The Euclid algorithm is "totally" gaussian

Brigitte Vallée

Laboratoire GREYC*, CNRS and University of Caen, France*

We consider Euclid's gcd algorithm for two integers (p, q) with $1 \leq p \leq q \leq N$, with the uniform distribution on input pairs. We study the distribution of the total cost of execution of the algorithm for an additive cost function d on the set of possible digits, asymptotically for $N \to \infty$. For any additive cost of moderate growth d, Baladi and Vallée obtained a central limit theorem, and –in the case when the cost d is lattice– a local limit theorem. In both cases, the optimal speed was attained. When the cost is non lattice, the problem was later considered by Baladi and Hachemi, who obtained a local limit theorem under an intertwined diophantine condition which involves the cost d together with the "canonical" cost c of the underlying dynamical system. The speed depends on the irrationality exponent that intervenes in the diophantine condition. We show here how to replace this diophantine condition by another diophantine condition, much more natural, which already intervenes in simpler problems of the same vein, and only involves the cost d. This "replacement" is made possible by using the additivity of cost d, together with a strong property satisfied by the Euclidean Dynamical System, which states that the cost c is both "strongly" non additive and diophantine in a precise sense. We thus obtain a local limit theorem, whose speed is related to the irrationality exponent which intervenes in the new diophantine condition. We mainly use the previous proof of Baladi and Hachemi, and "just" explain how their diophantine condition may be replaced by our condition. Our result also provides a precise comparison between the rational trajectories of the Euclid dynamical system and the real trajectories.

Keywords: Euclid algorithm, gcd, Gaussian limit law, dynamical system

Introduction

Every $x \in]0,1]$ admits a finite or infinite continued fraction expansion of the form

$$x = \cfrac{1}{m_1 + \cfrac{1}{m_2 + \cfrac{1}{\ddots + \cfrac{1}{m_n + \ldots}}}}\,. \tag{1}$$

Continued fraction expansions can be viewed as trajectories of a dynamical system defined by the Gauss map

$$T : [0,1] \to [0,1], \qquad T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, \ \text{ for } x \neq 0, \ \ T(0) = 0\,.$$

(Here, $\lfloor x \rfloor$ is the integer part of x.) For an irrational x, the trajectory $\mathcal{T}(x) = (x, T(x), \ldots, T^n(x), \ldots)$ never meets 0 and is encoded by the infinite sequence of *digits*

$$\mathcal{M}(x) := (m_1(x), m_2(x), m_3(x), \ldots, m_n(x), \ldots), \qquad \text{with} \quad m_i(x) := \left\lfloor \frac{1}{T^{i-1}(x)} \right\rfloor.$$

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.

If we associate a real value $d(m)$ to each possible digit $m \geq 1$, we say that d is the digit cost, and we may also regard d as a function on the set $\mathcal{H}$ of inverse branches of T. A digit-cost d is of moderate growth if one has $d(m) = O(\log m)$ for $m \to \infty$.

Real truncated trajectories. For an input x, and an integer $n \geq 1$, we consider the truncated trajectory $\mathcal{M}_n(x)$, and its cost $D_n(x)$

$$\mathcal{M}_n(x) := (m_1(x), m_2(x), \ldots, m_n(x)), \qquad D_n(x) := \sum_{i=1}^{\min(n, P(x))} d(m_i(x)).$$

We fix a reference probability measure on $[0, 1]$, with a smooth density f, and we wish to study the asymptotic distribution of the cost D_n, as $n \to \infty$. The main results are obtained in [9], [6] and [1] and can be described as follows: There is a Central Limit Theorem (CLT) for the cost D_n, and also a Local Limit Theorem (LLT), at least for a lattice digit-cost d (We recall that a cost d is lattice if there exist $d_0, L \in \mathbb{R}$ with $L > 0$, such that any value $d(m)$ satisfies $d(m) = d_0 + L_m \cdot L$ for some integer L_m). In both cases (CLT and LLT), the optimal speed $O(1/\sqrt{n})$ is attained. However, when the cost d is not lattice, the situation appears to be much less precise for the LLT : there exists a LLT for the costs D_n, but the speed of convergence is not precisely studied. We study this question here.

Rational trajectories. An execution of the Euclid algorithm on the input (p, q) formed with two integers p, q such that $p/q = x$ 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

$$\widehat{D}(x) := \sum_{i=1}^{P(x)} d(m_i(x)). \tag{2}$$

The natural reference parameter is no longer the truncation degree n, but the size $N := \max(p, q)$ of the input (p, q), and the probability measure $\mathbb{P}_N$ is now the uniform discrete measure on the (finite) set of inputs of size at most N. We also wish to study the asymptotic distribution of costs $\widehat{D}_N$, as $N \to \infty$, and compare it to the distribution of costs D_n, for $\log N \approx n$. The discrete study (rational trajectories) is more difficult to deal with than the previous continuous study (real trajectories), and requires estimates related to the *UNI* Property which states that the branches of the dynamical system do not have “too often” the “same” shape. In this discrete model, and for a lattice cost d, the main results are obtained in [2]: There is a CLT for the cost $\widehat{D}_N$, and also a LLT, at least for a lattice cost d. As previously, and in both cases (CLT and LLT), the optimal speed $O(1/\sqrt{\log N})$ is attained. Finally, the CLT and the LLT for a lattice cost are similar in a strong sense, in the real case and the rational case, for $\log N \approx n$.

Local limit theorems for non lattice costs. Even if the main natural cost digit-costs d are integer, and thus lattice, it is important to study Local Limit Theorems for general (non lattice) costs. Such costs give rise to general observables $D_n, \widehat{D}_N$ for the trajectories of the Euclid dynamical system, and a Local Limit Theorem (with a speed of convergence) would be a very precise tool for comparing these observables. There are two processes, a continuous process, and a discrete process, and it is important to know if the rational inputs (of zero measure in the real unit interval) behave as their real counterparts. This is why we wish to obtain a Local Limit Theorem with control of the speed of convergence in the non lattice case, even if this result has no real "practical" consequences in the analysis of the Euclid Algorithm.

Speed of convergence in Local Limit Theorems is not often studied. In the particular case of a memoryless process, where the integers m_i would be independent and uniformly distributed, and after Carlsson [7], Breuillard [4][5] is the first author (to the best of our knowledge) who relates the speed of convergence of the LLT to asymptotic properties of the characteristic function $\phi(\tau)$ when $|\tau| \to \infty$, and exhibits diophantine conditions on the cost d which allow him to control the speed of convergence in the LLT.

In the dynamical system context, the basic tool to obtain all limit theorems mentioned here is the transfer operator $\mathbf{H}_{s,w}$ (with two complex parameters s, w, see Eqn (6)) that allows to implement the characteristic function method. And there is a key difference between real and rational trajectories. For real trajectories, one only deals with one parameter $w = i\tau$ from the characteristic function, while the parameter s is fixed and equal to 1. However, in the transfer operators which appear in the study of rational trajectories, the parameter s ranges in a half-plane containing $s = 1$. The parameter s comes from a Dirichlet series which can be viewed as a generating function for rational trajectories. Then, the study of the LLT for rational trajectories requires estimates on the transfer operator, particularly on the vertical line $\Re s = 1$, when the transfer operator is of the form $\mathbf{H}_{1+it,i\tau}$, with both t and τ varying in $\mathbb{R}$.

When τ belongs to any compact set, or when $|t|$ is large enough, the bounds of [2] based on a first result of Dolgopyat [10] are sufficient. These bounds require a geometric condition on the shapes of the branches, the *UNI* Condition. For large $|\tau|$ and bounded $|t|$, Baladi and Hachemi already adapted in [3] other estimates of Dolgopyat [11] that have been previously generalized by Melbourne [13] to infinite alphabets. These last estimates require an arithmetic condition on the branches, of diophantine type. Finally, in their paper [3], Baladi and Hachemi obtain an intertwined condition which both involves the cost d and the canonical cost c related to branches of the underlying dynamical system: this is a mixing between the condition on cost d which already occur for simple memoryless systems and the condition on cost c which is "added" by the correlations of the dynamical system.

Main results. Even if we are interested by the Euclidean system, we study, as Baladi and Hachemi in [3], a general dynamical system, and we return to the Euclidean particular case at the end of the paper. We then deal with a dynamical system, its canonical cost c, and another cost d defined on the branches. Our purpose is to obtain natural conditions on the pair of costs (c, d) which would imply Local Limit Theorems with speed of convergence. We thus wish to obtain good estimates for the transfer operator $\mathbf{H}_{1+it,i\tau}$ for large $|\tau|$ and bounded $|t|$. We exhibit a set of two conditions, one condition (D) for cost d, and another one (C) for cost c, not intertwined. The first condition (D) is exactly the diophantine condition which already occurs in the memoryless case, or for real trajectories. The second condition (C), also of diophantine type, is a refinement of the *UNI* Condition. It deals with a new notion, the measure of non additivity of cost c, and appears to be original. Our main results are as follows:

(a) We relate the measure of non additivity of cost c to the *UNI* distance. We explain why Condition (C) can be viewed as a "strong diophantine" version of the *UNI* Condition.

(b) We prove that this set $(C), (D)$ of two conditions entails a good behavior for the transfer operator

Sections	Probabilistic Setting	Diophantine Properties of cost d	$\Longrightarrow$	Characteristic Properties of cost d	$\Longrightarrow$	Local Limit Theorem with speed of convergence
1.1	Memoryless case	Def. $1(b)$	Prop. 2	Def. $1(a)$	Prop. 3	
1.2	Real trajectories	Def. 4	Thm 1	Def. 3	Thm 2	
1.3	Rational trajectories of the Euclid system	Def. 4	Thm 3	Def. 5	Thm 4	

Figure 1: Summary of Section 1

$\mathbf{H}_{s,i\tau}$ for large $|\tau|$ and bounded $|t|$.

(c) We prove that the second condition (C) is satisfied by the Euclid Dynamical system: this system already satisfies the *UNI* condition and Baker's theorem shows that the diophantine extra condition on cost c is always fulfilled. Then the condition (C) "disappears" in the present case of interest, and we obtain a LLT (with a speed of convergence) for rational trajectories of the Euclid system, which holds under the unique condition (D) and is exactly of same type as in simpler cases (memoryless cases or real trajectories).

Plan of the paper. In Section 1, we recall the general framework of Local Limit Yheorems with speed of convergence, and replace our precise question (about trajectories of the Euclid Algorithm) inside this general context. We define the main objects of interest and present the statements of our results. Section 2 describes more precisely our strategy of proof, and explains the role of the main conditions (C) and (D) in the scheme of our proof. Section 2 does not contains the complete proof itself, which will be described in the full version of the paper.

1 Main statements

We describe here the general framework of the paper: Local Limit Theorems, with their speed of convergence. Even if we focus to rational trajectories of the Euclid dynamical system, we wish to replace our study in its general context, and begin by easier studies. Thus, we consider three cases, with increasing complexity: first, in Section 1.1, the case of independent variables (which can be related to a memoryless dynamical system), then, in Section 1.2, the (real) trajectories of a general dynamical system, and finally, in Section 1.3, the rational trajectories of the Euclid system. In each case, we are mainly interested by the relation between (i) the speed of convergence in the LLT, (ii) the asymptotic behaviour of the characteristic function, and (iii) the arithmetic properties of the basic cost (of diophantine type). We then show here that our main result on the Euclid algorithm (a LLT with a speed of convergence) is an exact extension of similar theorems in easier cases. The last section 1.4 describes the result which constitutes a main step for our Euclid result. This result holds for a general dynamical system and appears to be of independent interest.

1.1 LLT for a sum of independent variables with the same distribution.

Let (X_i) be a sequence of mutually independent random variables, with values in $\mathbb{N}$. These variables share the same distribution defined by the sequence $p_m := \mathbb{P}[X_i = m]$. We consider a cost $d : \mathbb{N} \to \mathbb{R}^+$, and we assume the following two conditions

$$\sigma_0 := \inf\{\sigma;\ \sum_{i=1}^{\infty} p_m^\sigma < \infty\} < 1, \qquad d(m) = O(|\log p_m|).$$

A distribution (p_m) that satisfies the first hypothesis belongs to the Good Class, and a cost that satisfies the second hypothesis is said to be of moderate growth. Such a cost possesses moments of any order, and we denote by $\mu[d]$ its mean and by $\sigma[d]$ its standard deviation. We wish to study the asymptotic distribution of the total cost

$$D_n := \sum_{i=1}^{n} d(X_i) \qquad (n \to \infty). \tag{3}$$

As soon as $\sigma[d]$ is not zero, there is a Central Limit Theorem (CLT) for D_n with a speed of convergence of order $O(1/\sqrt{n})$,

$$\mathbb{P}\left[\frac{D_n - n\mu[d]}{\sigma[d]\sqrt{n}} \le y\right] - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-t^2/2}dt = O\left(\frac{1}{\sqrt{n}}\right).$$

A Local Limit Theorem (LLT) deals with $Q(x,n) := \mu[d]n + \sigma[d]x\sqrt{n}$, for $x \in \mathbb{R}$, evaluates the probability that the difference $D_n - Q(x,n)$ belongs to some fixed interval $J \subset \mathbb{R}$, and compares it to $(|J|/\sqrt{2\pi n})e^{-x^2/2}$. Under the condition $\sigma[d] \neq 0$, there is a Local Limit Theorem (LLT) which proves

$$\sqrt{n}\,\mathbb{P}[D_n - \mathbf{Q}(x,n) \in J] - |J|\frac{e^{-x^2/2}}{\sigma[d]\sqrt{2\pi}} \to 0, \quad Q(x,n) := \mu[d]n + \sigma[d]x\sqrt{n} \qquad (n \to \infty).$$

However, the speed of convergence is less studied, and there are two main distinct cases, depending on arithmetical properties of the cost d: the lattice case, and the non–lattice case. We recall that a cost d is lattice if there exist $d_0, L \in \mathbb{R}$ with $L > 0$, such that any value $d(m)$ satisfies $d(m) = d_0 + L_m \cdot L$ for some integer L_m.

In the lattice case, the optimal speed of convergence, of order $O(1/\sqrt{n})$ is attained. In the non–lattice case, the speed in the Local Limit Theorem depends on the behaviour of the characteristic function

$$\phi(\tau) := \int_{\mathbb{R}} \exp[i\tau x]\, dP_d(x) = \sum_{m\ge 1} p_m \exp[i\tau d(m)],$$

as it is well explained in the works of Breuillard [4],[5]. We first recall an easy result :

Proposition 1 *The two following conditions are equivalent :*

(i) *The cost d is lattice*

(ii) *There exists $\tau_0 \neq 0$ for which $|\phi_d(\tau_0)| = 1$.*

Moreover, Condition (i) or (ii) entails the following Condition (iii):

(iii) *For any $h, k, \ell \in \mathbb{N}$, letting $d(h,k) := d(h) - d(k)$, the ratio $d(h,k)/d(h,\ell)$ is rational.*

It is then natural to introduce the following two properties, which provide respective reinforcements of the negation of Conditions (ii) or (iii) which occur in Proposition 1. Definition $1(a),(b)$ appear in Breuillard's works [4],[5], with other names. We recall that a number x is diophantine of exponent μ if there exists $C > 0$ such that the inequality $|x - (p/q)| > Cq^{-(2+\mu)}$ is satisfied for any pair $(p,q) \in \mathbb{N}^2$ (See for instance [8])

Definition 1 *There are two main exponents of interest for a cost d:*
(a) *A cost d is of characteristic exponent χ if there exist a constant K and a real $\tau_0 > 0$ for which its characteristic function satisfies*

$$|\phi(\tau)| \le 1 - \frac{K}{|\tau|^\chi} \qquad \text{for } |\tau| \ge \tau_0.$$

(b) *A cost d is of diophantine exponent μ if there exist three integers $(h,k,\ell) \in \mathbb{N}^3$ such that the ratio $d(h,k)/d(h,\ell)$ is diophantine of exponent μ.*

It is important to relate these two notions, because Definition $1(a)$ is easier to deal with for proving a LLT, whereas Definition 1 (b) is independent of the distribution, more natural and easier to check. The following result relates these two notions. This is a particular case of Theorem 7 stated in Section 2.

Proposition 2 *If a cost d is of diophantine exponent μ, it is of characteristic exponent χ for any χ that satisfies $\chi > 2(\mu+1)$.*

In the case when the cost d is of characteristic exponent χ, the following result, proven in Breuillard's paper or adapted in [3] to another context, shows that a Local Limit Theorem holds with a speed of convergence closely related to the characteristic exponent χ. With Proposition 2, this result can be applied to any cost of diophantine exponent μ, when replacing χ by $2(\mu+1)$.

Proposition 3 *Consider a distribution (p_m) of Good Class, and a non lattice cost d of moderate growth, with a mean value $\mathbb{E}[d]$ and a standard deviation $\sigma[d] \neq 0$. If moreover the cost d is of characteristic exponent χ, the following holds for the sum D_n defined in (3): For any ϵ with $\epsilon < 1/\chi$, for any compact interval $J \subset \mathbb{R}$, there exists a constant M_J so that, for every $x \in \mathbb{R}$ and all integers $n \ge 1$*

$$\left| \sqrt{n}\,\mathbb{P}[D_n - \mathbf{Q}(x,n) \in J] - |J| \frac{e^{-x^2/2}}{\sigma[d]\sqrt{2\pi}} \right| \le \frac{M_J}{n^\epsilon}.$$

The proof of Proposition 3 uses as a main tool the characteristic function $\phi_n(\tau)$ of the cost D_n that is equal to $\phi^n(\tau)$, and deals with a compactly supported $\psi \in \mathcal{C}^r(\mathbb{R})$ and the integral

$$I_n := \sqrt{n}\,\mathbb{E}[\psi(D_n(u) - \mathbf{Q}(x,n))] = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{\psi}(\tau) \exp[i\tau \mathbf{Q}(x,n)] \phi_n(\tau) d\tau,$$

where $\widehat{\psi}(\tau)$ is the Fourier transform of ψ. With some $\chi_0 \in]\chi, r-1]$, and $\ell_n := n^{1/\chi_0}$, there are four parts in the integral, respectively defined by the intervals

$$|\tau| \le \nu_0, \qquad |\tau| \in [\nu_0, 2] \qquad |\tau| \in [2, \ell_n], \qquad |\tau| > \ell_n.$$

The first interval gives the main contribution, via a saddle–point approximate. In the second interval, the aperiodicity of the cost is used, whereas the fourth integral deals with the regularity of the function ψ. The hypothesis about the characteristic exponent is crucial for the third integral.

1.2 Truncated trajectories of a dynamical system.

We consider now a dynamical system of the unit interval $\mathcal{I}$, associated to a denumerable partition $\mathcal{I}_m$, a shift T, and a coding mapping M. More precisely,

Definition 2 *A dynamical system of the interval $\mathcal{I} := [0,1]$ is defined by*
(a) A topological partition of $\mathcal{I}$ with disjoint open intervals $\mathcal{I}_m$, $m \in \mathbb{N}$, i.e. $\overline{\mathcal{I}} = \bigcup_{m\in\mathbb{N}} \overline{\mathcal{I}}_m$.
(b) A coding mapping $M : \mathcal{I} \to \mathbb{N}$, almost everywhere defined by the relation $M(x) = m$ for $x \in \mathcal{I}_m$.
(c) A mapping $T : \mathcal{I} \to \mathcal{I}$ whose restriction to each $\mathcal{I}_m$ is a $\mathcal{C}^2$ bijection from I_m to $\mathcal{I}$. The set $\mathcal{H} := \{h_m, m \in \mathbb{N}\}$ gathers all the local inverse branches h_m of T restricted to $\mathcal{I}_m$.

With an initial random variable X with values in the unit interval $\mathcal{I}$, the variables of interest are now

$$X_i := M[T^{i-1}(X)], \qquad D_n := \sum_{i=1}^{n} d(X_i). \tag{4}$$

The variables X_i are integral, and in the case when X is uniform, and the branches are both affine and surjective, we recover the framework of Section 1.1, with $p_m := |h'_m|$. For a general dynamical system, the variables X_i are no longer independent (these are the m_i's of the introduction in the Euclid case), and the cost $D_n(x)$ is just the cost of the trajectory $\mathcal{T}(x)$ truncated at depth n.

For a system of the Good Class (see Def. 6 in Section 2), any inverse branch of T^p is written as an element of $\mathcal{H}^p$: its depth $p(h)$ equals p. The set $\mathcal{H}^\star$ gathers all the inverse branches, of any depth. Any element $h \in \mathcal{H}^\star$ has a unique fixed point denoted by $h^\star$. This defines a canonical cost c, uniquely defined from the dynamical system, together with its normalized version $\widehat{c}$,

$$c(h) := -\log|h'(h^\star)|, \qquad \widehat{c}(h) := \frac{c(h)}{p(h)}. \tag{5}$$

This cost will play the rôle of a reference cost. In the memoryless case, the costs $c(m)$ coincide with the logarithms of probabilities p_m, and the cost c is additive, namely $c(h \circ k) = c(h) + c(k)$. Moreover, for a system of the Good Class, uniform contraction of branches h (see Def. 6) can be expressed by the property

$$\theta_0 := \inf\{\widehat{c}(h);\ \ h \in \mathcal{H}^\star\} > 0,$$

and the real θ_0 is called the logarithmic contraction of the system.

We also consider another cost d defined on $\mathbb{N}$, and, since there is a bijection between $\mathbb{N}$ and the set $\mathcal{H}$ of inverse branches, we may consider d as a cost on $\mathcal{H}$; we further extend it to the subset $\mathcal{H}^\star$ in an additive way, via the relation $d(h \circ k) := d(h) + d(k)$. We restrict ourselves to a cost d of moderate growth, for which $d(h) = O(c(h))$ for $h \in \mathcal{H}$, and we wish to study D_n defined in (4). The main tool is, as previously, its moment generating function Φ_n. In the present framework, it is written with the transfer operator $\mathbf{H}_{s,w}$ (s, w are complex parameters) defined by

$$\mathbf{H}_{s,w}[f](x) = \sum_{h\in\mathcal{H}} \exp[wd(h)]\,|h'(x)|^s\, f(h(x)), \tag{6}$$

under the form

$$\Phi_n(w) := \mathbb{E}[\exp(wD_n)] = \int_{\mathcal{I}} \mathbf{H}^n_{1,w}[f](x)dx,$$

and the characteristic function $\phi_n(\tau)$ is just

$$\phi_n(\tau) := \Phi_n(i\tau) = \int_{\mathcal{I}} \mathbf{H}^n_{1,i\tau}[f](x)dx. \tag{7}$$

If the system is of Good Class, and the cost d of moderate growth, then, for convenient values of pairs (s, w) of parameters, the transfer operator $\mathbf{H}_{s,w}$ acts on the space $\mathcal{C}^1(\mathcal{I})$ endowed with the usual norm $||.||_1$. The operator $\mathbf{H}_{1,w}$ admits spectral dominant properties when w is close to 0 that entail that $\mathbf{H}^n_{1,w}[f](x)$ is a uniform quasi-power when w is close to 0, and x belongs to $\mathcal{I}$. This entails a similar property for the moment generating function $\Phi_n(w)$, and there is a Central Limit Theorem for the costs D_n, with an optimal speed $O(1/\sqrt{n})$. For the LLT, there are, as in Section 1.1, two cases, according as the cost be lattice or non lattice. The optimal speed is attained in the lattice case, and, in the non–lattice case, it is then natural to introduce the following definition which provides a natural extension of Part (a) in Definition 1.

Definition 3 *The cost d is of characteristic exponent χ wrt to the dynamical system (I, T) if there exist $\gamma > 0, \tau_0 > 0$ such that the weighted transfer operator $\mathbf{H}_{s,w}$ relative to $(\mathcal{I}, T, d)$ and defined in Eqn (6) satisfies, when acting on the space $\mathcal{C}^1(\mathcal{I})$, with $n(\gamma, \tau) := \lfloor \gamma \log |\tau| \rfloor + 1$,*

$$||\mathbf{H}^{n(\gamma,\tau)}_{1,i\tau}|| \le 1 - \frac{1}{|\tau|^\chi} \qquad \textit{for any } \tau \textit{ with } |\tau| \ge \tau_0.$$

For a natural extension of Part (b) of Definition 1, we first extend the function $d(h, k)$ involved in Proposition 1 to any pair (h, k) of branches of any depth, via the relation $d(h, k) := p(h)d(k) - p(k)d(h)$. We also deal with the set of branches whose cost $\widehat{c}$ is bounded,

$$\mathcal{H}^\star(\theta) := \left\{ h \in \mathcal{H}^\star,\ \ \widehat{c}(h) := \frac{c(h)}{p(h)} < \theta \right\}. \tag{8}$$

The following definition extends in a natural way Part (b) of Definition 1,

Definition 4 [Condition $(D)(\mu, \theta)$] *Consider a dynamical system (I, T) and an additive cost d. Let $d(h, k) := p(h)d(k) - p(k)d(h)$. The cost d is of diophantine exponent (μ, θ) if there exists a triple (h, k, ℓ) for which the following three conditions hold:*

- (i) *the five branches $h, k, \ell, h \circ k, h \circ \ell$ belong to $\mathcal{H}^\star(\theta)$ defined in Eqn (8),*
- (ii) *$d(h, k)$ and $d(h, \ell)$ are both non zero,*
- (iii) *the ratio $d(h, k)/d(h, \ell)$ is diophantine of exponent μ.*

There are now two main results: The first one relates the notions introduced in the two previous definitions Def. 3 and Def. 4, and may be viewed as a natural extension of Proposition 2. This is a particular case of Theorem 7 stated in Section 2.

Theorem 1 *Consider a dynamical system of the Good Class, of logarithmic contraction θ_0, and an additive cost d of moderate growth defined on $\mathcal{H}^\star$. If d is of diophantine exponent (μ, θ), then it is of characteristic exponent χ with any $\chi > 2(\mu+1)(2+\theta/\theta_0)$.*

The second result is the analog of Proposition 3, and provides a LLT with a speed of convergence. With the previous Theorem 1, the conclusions hold as soon as d is of diophantine exponent (μ, θ), with $\chi > 2(\mu+1)(2+\theta/\theta_0)$. The proof of Theorem 2 starts with the expression of Eqn (7) and follows the same lines as Proposition 3.

Theorem 2 *Consider a dynamical system of the Good Class, and a non lattice cost d of moderate growth, defined on $\mathcal{H}$, extended to $\mathcal{H}^\star$ in an additive way. Denote by $\mu[d]$ and $\sigma[d]$ the mean and the standard deviation of digit-cost wrt to the invariant density φ of the dynamical system, and let*

$$Q(x,n) := \mu[d]n + \sigma[d]x\sqrt{n}.$$

If, moreover, the cost d is of characteristic exponent χ, then, the following holds for the cost D_n defined in (4): for any ϵ, with $\epsilon < 1/\chi$, for any compact interval $J \subset \mathbb{R}$, there exists a constant M_J so that, for every $x \in \mathbb{R}$ and all integers $n \geq 1$

$$\left| \sqrt{n}\,\mathbb{P}[D_n - \mathbf{Q}(x,n) \in J] - |J| \frac{e^{-x^2/2}}{\sigma[d]\sqrt{2\pi}} \right| \leq \frac{M_J}{n^\epsilon}.$$

1.3 Rational trajectories of the Euclid system.

Since the Euclid dynamical system belongs to the Good Class, the previous section applies, and describes the asymptotic behavior of its truncated trajectories $\mathcal{T}_n(x)$ when x is a random real of the unit interval. Here, the random variables X_i are just the quotients that occur in the continued fraction expansion of the real x, that are classically denoted by m_i.

We now focus on our main subject of interest, when the input x is no longer a random real of the unit interval, but a random rational p/q of the interval. In this case, the trajectory $\mathcal{T}(p/q)$ ends "itself" at the depth $P(p,q)$. This depth equals the number of steps of the Euclid algorithm on the pair (p,q). We consider a pair (p,q) of integers and we wish to study the cost of the trajectory $\mathcal{T}(p/q)$, namely

$$\widehat{D}(p,q) := \sum_{i=1}^{P(p,q)} d(m_i(p/q)). \tag{9}$$

For any $N \geq 1$, we consider the restriction $\widehat{D}_N$ of $\widehat{D}$ to the finite subset

$$\Omega_N := \{(p,q); \qquad 1 \leq p \leq q, \quad \gcd(p,q) = 1, \quad q \leq N\},$$

endowed with the uniform probability, and we study the asymptotic behaviour of the cost $\widehat{D}_N$ when $N \to \infty$. In particular, we compare it to the previous cost D_n when $n \to \infty$, and $\log N$ is close to n.

As it is now well-known from works [17], [18], [2], the main object for this discrete study is the transfer operator $\mathbf{H}_{1+it,i\tau}$, defined in (6) or more generally, the operator $\mathbf{H}_{s,i\tau}$, when $\Re s$ is close to 1, and we deal with its quasi-inverse

$$S(t,\tau) := (I - \mathbf{H}_{1+it,i\tau})^{-1}.$$

Distribution results for the cost $\widehat{D}$ depend on nice properties of $S(t,\tau)$, and there are five regions of interest for the pair (t,τ): region R_0 where $t=0$, region R_1 where (t,τ) is close to $(0,0)$, region R_2 where (t,τ) remains bounded, but not small, region R_3 where $|t|$ is large, and finally region R_4 where $|\tau|$ is large, and $|t|$ bounded.

For the first four regions, there already exist convenient estimates for the quasi-inverse $S(s,\tau)$. The estimates in region R_0 already obtained in Section 1.2 are (only) related to properties of cost d; the estimates in region R_1 are only due to the Good Class properties, whereas the role of the *UNI* Condition, later described in Section 2.2, is crucial for estimates in regions R_2 or R_3. These estimates are used for studying distribution of costs $\widehat{D}_N$, and they are sufficient to entail the Central Limit Theorem (for any cost d of moderate growth), or the Local Limit Theorem (for any lattice cost d of moderate growth). After Hensley [12], Baladi and Vallée obtained in [2] a Central Limit Theorem, and a Local Limit Theorem, both with optimal speed(i),

$$\mathbb{P}_N\left[\frac{D_N-\widehat{\mu}[d]\,\log N}{\widehat{\sigma}[d]\,\sqrt{\log N}}\le y\right]-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}e^{-t^2/2}dt=O\left(\frac{1}{\sqrt{\log N}}\right),$$

where $\widehat{\mu}[d],\widehat{\sigma}[d]$ can be expressed with their "continuous" counterparts $\mu[d]$ and $\sigma[d]$, and also as a function of the distribution of the length P of the rational trajectory.

However, in order to obtain a LLT in the case of a non lattice cost d (with the control of the speed of convergence), we need estimates in region R_4, which leads to the following notion, already introduced in the Baladi-Hachemi paper (without its name):

Definition 5 *The cost d is of uniform characteristic exponent χ wrt to the dynamical system (I,T) if there exists $\gamma>0$ for which, for any $a\ge 0$, there is τ_0 such that the weighted transfer operator relative to $(\mathcal{I},T,d)$ satisfies, when acting on the space $\mathcal{C}^1(\mathcal{I})$, with $n(\gamma,\tau):=\lfloor\gamma\log|\tau|\rfloor+1$,*

$$||\mathbf{H}^{n(\gamma,\tau)}_{1+it,i\tau}||\le 1-\frac{1}{|\tau|^{\chi}},\qquad \text{for any } (t,\tau) \text{ with } |t|\le a \text{ and } |\tau|\ge\tau_0.$$

Definition 5 deals with any $a\ge 0$, and is clearly stronger than Definition 3 that only deals with the case $a=0$. There are now two main facts:

$(F1)$ Definition 5 is well-adapted to obtain a LLT for the costs $\widehat{D}_N$, as Baladi and Hachemi already proved it in [3].

$(F2)$ We have already seen in Theorem 1 that properties described in Def. 4 entail those of Def. 3. The following result, Theorem 3, is one of the main steps for our final result. This is an extension of Theorem 1, and it is obtained as a particular consequence of Theorem 6, stated in Section 2. Remark that the bound obtained for the exponent χ in Theorem 3 is the same as the bound obtained for the exponent χ in Theorem 1.

Theorem 3 *Consider the Euclidean dynamical system, and an additive cost d of moderate growth, defined on $\mathcal{H}$ and extended to $\mathcal{H}^\star$ in an additive way. If d is of diophantine exponent (μ,θ), then it is of uniform characteristic exponent χ with any $\chi>2(\mu+1)(2+\theta/\theta_0)$, where $\theta_0=\log(1+\sqrt{5})/2$ is the logarithmic contraction of the Euclidean Dynamical system.*

(i) Hensley studied only the length of the rational trajectory, and did not attain the optimal speed of convergence

With facts $(F1)$ and $(F2)$, it is now easy to obtain the following Theorem 4. This is one of the main results of the paper, and it proves a LLT for costs $\widehat{D}_N$, exactly of the same type as for costs D_n, obtained in Theorem 2 of the previous section. When one replaces in Theorem 4 the sentence "If moreover the cost d is diophantine with parameters (μ, θ), let $\chi := 2(\mu + 1)(2 + \theta/\theta_0)$" by the sentence "If moreover the cost d is of uniform characteristic exponent χ", we obtain the result proven by Baladi and Hachemi in [3]. Here, we "just" provide a natural sufficient condition for applying it.

Theorem 4 *Consider the Euclidean dynamical system, for which $\theta_0 := \log(1 + \sqrt{5})/2$. Consider a cost d of moderate growth defined on $\mathcal{H}$ and extended to $\mathcal{H}^\star$ in an additive way. With $\widehat{\mu}[d], \widehat{\sigma}[d]$ previously defined, let*

$$\widehat{Q}(x, N) := \widehat{\mu}[d] \log N + \widehat{\sigma}[d] x \sqrt{\log N}.$$

If moreover the cost d is diophantine with parameters (μ, θ), let $\chi := 2(\mu + 1)(2 + \theta/\theta_0)$. Then, the following holds for the cost $\widehat{D}_N$ of rational trajectories defined in (9): For any ϵ with $\epsilon < 1/\chi$, for any compact interval $J \subset \mathbb{R}$, there exists a constant M_J so that, for every $x \in \mathbb{R}$ and all integers $n \geq 1$,

$$\left| \sqrt{\log N}\, \mathbb{P}_N[\widehat{D}_N - \widehat{Q}(x, N) \in J] - |J| \frac{e^{-x^2/2}}{\widehat{\sigma}[d]\sqrt{2\pi}} \right| \leq \frac{M_J}{\log^\epsilon N}.$$

1.4 A more general result, of independent interest.

We focus here to the Euclidean dynamical system. But, in our proof, we are led to answer a more general question that may be asked for any dynamical system: Is it possible to find natural conditions on the triple (I, T, d) which entail good estimates of the transfer operator in Region R_4? Region R_4 is an intermediary region between region R_3, where $|t|$ is large and Region R_0 where $t = 0$. In Region R_3, the *UNI* Condition on the system (I, T) is needed, whereas only conditions on cost d are needed in Region R_0, as we already proved in Theorem 1.

Baladi and Hachemi already asked this general question, and they exhibit an intertwined condition between the canonical cost c related to branches of system (I, T) and the digit-cost d that is sufficient to entail properties of Def. 4 (See their condition described later in Prop. 6of Section 2). Here, we exhibit in Section 2.6 a double condition, not intertwined, which involves, on the one hand, the cost d –this is the condition $(D)(\theta, \mu)$– and on the other hand, the canonical cost c of the dynamical system (I, T) –this is the condition $(C)(\theta)$ that deals with a new concept, namely the measure of non additivity of cost c, defined in (11)–. Then, the main scheme of our proof is as follows:

(i) we relate Condition $(C)(\theta)$ to the *UNI* Condition in Proposition 4.

(ii) we prove in Theorem 7 that these two conditions $(C)(\theta)$ and $(D)(\theta, \mu)$ are sufficient to entail good estimates for the quasi-inverse in region R_4.

(iii) we prove that Conditions $(C)(\theta)$ are always true for any $\theta > \theta_0$ in the Euclid dynamical system.

2 Strategies for the proof.

We introduce in the first three sections (Sections 2.1, 2.2, 2.3), the precise definitions for the objects for interest: Good Class, *UNI* distance, *UNI* Condition, measure Γ of non additivity. We then describe a basic result which relate various notions, and explain why it is interesting to consider "reinforcements of their negations". Following Dolgopyat [11], Melbourne [13], Baladi and Hachemi [3], and using the teminology of Roux and Vallée [15], [16], we are led to introduce in Section 2.5 subsets of type $\mathcal{F}$ (whose boundless is related to the characteristic properties of cost d) and $\mathcal{A}$ (related to the existence of approximate eigenfunctions). We first state in Section 2.5 the result (Theorem 5) which proves the inclusion between subsets $\mathcal{F}^c$ and subsets $\mathcal{A}$. Then, we introduce in Section 2.6 our original diophantine conditions (C) and (D), and prove that they ensure $\mathcal{A}$ to be bounded. Finally, this Section describes the complete chain between diophantine properties and characteristic properties of cost d, needed to obtain LLT with the control of speed, as it is summarized in Figure 2.

(C) [and (D)] Conditions	$\Longrightarrow$ Thm 6	Subset $\mathcal{A}$ is bounded	$\Longrightarrow$ $\mathcal{F}^c \subset \mathcal{A}$ Thm 5	Subset $\mathcal{F}^c$ is bounded	$\Longrightarrow$ Def of $\mathcal{F}$	Characteristic properties of cost d

Figure 2: Summary of Section 2

2.1 Good Class

We first define the Good Class, for which the shift is expansive, and gives rise to a chaotic behaviour for the trajectories. The canonical cost c, and its normalized version $\widehat{c}$, already defined in Eqn (5), play a central role:

Definition 6 [Good Class]. *A dynamical system of the interval $(\mathcal{I}, T)$ belongs to the Good Class if it is complete, with a set $\mathcal{H}$ of inverse branches which satisfies the following:*

$(G1)$ *The set $\mathcal{H}$ is uniformly contracting, i.e., there exists a constant $\theta_0 > 0$, for which*

$$\inf\{\widehat{c}(h), h \in \mathcal{H}^\star\} = \theta_0$$

The real θ_0 is called the logarithmic contraction, and $\rho := e^{\theta_0}$ is the contraction ratio.

$(G2)$ *There is a constant $A > 0$, so that every inverse branch $h \in \mathcal{H}$ satisfies $|h''| \le A|h'|$.*

$(G3)$ *There exists $\sigma_0 < 1$ for which the series $\sum_{h\in\mathcal{H}} e^{-c(h)s}$ converges on $\Re s > \sigma_0$.*

If the system is of Good Class, and the cost d of moderate growth (we recall that it means $d(h) = O(c(h))$ for $h \in \mathcal{H}$), then, for convenient values of pairs (s, w) of parameters, the transfer operator $\mathbf{H}_{s,w}$ acts on the space $\mathcal{C}^1(\mathcal{I})$ endowed with the usual norm $||.||_1$. The operator $\mathbf{H}_{1,w}$ admits spectral dominant properties, with a spectral gap, when w is close to 0, which entails that $\mathbf{H}^n_{1,w}[f](x)$ is a uniform quasi-power when w is close to 0, and x belongs to $\mathcal{I}$.

2.2 UNI Condition.

One first defines a probability Π_n on each set $\mathcal{H}^n$, in a natural way, and lets $\Pi_n[h] := |h(\mathcal{I})|$ where $|\mathcal{J}|$ denotes the length of the interval $\mathcal{J}$. Furthermore, $\Delta(h, k)$ denotes the "distance" between two inverse branches h and k of same depth, defined as

$$\Delta(h,k) = \inf_{x\in\mathcal{I}} |\Psi'_{h,k}(x)| \qquad \text{with} \quad \Psi_{h,k}(x) = \log\left|\frac{h'(x)}{k'(x)}\right|. \tag{10}$$

The distance $\Delta(h,k)$ is a measure of the difference between the "shape" of the two branches h, k. The *UNI* Condition, stated as follows [10], is a geometric condition which expresses that the probability that two inverse branches have almost the same "shape" is very small:

Definition 7 [Condition *UNI*]. *A dynamical system $(\mathcal{I}, T)$ of the Good Class, with a contraction ratio ρ satisfies the* UNI *condition if its set $\mathcal{H}$ of inverse branches satisfies*

$(U1)$ *For any $\widehat{\rho}$ with $\rho < \widehat{\rho} < 1$, for any integer n, for any $h \in \mathcal{H}^n$, one has $\Pi_n[k;\ \Delta(h,k) \le \widehat{\rho}^n] \ll \widehat{\rho}^n$.*

$(U2)$ *Each $h \in \mathcal{H}$ is of class $\mathcal{C}^3$ and for any n, there exists B_n for which $|h'''| \le B_n |h'|$ for any $h \in \mathcal{H}^n$.*

For a dynamical system with affine branches, the "distance" Δ is always zero, and the probabilities of Assertion $(U1)$ are all equal to 1. Such a source never satisfies the Condition *UNI*. Conversely, a dynamical system of the *Good-UNI* Class cannot be conjugated to a system with affine branches, as it is proven by Baladi and Vallée [2].

2.3 Measure Γ of non additivity.

For a system with affine branches, the canonical cost c is additive: for any $h, k \in \mathcal{H}^\star$ the equality $c(h \circ k) = c(h) + c(k)$ holds. It is now natural to introduce the measure of non additivity of cost c, denoted by Γ,

$$\Gamma(h,k) := c(h) + c(k) - c(h \circ k). \tag{11}$$

This measure will play a central role in the sequel of the paper. The following Proposition 4 relates the distance $\Delta(h,k)$ and the measure $\Gamma(h,k)$. It will be proven in the full paper.

Proposition 4 *Consider a dynamical system of the Good Class, with contraction ratio ρ, the UNI distance Δ defined in Eqn (7), and the Γ measure of non additivity defined in Eqn (11). For two branches h, k of same depth, consider the two "distances"*

$$\delta_1(h,k) := \sup\{|h(x) - k(y)|; (x,y) \in \mathcal{I} \times \mathcal{I}\}, \qquad \delta_2(h,k) := \inf\{|h(x) - k(y)|; (x,y) \in \mathcal{I} \times \mathcal{I}\}$$

Then, the following holds:

(a) *For two branches h, k of same depth, the UNI distance Δ and the measure Γ of non additivity are related, (L is the distortion constant),*

$$|\Delta(h,k)\, \delta_2(h,k) - L\,(|h(\mathcal{I})| + |k(\mathcal{I})|)\, \delta_1(h,k)| \le |\Gamma(h,k)| \le L\,(|h(\mathcal{I})| + |k(\mathcal{I})|)\, \delta_1(h,k)$$

(b) *If the dynamical system satisfies the* `UNI` *condition, then,*

$(b1)$ *Define the probability Π_n on $\mathcal{H}^n$ by $\Pi_n[h] := |h(\mathcal{I})|$. For any $\widehat{\rho}$ with $\rho < \widehat{\rho} < 1$, for any n and any $h \in \mathcal{H}^n$,*

$$\Pi_n[k; \quad |\Gamma(h,k)| \le \widehat{\rho}^n] \ll \widehat{\rho}^{n/2}.$$

$(b2)$ *For any n large enough, and for any $h \in \mathcal{H}^n$, there exists a pair $(k, \ell) \in \mathcal{H}^n \times \mathcal{H}^n$ for which $\Gamma(h,k)$ and $\Gamma(h,\ell)$ are both non zero.*

2.4 A basic proposition

The following result, next described in Proposition 5, is the analog of Proposition 1. The first assertion [equivalence between (i) and (ii)] is classical and its proof is based on the converse of the triangular inequality. It shows that it is sufficient to deal with "component" operators

$$\mathbf{H}_{it,i\tau,h}[f](x) := \exp[i\tau d(h)]\,|h'(x)|^{it}\, f\circ h(x), \qquad \text{for } h\in\mathcal{H}^\star \text{ and } x\in\mathcal{I}. \tag{12}$$

The second assertion deals with Condition (d), already introduced in Proposition 1, but, also, with Condition (c) that involves the measure of non–additivity Γ introduced in (11).

Proposition 5 *Consider the weighted transfer operator $\mathbf{H}_{s,w}$ relative to a dynamical system of the Good Class, related to a canonical cost c, together with an additive cost d of moderate growth. Let*

$$d(h,k) := p(k)d(h) - p(h)d(k), \qquad \Gamma(h,k) := c(h) + c(k) - c(h\circ k).$$

Let η be a complex number with $|\eta| = 1$. For any $t,\tau\in\mathbb{R}$, with $(t,\tau)\neq(0,0)$, the following assertions are equivalent:
(i) The operator $I-\eta\mathbf{H}_{1+it,i\tau}$ is not invertible
(ii) There exists a function $f\in\mathcal{C}^1(\mathcal{I})$ with $|f|=1$ which satisfies the equality

$$\mathbf{H}_{it,i\tau,h}[f] = \eta^{-p(h)} f, \quad \textit{for any } h\in\mathcal{H}^\star. \tag{13}$$

Moreover, Condition (i) or (ii) entails the following Conditions (c) (for $t\neq 0$) and (d) (for $t=0$):
(c) For any (h,k,ℓ) for which $\Gamma(h,k)$ and $\Gamma(h,\ell)$ are not zero, the ratio $\Gamma(h,k)/\Gamma(h,\ell)$ is rational.
(d) For any (h,k,ℓ) for which $d(h,k)$ and $d(h,\ell)$ are not zero, the ratio $d(h,k)/d(h,\ell)$ is rational.

Proof. We fix (t,τ) with $t\neq 0$. Assume that Assertion (ii) holds. The equality (13) taken at the fixed point $x=h^\star$ entails the relation, for any $h\in\mathcal{H}^\star$,

$$\exp[itc(h) - i\tau d(h)] = \exp[ip(h)\theta] \qquad \text{and then} \quad tc(h) - \tau d(h) = p(h)\theta + 2\pi L_h$$

for some integer $L_h = L_h(t,\tau)$. Writing this relation for h,k and $h\circ k$, and using the additivity of cost d, we obtain

$$t\Gamma(h,k) := t[c(h)+c(k)-c(h\circ k)] = 2\pi[L_h + L_k - L_{h\circ k}],$$

and thus Condition (c). We fix (t,τ) with $t=0$ and $\tau\neq 0$. Assume that Assertion (ii) holds with a function f. Taken at the fixed point $x=h^\star$, this entails the relation, for any $h\in\mathcal{H}^\star$,

$$\exp[i\tau d(h)] = \exp[-ip(h)\theta] \qquad \text{and then} \quad \tau d(h) + p(h)\theta = 2\pi L_h, \quad d(h) + p(h)\frac{\theta}{\tau} = 2\frac{\pi}{\tau}L_h,$$

for some integer L_h. This proves that d is lattice, and the relations

$$\tau d(h,k) = \tau[p(k)d(h) - p(h)d(k)] = 2\pi[L_h p(k) - L_k p(h)],$$

$$\tau d(h,\ell) = \tau[p(\ell)d(h) - p(h)d(\ell)] = 2\pi[L_h p(\ell) - L_\ell p(h)],$$

entail Condition (d).

2.5 Subsets $\mathcal{F}$ and $\mathcal{A}$

We wish to deal with reinforcements of the negations of various conditions which intervene in Proposition 5. We begin with conditions (i) and (ii) and we define two subsets $\mathcal{A}$ and $\mathcal{F}$ of $\mathbb{R}^2$: Definition 5 of Section 1.3 (diophantine exponent) is related to a reinforcement of the negation of Condition (i), that leads to subsets of type $\mathcal{F}$. It is then natural to deal with a reinforcement of the negation of Condition (ii) that introduces sets of type $\mathcal{A}$. We then may expect that subsets $\mathcal{A}$ and $\mathcal{F}$ are related, which will be proven to hold in Proposition 5.

We introduce later (in Section 2.6) reinforcements of the negation of Conditions (c) and (d), and observe their relation with reinforcements introduced in this section 2.5.

Condition (i) : Subsets $\mathcal{F}$. Letting $n(\gamma,\tau) := \lfloor \gamma \log|\tau| \rfloor + 1$, we associate to any $a \geq 0$, $\chi > 0$, $\gamma \geq 0$, the set

$$\mathcal{F}_a(\chi,\gamma) := \left\{ (t,\tau) \in \mathbb{R}^2; \quad |t| \leq a,\ |\tau| \geq 2,\ ||\mathbf{H}^{n(\gamma,\tau)}_{1+it,i\tau}||_1 \leq 1 - \frac{K}{|\tau|^{\chi}} \right\}.$$

A possible reinforcement of the negation of Condition (i) is :
There exists a pair (χ,γ) with $\chi > 0, \gamma \geq 0$ such that, for any $a \geq 0$, the subset $\mathcal{F}^c_a(\chi,\gamma)$ is bounded.

This exactly means that the cost d is of uniform characteristic exponent χ wrt to the dynamical system (I,T) (See Def. 5).

Condition (ii): Subsets $\mathcal{A}$. We first consider the subset of $\mathcal{H}^\star$,

$$\mathcal{H}(\tau,\beta,\delta) := \{h; \quad p(h) = n(\beta,\tau), \quad c(h) \leq n(\delta,\tau)\},$$

which gives rise to fundamental intervals $h(\mathcal{I})$ not too "small" with respect to τ.

The set $\mathcal{A}_a(\alpha,\beta,\delta)$ gathers the pairs (t,τ) with $|t| \leq a, |\tau| \geq 2$ for which there exists a pair $(w_{t,\tau},\eta_{t,\tau})$, with $w_{t,\tau} \in \mathcal{W}$ and $|\eta_{t,\tau}| = 1$, which satisfies,

$$||\mathbf{H}_{it,i\tau,h}[w_{t,\tau}] - \eta_{t,\tau} w_{t,\tau}||_\infty \leq \frac{1}{|\tau|^\alpha}, \qquad \text{for any } h \in \mathcal{H}(\tau,\beta,\delta). \tag{14}$$

The pair $(w_{t,\tau},\eta_{t,\tau})$ plays the role of an approximate eigenfunction, and the degree of approximation is described by the triple (α,β,δ).

A possible reinforcement of the negation of Condition (ii) is :
There exists a triple (α,β,δ), such that, for any $a \geq 0$, the subset $\mathcal{A}_a(\alpha,\beta,\delta)$ is bounded.

The following result relates the two types of subsets $\mathcal{F}$ and $\mathcal{A}$. This is a result closely related to Dolgopyat's paper [11] that includes the extensions of Melbourne [13] to an infinite alphabet. A slightly different version is proven in the Baladi-Hachemi paper ([3], Lemma 2.2). However, we do not use here exactly the same definition of subsets $\mathcal{A}$, $\mathcal{F}$, and we prefer to deal with the framework introduced by Roux and Vallée in [15], [16]. The proof of this result will be found in the full paper; it is very similar to proofs of Roux' thesis [15].

Theorem 5 [Subsets $\mathcal{A}$ and $\mathcal{F}$] *Consider any dynamical system of the Good Class with logarithmic contraction θ_0. Then, for any real $a \geq 0$, the following inclusions holds between sets $\mathcal{A}$ and $\mathcal{F}$:*

[Case of memoryless sources.] *For any pair (α, δ) with $\delta > 0$, there exists $\tau_1 > 0$ for which the following inclusion holds,*

$$\mathcal{F}_a^c(2\alpha + \delta, 0) \cap [-a, +a] \times \{\tau; |\tau| \geq \tau_1\} \subset \mathcal{A}_a(\alpha, 0, \delta) \cap [-a, +a] \times \{\tau; |\tau| \geq \tau_1\}.$$

[General case.] *For any triple (α, β, δ) with $\beta\theta_0 \geq \alpha$, there exists $\tau_1 > 0$ for which the following inclusion holds,*

$$\mathcal{F}_a^c(4\alpha + 2\delta, 4\alpha + 2\delta + 3\beta) \cap [-a, +a] \times \{\tau; |\tau| \geq \tau_1\} \subset \mathcal{A}_a(\alpha, \beta, \delta) \cap [-a, +a] \times \{\tau; |\tau| \geq \tau_1\}.$$

It is now sufficient to describe conditions which ensure subset $\mathcal{A}$ to be bounded: With Theorem 5, these conditions will entail subset $\mathcal{F}^c$ to be bounded, and thus the cost d will be of uniform characteristic exponent χ wrt to the dynamical system (I, T).

2.6 Sufficient conditions which ensure $\mathcal{A}$ to be bounded.

The paper of Baladi and Hachemi exhibits a (diophantine) condition which deals with the pair (c, d) of costs and entails that the set $\mathcal{A}$ is bounded.

Proposition 6 [Baladi-Hachemi] *For a set $\{h_i; 1 \leq i \leq 4\}$ of four branches of $\mathcal{H}^\star$, let us adopt the following notations :* $c(1, i) := p(h_1)c(h_i) - p(h_i)c(h_1)$, $d(1, i) := p(h_1)d(h_i) - p(h_i)d(h_1)$, $e(i, j) := c(1, i)d(1, j) - c(1, j)d(1, i)$.

If there exist four branches of $\mathcal{H}^\star(\theta)$ for which the three ratios $\dfrac{d(1,2)}{d(1,3)}, \dfrac{e(2,4)}{e(2,3)}, \dfrac{e(3,4)}{e(2,3)}$ *are diophantine with exponent μ, then, for any $a \geq 0$, for any triple (α, β, δ) with $\alpha > \mu + 1$, $\delta/\beta \geq \theta$, the set $\mathcal{A}_a(\alpha, \beta, \delta)$ is bounded.*

We now exhibit new sufficient conditions which ensure $\mathcal{A}$ to be bounded, based on reinforcements of negation of Conditions (c) and Condition (d) of Prop. 5. Observe that these conditions involve cost c and cost d in a similar way.

[Condition $(C)(\theta)$] *There exists a triple (h, k, ℓ) for which*

- (i) *the five branches $h, k, \ell, h \circ k, h \circ \ell$ belong to $\mathcal{H}^\star(\theta)$ defined in (8).*
- (ii) *$\Gamma(h, k)$ and $\Gamma(h, \ell)$ are non zero,*
- (iii) *the ratio $\Gamma(h, k)/\Gamma(h, \ell)$ is diophantine.*

[Condition $(D)(\mu, \theta)$]*: There exists a triple (h, k, ℓ) for which*

- (i) *the five branches $h, k, \ell, h \circ k, h \circ \ell$ belong to $\mathcal{H}^\star(\theta)$,*
- (ii) *$d(h, k)$ and $d(h, \ell)$ are non zero,*
- (iii) *the ratio $d(h, k)/d(h, \ell)$ is diophantine of exponent μ.*

The main result of the paper deals with these two reinforcements and shows that they together entail that subsets $\mathcal{A}$ are bounded. It will be proven in the full paper, but we describe its main steps in an informal way in the following Section.

Theorem 6 *For any dynamical system of the Good Class, and any cost d of moderate growth, the following conditions entail the boundness of the set $\mathcal{A}_a(\alpha,\beta,\delta)$:*
(i) *If Condition $(D)(\mu,\theta)$ holds, then*

(ia) [case $a=\beta=0$]: *the set $\mathcal{A}_0(\alpha,0,\delta)$ is bounded, as soon as $\alpha>\mu+1$.*

(ib) [case $a=0,\beta>0$]: *the set $\mathcal{A}_0(\alpha,\beta,\delta)$ is bounded, as soon as $\alpha>\mu+1$ and $\delta/\beta>\theta$.*

(ii) *If both Conditions $(C)(\theta)$ and $(D)(\mu,\theta)$ hold, then,*

[case $a\geq 0,\beta\geq 0$]: *the set $\mathcal{A}_a(\alpha,\beta,\delta)$ is bounded, as soon as $\alpha>\mu+1$ and $\delta/\beta>\theta$.*

Together with Theorem 5, Theorem 6 allows us to obtain the main results which are described in Sections 1.1 and 1.2 . We gather them in the following Theorem 7, and we state them more precisely. Proposition 2 is obtained with Theorem 6 (ia) and Theorem 5 whereas Theorem 1 is obtained with Theorem 6 (ib) and Theorem 5.

Theorem 7 *Consider a dynamical system of the Good Class, with logarithmic contraction θ_0, and a cost d of moderate growth.*

(a) [Proposition 2]. *In the memoryless case, if Condition $(D)(\mu,\theta)$ holds, then, the cost d is of characteristic exponent χ wrt to the dynamical system (I,T), for any χ with $\chi>2(\mu+1)$.*

(b) [Theorem 1]. *If Condition $(D)(\mu,\theta)$ holds, then the cost d is of characteristic exponent χ wrt to the dynamical system (I,T), for any χ with $\chi>2(\mu+1)(2+\theta/\theta_0)$.*

(c) *If both Conditions $(C)(\theta)$ and $(D)(\mu,\theta)$ hold, then the cost d is of uniform characteristic exponent χ wrt to the dynamical system (I,T), for any χ with $\chi>2(\mu+1)(2+\theta/\theta_0)$.*

2.7 Main steps for the proof of Theorem 6

We prove it by contradiction. We consider $a\geq 0$ and a triple (α,β,δ) for which $\mathcal{A}_a(\alpha,\beta,\delta)$ is not bounded: there exists a sequence $\mathcal{T}\subset\mathcal{A}_a(\alpha,\beta,\delta)$ of points (t,τ) for which $|\tau|\to\infty$. This entails the existence of a triple (h,k,ℓ) for which, for any $(t,\tau)\in\mathcal{T}$, there exist real $\theta_{t,\tau}$ and five integers $L_h,L_k,L_\ell,L_{h\circ k},L_{h\circ\ell}$ for which

$$t\,n(\tau)\,[c(h)+c(k)]+\tau\,n(\tau)\,[d(h)+d(k)]=[p(h)+p(k)]\theta_{t,\tau}+2\pi[L_h(t,\tau)+L_k(t,\tau)]+O(\tau^{-\alpha}),$$
$$t\,n(\tau)\,[c(h)+c(\ell)]+\tau\,n(\beta,\tau)\,[d(h)+d(\ell)]=[p(h)+p(\ell)]\theta_{t,\tau}+2\pi[L_h(t,\tau)+L_\ell(t,\tau)]+O(\tau^{-\alpha}).$$
$$t\,n(\tau)\,[c(h\circ k)]+\tau\,n(\tau)\,[d(h\circ k)]=[p(h\circ k)]\theta_{t,\tau}+2\pi L_{h\circ k}(t,\tau)+O(\tau^{-\alpha}),$$
$$t\,n(\tau)\,[c(h\circ\ell)]+\tau\,n(\tau)\,[d(h\circ\ell)]=[p(h\circ\ell)]\theta_{t,\tau}+2\pi L_{h\circ\ell}(t,\tau)+O(\tau^{-\alpha}).$$

Using in a strong way the additivity of cost d and depth p, we obtain two final relations which involve $\Gamma(h,k)$ and $\Gamma(h,\ell)$,

$$t\,n(\tau)\,\Gamma(h,\ell)=2\pi L_{h,\ell}(t,\tau)+O(\tau^{-\alpha}),\qquad t\,n(\tau)\,\Gamma(h,k)=2\pi L_{h,k}(t,\tau)+O(\tau^{-\alpha}).$$

We consider now three possible cases for the sequence $\mathcal{T}$, and this leads in each case to a contradiction:

– the first case $|t|n(\tau)\to\infty$ entails that the ratio $\Gamma(h,k)/\Gamma(h,k)$ is diophantine, and contradicts Condition $(C)(\theta)$.

– the second case $A\leq|t|n(\tau)\leq B$ for some $A,B>0$ entails that the ratio $\Gamma(h,k)/\Gamma(h,k)$ is rational and contradicts Condition $(C)(\theta)$.

– the third case $|t|n(\tau) \to 0$ entails that the ratio $d(h,k)/d(h,\ell)$ is diophantine with exponent μ with $\mu + 2 \geq \alpha + 1$ and contradicts Condition $(D)(\mu, \theta)$.

2.8 Rational trajectories of the Euclid dynamical system

The following result describes a general setting which will be further applied to the Euclid Dynamical system. This will imply that Condition $(C)(\theta)$ holds for for any $\theta > \theta_0$. Together with Theorem 7, it entails Theorem 3 of Section 1.3.

Proposition 7 *Consider a system of the Good Class. If the following two conditions are satisfied:*

(a) *The set $\mathcal{H}$ is formed with linear fractional transformations with integer coefficients.*

(b) *There exists a triple (h, k, ℓ) for which*

$(b1)$ *the five branches $h, k, \ell, h \circ k, h \circ \ell$ belong to $\mathcal{H}^\star(\theta)$,*

$(b2)$ *the ratio $\Gamma(h,k)/\Gamma(h,\ell)$ is irrational,*

then Condition $(C)(\theta)$ holds.
In the case of the Euclid system, Condition (a) holds and Condition (b) holds for any $\theta > \theta_0$.

The proof of Proposition 7 is mainly based on three easy lemmas, and will be found in the full paper.

Lemma 1 *Consider a system of the Good Class for which the set $\mathcal{H}$ is formed with linear fractional transformations with integer coefficients. Then, for any pair (h,k) of branches of $\mathcal{H}^\star$, the number $\Gamma(h,k)$ is the logarithm of an algebraic number.*

The second lemma is an easy consequence of Baker's theorem:

Lemma 2 *Let a and b be two algebraic numbers. If the ratio $\log b/\log a$ is irrational, it has a finite irrationality exponent.*

The third lemma returns to the Euclid system and exhibits for each θ_1 a possible triple (h, k, ℓ).

Lemma 3 *In the case of the Euclid system, Condition (b) of Proposition 7 holds for any $\theta_1 > \theta_0$.*

Conclusion

For a digit cost of moderate growth, we have obtained a Local Limit Theorem, with control of the speed, for rational trajectories of the Euclid System, under a diophantine hypothesis on cost d that is similar to the hypothesis which already occurs in simpler frameworks (memoryless cases, real trajectories). As Baladi and Hachemi mentioned it in [3], this would be possible to replace the hypothesis "The cost d is of moderate growth", by the weaker hypothesis "The cost d has strong moments up to order 3."
Other discrete trajectories are of great interest for a general dynamical system, for instance periodic trajectories. This class of trajectories, when weighted by a digit cost d, can be studied with the help of the transfer operator $\mathbf{H}_{s,w}$, in the same vein as in [14], where the plain transfer operator is used for (unweighted) periodic trajectories. It would be then possible to obtain a Local Limit Theorem for weighted periodic trajectories (with speed of convergence), of the same type as the present results for rational trajectories. This is a work in progress with Eda Cesaratto.

Acknowledgements

I wish to thank Mathieu Roux for many stimulating discussions during his PhD thesis. The paper is dedicated to the memory of Philippe Flajolet: *This is for you, Philippe.*

References

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

[2] V. Baladi and B. Vallée. Euclidean algorithms are Gaussian. *J. Number Theory* 110 (2005) 331–386.

[3] V. Baladi and A. Hachemi. A local limit theorem with speed of convergence for Euclidean algorithms and diophantine costs, *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* 2008, Vol. 44, No. 4, 749–770

[4] E. Breuillard. Distributions diophantiennes et théorème limite local sur R^d. *Probab. Theory Related Fields* 132 (2005) 39–73.

[5] E. Breuillard. Local limit theorems and equidistribution of random walks on the Heisenberg group. *Geom. Funct. Anal.* 15 (2005) 35–82.

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

[7] H. Carlsson. Remainder term estimates of the renewal function. *Ann. Probab.* 11 (1983) 143–157.

[8] J. W. S. Cassels. An Introduction to Diophantine Approximation. *Cambridge Univ. Press*, New York, 1957.

[9] P. Collet. 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.

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

[11] D. Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergodic Theory Dynam. Systems* 18 (1998) 1097–1114.

[12] D. Hensley. The number of steps in the Euclidean algorithm. *J. Number Theory* 49 (1994) 142–182.

[13] I. Melbourne. Rapid decay of correlations for nonuniformly hyperbolic flows. *Trans. Amer. Math. Soc.* 359 (2007) 2421– 2441.

[14] M. Pollicott. On the rate of mixing of Axiom A flows. *Invent. Math.* 81 (1985) 413–426.

[15] M. Roux. Théorie de l'Information, séries de Dirichlet, et analyse d'algorithmes. *PhD Thesis, University of Caen*, 2011.

[16] M. Roux and B. Vallée. Information theory : Sources, Dirichlet series, and realistic analysis of data structures, *Proceedings of the Conference Words11*, Vol. 63 of Electronic Proceedings of Theoretical Computer Science, pp 199-214 (2011)

[17] B. Vallée. Euclidean dynamics. *Discrete Contin. Dyn. Syst.* 15 (2006) 281–352.

[18] B. Vallée. Digits and continuants in Euclidean algorithms. Ergodic versus Tauberian theorems. . *J. Théor. Nombres Bordeaux* 12 (2000) 531–570.