Cesaratto [10] involved an extensive use of the weighted transfer operator, and some of the ideas that we shared on that occasion proved very useful for the present paper. Finally, we have had many stimulating discussions with Philippe Flajolet regarding the saddle-point method and the notion of smoothed costs. This work has been supported in part by two CNRS MATHSTIC grants and by the European Union under the Future and Emerging Technologies programme of the Fifth Framework (ALCOM-FT Project IST-1999-14186).

REFERENCES

[1] AARONSON, J., AND DENKER, M., Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps, *Stoch. Dyn.* 1 (2001) 193–237

[2] AKHAVI, A. AND VALLÉE, B., Average bit–complexity of Euclidean Algorithms, Proceedings ICALP'00, *Lecture Notes Comp. Science* 1853, 373–387, Springer, 2000.

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

[4] BALADI, V., *Positive Transfer operators and decay of correlations,* Advanced Series in non linear dynamics, World Scientific, 2000.

[5] BALADI, V., AND VALLÉE, B., Distributional analyses of Euclidean algorithms, *Proceedings ANALCO04,* submitted 2003.

[6] BEDFORD, T., KEANE, M., AND SERIES, C., Eds., *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces,* Oxford University Press, 1991.

[7] BOURDON, J. DAIREAUX, B., AND VALLÉE, B., Dynamical analysis of α-Euclidean algorithms, *Journal of Algorithms* 44 (2002) 246–285.

[8] BRENT, R.P. Analysis of the Binary Euclidean algorithm, *Algorithms and Complexity, New Directions and Recent Results,* J.F. Traub, ed., Academic Press, 1976.

[9] BROISE, A., Transformations dilatantes de l'intervalle et théorèmes limites, *Astérisque* 238, 5–109, Société Mathématique de France, 1996.

[10] CESARATTO, E., AND VALLÉE, B. Hausdorff dimension of sets defined by constraints on digits related to large additive costs, Les Cahiers du GREYC, 2003, submitted.

[11] CHERNOV, N. *Markov approximations and decay of correlations for Anosov flows,* Ann. of Math. (2) 147 (1998) 269–324.

[12] 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.

[13] CHAZAL, F., MAUME, V., Statistical properties of General Markov dynamical sources: applications to information theory, Preprint Université de Bourgogne, 2003, to appear in *DMTCS.*

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

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

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

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

[18] Finch, S. R., *Mathematical Constants*, Cambridge University Press, 2003.

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

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

[21] Gouezel, S. Central limit theorem and stable laws for intermittent maps, *Prob. Theory and Applications* (2003), to appear.

[22] Heilbronn, H., On the average length of a class of continued fractions, *Number Theory and Analysis,* ed. by P. Turan, 87–96, Plenum, New York, 1969.

[23] Henrici, P., *Applied and Computational Complex Analysis*, vol. 2, John Wiley, 1974.

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

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

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

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

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

[29] Khinchin, A. I., *Continued Fractions,* University of Chicago Press, Chicago, 1964.

[30] Knuth, D.E., *The Art of Computer programming,* Volume 2, 3rd edition, Addison Wesley, Reading, Massachusetts, 1998.

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

[32] Lhote, L., Modélisation et approximation de sources complexes, Mémoire de DEA, Université de 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) 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., 175–222, Oxford University Press, 1991.

[35] Nagaev, S.V., Some limit theorems for stationary Markov chains, *Theor. Probab. Appl.* 2 (1957) 378–406.

[36] Parry, W., and Pollicott, M., *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics,* Astérisque, 1990.

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

[38] Rieger, G. J., Über die mittlere Schrittanzahl bei Divisionalgorithmen, *Math. Nachr.* (1978) 157–180.

[39] Rieger, G. J.,, Über die Schrittanzahl beim Algorithmus von Harris und dem nach nächsten Ganzen, *Archiv der Mathematik* 34 (1980), 421–427.

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

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

[42] Stein, E., *Harmonic analysis: Real Variable Methods, Orthogonality and Oscillatory integrals,* Princeton University Press, Princeton, 1993.

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

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

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

[46] Vallée, B., Dynamique des fractions continues à contraintes périodiques, *Journal of Number Theory* (1998) 2, 183–235.
[47] Vallée, B., Dynamical Analysis of a Class of Euclidean Algorithms, *Theoretical Computer Science* vol 297/1-3 (2003) 447–486.
[48] Vallée, B., Digits and Continuants in Euclidean Algorithms. Ergodic Versus Tauberian Theorems, *Journal de Théorie des Nombres de Bordeaux* 12 (2000) 531–570.
[49] Vardi, I., A relation between Dedekind sums and Kloosterman sums, *Duke Math. J.* 55 (1987), 189–197.
[50] 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 de Mathématiques de Jussieu, F-75251 Paris, France
E-mail address: `baladi@math.jussieu.fr`

Brigitte Vallée: CNRS UMR 6072, GREYC, Université de Caen, F-14032 Caen, France
E-mail address: `brigitte.vallee@info.unicaen.fr`