Gaussian behavior of quadratic irrationals

by

EDA CESARATTO (Buenos Aires) and BRIGITTE VALLÉE (Caen)

1. Introduction

1.1. General framework. This paper studies particular quadratic irrationals (rqi for short) defined by purely periodic continued fraction expansions. When a cost is defined on continued fraction expansions, in an “additive” way, each rqi number inherits the cost defined on its period, and the paper focus on such costs, with a “moderate growth”. Moreover, we consider the usual notion of size that is defined in number theory contexts for an rqi number, and closely related to the fundamental unit of the associated quadratic field, and to the length of the corresponding closed geodesic in the modular surface.

Our general framework is described as follows: we consider the set $\mathcal{P}$ of rqi numbers, the subset $\mathcal{P}_N$ of rqi numbers of size at most $N$, and the restriction $C_N$ to $\mathcal{P}_N$ of an additive cost $C$ defined on $\mathcal{P}$. We perform a probabilistic analysis of $C_N$ when the set $\mathcal{P}_N$ is endowed with the uniform probability. We prove in Theorem 1.1 that $C_N$ asymptotically follows a Gaussian law (when $N \to \infty$). Moreover, Theorem 1.1 exhibits precise asymptotic expansions for its expectation and its variance. Both are of order $\Theta(\log N)$. Theorem 1.2 provides a computable description of the constants that appear in the expansions. The main term of the expectation is already well-studied, but, to the best of our knowledge, the estimates on the variance are new, and this is the first study which exhibits a limit Gaussian law in this context.

All these results are first proven within the unrestricted framework, with no “constraints” on the continued fraction expansion of the rqi numbers. But we also present in Section 5 similar results that hold for rqi numbers with
bounded digits. These results combine the general methods of this paper with previous results that are specific to the constrained case, described in [11, 27, 4]. In the same section, we also study an important (non-additive) cost, namely the Lévy constant, and obtain asymptotic expansions for its expectation and its variance. Finally, we discuss the occurrence of local limit laws, according to the type of costs (lattice or non-lattice).

The central object in our analysis is a bivariate Dirichlet generating function $P(s, w)$ where the complex parameter $s$ “marks” the size and the complex parameter $w$ “marks” the cost. Our general methodology is described by means of three main steps:

**Step 1.** Relate the series $P(s, w)$ to the weighted transfer operator $H_{t, w}$, associated with the underlying dynamical system defined by the Gauss map (with $t = s/2$).

**Step 2.** Transfer the analytic properties of the operator $H_{t, w}$ (acting on convenient function spaces) to obtain analytic properties of $P(s, w)$ about:

- dominant singularity (when $s$ is close to the real axis), and
- polynomial growth (when $s$ is far from the real axis).

**Step 3.** Extract the asymptotics of the coefficients (with the Landau Theorem) and exhibit a quasi-powers behavior for the coefficients. With the Quasi-Powers Theorem, this entails the Gaussian law.

There are strong similarities to the approach introduced by Baladi and Vallée [1], who perform a similar probabilistic analysis on rationals, with an analogous series $Q(s, w)$. This explains the strong similarity between the two results: the Gaussian law obtained in the present paper, together with the precise asymptotic expansion of the expectation and the variance, is of the same type as in the rational case.

Even though Step 3 is the same in the two analyses, the present rqi framework introduces new important issues in the first two steps. In Step 1, the periodicity phenomenon leads to several Dirichlet series—not only the initial series $P(s, w)$—which take into account the “primitivity” vs “non-primitivity” of the period (as described in Section 2.5). In Step 2, we deal with two different function spaces:

- When $s$ is close to the real axis, we relate the bivariate series $P(s, w)$ to traces of operators. We then need spaces of analytic functions, where such traces are well-defined, as shown by Mayer [18].
- When $s$ is far from the real axis, the useful space is the space $C^1(I)$ of functions on the unit interval, as in [1]. Here, we adapt the results due to Pollicott and Sharp to obtain polynomial growth in the rqi framework (see Theorem 3.5).
Finally, whereas the Baladi–Vallée analysis in [1] uses a unique Dirichlet series $Q(s,w)$ and “stays” in the function space $C^1(I)$, the present, more involved analysis deals with various bivariate generating functions and two different function spaces. Each generating function, and each function space, plays a specific role. The occurrence of the trace in the present study explains the resemblances and the differences between constants which appear in the asymptotic estimates of the expectation and the variance, in the rqi case and in the rational case. This is made precise in Section 4.4.

1.2. Continued fractions and quadratic irrationals. Every real number $x \in ]0,1[$ admits a continued fraction expansion of the form

$$x = \frac{1}{m_1 + \frac{1}{m_2 + \frac{1}{m_3 + \cdots}}}$$

denoted as $x = [m_1, m_2, \ldots]$. Here, the coefficients $m_k$ are positive integers known as partial quotients and also called digits. Rational numbers have a finite continued fraction expansion. When the number $x$ is irrational, the continued fraction expansion is infinite and $x$ is completely determined by the whole sequence $(m_k)$ of its digits.

In 1770, Lagrange proved that a number is quadratic irrational if and only if its continued fraction expansion is eventually periodic. We are mainly interested here in particular quadratic irrational numbers, whose continued fraction expansion is purely periodic. Such numbers are called reduced (rqi for short). This paper is devoted to their study; in particular, we describe the probabilistic behavior of such numbers in terms of “costs” that are defined via their continued fraction expansion.

Size. If $p(x)$ is the length of the smallest period of an rqi number $x$, then $x$ is defined by the relation $[m_1, m_2, \ldots, m_k + x] = x$ with $k = p(x)$ and is denoted by $(m_1, \ldots, m_p)$. This relation can be rewritten as a quadratic polynomial equation of the form $Ax^2 + Bx + C = 0$ with a triple $(A, B, C)$ of relatively prime integers. Then $x$ belongs to the quadratic field $\mathbb{Q}(\sqrt{\Delta})$ where $\Delta = B^2 - 4AC$ is the discriminant of the polynomial, and one associates with $x$ the fundamental unit $\epsilon(x) > 1$ of this quadratic field. This fundamental unit $\epsilon(x)$ plays, for rqi numbers, the same role as the denominator for rationals, and it defines a natural notion of size.

In this paper, we consider the set $\mathcal{P}$ of rqi numbers $x$, together with its finite \(^{(1)}\) subsets $\mathcal{P}_N$ formed by the rqi numbers of size $\epsilon(x)$ at most $N$:

$$\mathcal{P} := \{x \in I \mid x \text{ is a rqi number}\}, \quad \mathcal{P}_N := \{x \in \mathcal{P} \mid \epsilon(x) \leq N\}.\)  

\(^{(1)}\) The finiteness of $\mathcal{P}_N$ is proven at the end of Section 2.3.
Cost on digits. Any non-zero map $c : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is called a digit-cost. With a given digit-cost $c$, we associate the cost $C(x)$ of an rqi number $x$ defined as the “total cost” on the (smallest) period of $x$:

$$C(x) := c(m_1) + \cdots + c(m_p) \quad \text{if} \quad x = (m_1, \ldots, m_p).$$

This defines a non-zero cost $C : \mathcal{P} \to \mathbb{R}_{\geq 0}$. Moreover, we restrict ourselves to digit-costs $c$ that satisfy $c(m) = O(\log m)$ and are said to be of moderate growth\(^{(2)}\).

There are three main examples of digit-costs $c$ that can be treated by our methods: the unit cost $c = 1$; the characteristic function $\chi_n$ of a given digit $n$; the length $\ell$ of the binary expansion of integers, defined as $\ell(m) := \lfloor \log_2 m \rfloor + 1$. The associated costs $C$ are natural and interesting. For $c = 1$, $C(x)$ coincides with the period length $p(x)$. For $c = \chi_n$, $C(x)$ equals the number of digits equal to $n$ in the smallest period of $x$. Finally, for $c = \ell$, $C(x)$ is the number of binary digits needed to store the rqi $x$.

Interpretation in the hyperbolic plane. We now describe the classical geometric interpretation of the size $\epsilon(x)$ and the digits $m_j(x)$ of an rqi number $x$ in the hyperbolic plane.

We first recall the coding of a curve in the hyperbolic plane $\mathbb{H} := \{z = x + iy \in \mathbb{C} \mid y > 0\}$. When $\mathbb{H}$ is endowed with the usual metric $ds = |dz|/y$, the geodesics are the semicircles centered on the real axis together with the vertical lines. Consider the hyperbolic triangle $\Delta$ with the three cusps $i\infty$, 0 and 1, represented in Fig. 1. Together with all the triangles $h(\Delta)$ which are the transforms of $\Delta$ with $h \in \text{SL}_2(\mathbb{Z})$, it defines the Farey tessellation of $\mathbb{H}$ represented in Fig. 2. When a “good” oriented curve $\gamma$ of $\mathbb{H}$ goes through such a triangle $h(\Delta)$ in a general position, there are two possibilities for the position of the curve $h(\Delta) \cap \gamma$ with respect to the three cusps $h(i\infty)$, $h(0)$ and $h(1)$: if there is only one cusp on the right of the curve (and then two

\[^{(2)}\] The reasons for such a restriction are explained at the end of Section 3.2.
cusps on the left), we code the portion of the curve by $R$; otherwise, we code it by $L$ (see Fig. 1).

We now return to an rqi $x$, and consider its minimal even period, of length $e(x)$, with $e(x) := p(x)$ if $p(x)$ is even, and $e(x) := 2p(x)$ if $p(x)$ is odd. First, if $x \in ]0,1[$ is an rqi number, its conjugate $\bar{x}$ satisfies $\bar{x} < -1$ and

$$x = \langle m_1, \ldots, m_e \rangle \implies -1/\bar{x} = \langle m_e, m_{e-1}, \ldots, m_1 \rangle.$$  

The geodesic $\gamma(x)$ which links $\bar{x}$ to $x$ (with this orientation) intersects the imaginary axis at it$(x)$. This defines two oriented curves: $\gamma_+(x)$ that links it$(x)$ to $x$, and $\gamma_-(x)$ that links it$(x)$ to $\bar{x}$. As shown by Series [25] and Pollicott [21], the coding of the geodesics “copies” the continued expansions of the associated rqi numbers, and the codings of the two curves $\gamma_+(x)$ and $\gamma_+(-1/\bar{x})$ are the respective periodic infinite words

$$(L^{m_1}R^{m_2}L^{m_3} \ldots R^{m_e})^\mathbb{N}, \quad (L^{m_e}R^{m_{e-1}}L^{m_{e-2}} \ldots R^{m_1})^\mathbb{N}.$$  

The close connection between the codings of the curves $\gamma_+(-1/\bar{x})$ and $-\gamma_-(x)$ finally entails a coding for the concatenation $\gamma(x)$ of the curves $(-\gamma_-(x))$ and $\gamma_+(x)$ as the bi-infinite periodic word of period $R^{m_1}L^{m_2}R^{m_3} \ldots L^{m_e}$ (see Fig. 2).

![Fig. 2. Farey tessellation and continued fractions. Here, we consider $x := \phi^{-1}$ (\phi is the golden ratio) and the geodesic which links $x$ to its conjugate $-\phi$ as it goes through the Farey tessellation. On the right, the coding of $x$ in terms of $R$ and $L$.](image)

Moreover, the length of the “primitive” part of the geodesic $\gamma(x)$—associated with the coding $L^{m_1}R^{m_2}L^{m_3} \ldots R^{m_e}$ of the minimal even period—equals $2 \log \epsilon(x)$, where $\epsilon(x)$ is the fundamental unit associated with the rqi $x$.

These interpretations of the size $\epsilon(x)$ and the digits $m_j(x)$ of the rqi $x$ on its geodesic $\gamma(x)$ provide a geometric framework for the whole present study.
1.3. **Statement of the main results.** We study the probabilistic behavior of the continued fraction expansion of an rqi number, with respect to some cost $C$. When the set $\mathcal{P}_N$ is endowed with the uniform probability $\mathbb{P}_N$, the restriction $C_N$ of the cost $C$ to $\mathcal{P}_N$ becomes a random variable. Our first main result exhibits the asymptotic Gaussian behavior of the sequence $(C_N)$.

**Theorem 1.1.** Consider the set $\mathcal{P}$ of reduced quadratic irrational numbers $x$, endowed with the size $\epsilon$. With a non-zero digit-cost of moderate growth, associate the additive cost $C$ on $\mathcal{P}$. Then the following holds on the set $\mathcal{P}_N$ of reduced quadratic irrational numbers $x$ with $\epsilon(x) \leq N$, for $N \to \infty$:

(i) There are four constants, two dominant constants $\mu(c)$ and $\nu(c)$ that are strictly positive, and two subdominant constants, $\mu_1(c)$ and $\nu_1(c)$, for which the expectation $\mathbb{E}_N[C]$ and the variance $\mathbb{V}_N[C]$ satisfy the following asymptotic estimates, for some $\beta > 0$:

\[
\mathbb{E}_N[C] = \mu(c) \log N + \mu_1(c) + O(N^{-\beta}), \\
\mathbb{V}_N[C] = \nu(c) \log N + \nu_1(c) + O(N^{-\beta}).
\]

(ii) The distribution of $C$ is asymptotically Gaussian,

\[
\mathbb{P}_N \left[ x \left| \frac{C(x) - \mu(c) \log N}{\sqrt{\nu(c) \log N}} \leq t \right. \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2} du + O \left( \frac{1}{\sqrt{\log N}} \right).
\]

We remark that the speed of convergence towards the Gaussian law is optimal, of the same order $(\log N)^{-1/2}$ as in the usual Central Limit Theorem.

1.4. **Main constants.** The next theorem gives a precise description of the four constants that appear in Theorem 1.1. We gather here various results of different types. The first two items (i) and (ii) are classical; items (iii) and (iv) show a strong analogy to results in [1].

**Theorem 1.2.** Consider a digit-cost $c$ of moderate growth and the associated additive cost $C$.

(i) Consider the dynamical system related to the Gauss map that underlies the continued fraction expansion, and the (weighted) transfer operator $H_{t,w}$ which associates with a function $f$ the function $H_{t,w}[f]$ defined as

\[(1.3) \quad H_{t,w}[f](x) := \sum_{m \geq 1} \frac{e^{w c(m)}}{(m + x)^2 t} f \left( \frac{1}{m + x} \right).
\]

When it acts on the function space $A_{\infty}(\mathcal{V})$ defined in (3.4), the operator $H_{t,w}$ has, for $(t, w)$ close to $(1, 0)$, a dominant eigenvalue denoted as $\lambda(t, w)$. 

(ii) The two partial derivatives of the map \((t, w) \mapsto \lambda(t, w)\) at \((t, w) = (1, 0)\) are related to the entropy \(E\) of the system together with the mean value \(\mathbb{E}[c]\) of the cost with respect to the Gauss density \(\psi : x \mapsto (1/\log 2)(1/(1 + x))\),

\[-\lambda'_t(1, 0) = E = \frac{\pi^2}{6 \log 2}, \quad \lambda'_w(1, 0) = \mathbb{E}[c] = \sum_{m \geq 1} c(m) \int_{1/m}^{1/(m+1)} \psi(x) \, dx.\]

(iii) Consider the map \(w \mapsto \sigma(w)\) defined from the equation \(\lambda(\sigma(w), w) = 1\) with \(\sigma(0) = 1\), together with the two mappings

\[U(w) := 2(\sigma(w) - 1), \quad V(w) := \log \left( \frac{\lambda'_t(1, 0)}{\lambda'_t(\sigma(w), w)} \right) - \log \sigma(w).\]

The constants of Theorem 1.1 can be expressed in terms of the first two derivatives of \(U\) and \(V\):

\[
\mu(c) = U'(0) = \frac{2}{E} \mathbb{E}[c], \quad \nu(c) = U''(0), \quad \mu_1(c) = V'(0), \quad \nu_1(c) = V''(0).
\]

(iv) The dominant constants \(\mu(c), \nu(c)\) coincide with their analogs that appear in the analysis of rational trajectories described in [1].

We first observe the equality \(\mu(c) = (2/E)\mathbb{E}[c]\) that entails the estimate \(\mathbb{E}_N[C] \sim \mathbb{E}_N[p] \cdot \mathbb{E}[c]\). This estimate may be compared to the ergodic relation that holds on almost any generic trajectories. This leads to the following (informal) statement: “The periodic trajectories behave on average as the generic trajectories behave almost everywhere”, which holds in a very general context of “dynamical analysis”.

All the constants that appear in Theorem 1.1 are computable. In particular, the constants \(\mathbb{E}[c]\) related to the three costs of interest are

\[
\mathbb{E}[c] = 1, \quad \mathbb{E}[\chi_n] = \frac{1}{\log 2} \log \left[ \frac{(n+1)^2}{n(n+2)} \right], \quad \mathbb{E}[\ell] = \frac{1}{\log 2} \prod_{i=1}^{\infty} \log \left( 1 + \frac{1}{2^k} \right).
\]

In many natural situations, the mean value \(\mathbb{E}[c]\) admits an explicit expression, and hence so does the constant \(\mu(c)\) itself.

The situation is completely different for the other three constants, as there are no longer closed expressions for the second (and third) derivatives of \((t, w) \mapsto \lambda(t, w)\) at \((1, 0)\). However, these derivatives are computable in polynomial time with methods first described in [11] and [5], and further developed in [16]. All these methods are based on a very natural idea: use the truncated Taylor expansion of analytic functions to approximate the operator by matrices (which now act on polynomials); compute the spectrum of the matrices with classical tools of linear algebra (in finite dimensions). This should provide approximations of the (upper part of) the spectrum of the operator (and this is proven to be true).
1.5. Description of the methods. The present work deals with methods that come from analytic combinatorics, along the lines of the book of Flajolet and Sedgewick [8]. We associate with the set $\mathcal{P}$, endowed with the size $\epsilon$ and the cost $C$, a bivariate generating function; in the present number theory framework, it is of Dirichlet type,

$$P(s, w) := \sum_{x \in \mathcal{P}} \exp(wC(x))\epsilon(x)^{-s}.$$ 

Cumulative costs over the set $\mathcal{P}_N$ are written as sums of coefficients, and we define

$$S^{[C]}_w(N) := \sum_{x \in \mathcal{P}, \epsilon(x) \leq N} \exp(wC(x));$$

the moment generating function of the cost $C$ on $\mathcal{P}_N$ is then written as the quotient of such sums,

$$\mathbb{E}_N[\exp(wC)] = \frac{S^{[C]}_w(N)}{S^{[C]}_0(N)}.$$ 

The distribution of the cost $C$ on $\mathcal{P}_N$ is studied here with this moment generating function: we use the Quasi-Powers Theorem [12] that states that an asymptotic Gaussian law for $C$ on $\mathcal{P}_N$ holds if the moment generating function has a “uniform quasi-powers” form (when $w$ is a complex number near 0).

We therefore need precise asymptotic estimates on the sums $S^{[C]}_w(N)$ of the series $P(s, w)$; analytic combinatorics principles relate them to the analytic properties of $P(s, w)$. The strong tool which enacts this transfer here, between analytic properties of $P(s, w)$ and asymptotic properties of its coefficients, is the Landau Theorem [15]. The analytical properties that are needed for $s \mapsto P(s, w)$ are of two types: a precise knowledge of its singularities (here, its poles, located near the real axis) together with a good knowledge of its behavior for $|\Im s| \to \infty$ (here, polynomial growth for $|\Im s| \to \infty$).

We then have to study the series $P(s, w)$. This will be done with “dynamical analysis”, which has already been used in many works, in particular in the analogous rational study (see for instance [11, 29, 28]). The idea is to use the underlying dynamical system, associated with the continued fraction transformation (the Gauss map), and relate the series $P(s, w)$ to the (weighted) transfer operator of the dynamical system described in (1.3). One then has to study the transfer operator itself, for $w$ close to 0, and various values of $s$.

Here, as already mentioned in Section 1.1 we need two different function spaces: the first one, useful when $s$ is close to the real axis, is the set $\mathcal{A}_\infty(\mathcal{V})$ of analytic functions (defined in (3.4)) where the transfer operator
has dominant spectral properties and a well-defined trace (in the sense of Grothendieck); this study provides a precise knowledge of the singularities of \( s \mapsto P(s, w) \) located near the real axis; the second function space, useful for \(|\Im s| \to \infty\), is the set of \( C^1 \) functions, where the transfer operator admits bounds à la Dolgopyat [6], which entail the polynomial growth of \( P(s, w) \) for \(|\Im s| \to \infty\).

### 1.6. Comparison with the Baladi–Vallée approach.

As already mentioned in Section 1.1, our methods are rather similar to those used for rational trajectories in [1]. The paper [1] deals with the generating function of the rational set \( Q := \mathbb{Q} \cap [0, 1[ \), endowed with the denominator size \( q(x) \), namely

\[
Q(s, w) := \sum_{x \in \mathbb{Q}} \exp(wC(x))q(x)^{-s}.
\]

This series can be directly expressed in terms of the weighted transfer operator \( \mathbf{H}_{t,w} \) introduced in (1.3), as the equality \( Q(s, w) = (I - \mathbf{H}_{t,w})^{-1}[1][0]^{(3)} \) holds with \( t = s/2 \). This explains why the first two steps of the rational study are easier than in the present one, which is more involved due to the three issues described in Section 1.1.

In Step 3, the main tools are essentially the same, namely, the Landau Theorem \( ^{(4)} \) and the Quasi-Powers Theorem.

### 1.7. Comparison with already known results

**Plain periodic trajectories.** The periodic trajectories of dynamical systems are very well studied, by means of various zeta series, as pioneered by Ruelle [24]. In the case of the Gauss map, the Selberg zeta series is a powerful tool that is well-adapted to plain periodic trajectories (without a cost \( C \)), and there is a close connection exhibited by Mayer [17] between the plain Dirichlet series \( P(s) := P(s, 0) \), the Selberg zeta series and the Riemann zeta function. This leads to the asymptotic estimate of the cardinality, \( |\mathcal{P}_N| \sim \left(\frac{3 \log 2}{\pi^2}\right) \cdot N^2 \). The study of the remainder term is more involved and was performed by Boca [3] who obtained an error term of order \( O_{\varepsilon}(N^{7/4+\varepsilon}) \) for any \( \varepsilon > 0 \). His proof relies on precise estimates on the distribution of pairs \((x, y)\) of coprime integers that satisfy \( xy = 1 \mod q \) (for an integer \( q \)), together with classical results on the Riemann zeta function. Following this approach, Ustinov [26] improved the estimate of the error term and obtained an order \( O(N^{3/2} \log^4 N) \).

\( ^{(3)} \) This is not completely exact, as the final operator \( \mathbf{F}_{t,w} \) also intervenes (see Section 4.4).
\( ^{(4)} \) The paper [1] does not use the “ready for use” version of the Landau Theorem described in Section 4.2 and, as in many other works (see for instance [22]), it proves the analogous result “by hand”.

---

\( \mathbf{Gaussian \ behavior \ of \ quadratic \ irrationals} \)
Our series $P(s, w)$ can be viewed as a “twisted Selberg zeta series”, where the twist is effected by the cost $C$; many geometric properties of the Selberg zeta series “disappear” with this twist and it is not clear how to extend Boca’s and Ustinov’s methods to weighted periodic trajectories.

**Weighted periodic trajectories.** To the best of our knowledge, Parry and Pollicott [20] and Pollicott [21] were the first to study weighted periodic trajectories with analytic combinatorics methods. Pollicott indeed introduced a weighted generating function, and, as he only performed average-case analysis, he could use Tauberian theorems that provide estimates of the mean values, as in Theorem 1.1(i), but with only the first (dominant) term, without remainder terms. Later on, Faiivre [7] used slightly different methods, and obtained the same results.

With similar methods, Kelmer [14] studied the set $P^{(n)}_N$ of rqi numbers whose alternate sum of partial quotients is fixed and equal to $n$. He considered the restriction of a function $f \in C([0,1])$ to the set $P^{(n)}_N$ and proved that, for $n$ fixed and $N \to \infty$, its mean value tends to the mean value $E[f]$ of $f$ with respect to the Gauss density $\psi$. He did not specify the speed of convergence.

**Remainder terms?** One may expect more precise asymptotic estimates for the mean value (estimates with remainder terms) or for the variance (with even a rough estimate); such estimates are also needed for a distributional study. They appear to be based on a more precise knowledge of the weighted transfer operator $H_{t,w}$ for $|3t| \to \infty$. This better knowledge was indeed brought by works of Dolgopyat, notably [6], first for unweighted transfer operators associated with Markovian dynamical systems with a finite number of branches. Then Pollicott and Sharp used Dolgopyat’s result (see for instance [22]), and performed various average-case analyses, obtaining estimates for various mean values (related to restricted periodic trajectories) with remainder terms.

Dolgopyat’s results were further adapted in [1] to the present case of interest, a weighted transfer operator associated with a Markovian dynamical system with an infinite denumerable number of branches. Then Baladi and Vallée [1] used this extension of Dolgopyat’s result in distributional studies of weighted rational trajectories and derived asymptotic Gaussian laws. To the best of our knowledge, the present work is the first one devoted to distributional studies of weighted periodic trajectories.

**1.8. Plan of the paper.** Section 2 presents the main objects, and the general methodology; it introduces various generating functions of interest, describes their relations and their first properties. Section 3 introduces the weighted transfer operator, the two useful function spaces $A_{\infty}(\mathcal{V})$ and $C^1(\mathcal{I})$,
and describes their role in the analysis. It relates analytic properties of the generating functions to fine properties of the weighted transfer operator: bounds à la Dolgopyat in $C^1(I)$, and properties of the traces in $A_{\infty}(V)$. As our work requires a precise (and long) study of traces, we perform it in the Appendix and just summarize the main results in Section 3. Section 4 uses two main tools (the Landau Theorem and the Quasi-Powers Theorem) to transfer analytic properties of the bivariate generating functions into probabilistic properties of the cost, notably the asymptotic Gaussian law (Theorem [1.1]). It also contains a precise study of the main constants of interest (Theorem [1.2]). In Section 5, we explain how to “transfer” to periodic trajectories some results that are already proven for rational trajectories. We study in particular the case of “constrained” periodic trajectories. The paper ends with an Appendix, which is a self-contained study of traces of operators. It is needed in Section 3 and may be of independent interest for people who are not specialists of this theory.

The results that are already known and directly used in this paper are labeled with letters, whereas the new results that are proven in the paper are labeled with numbers.

2. Presentation of the main objects. The present section introduces the main actors: first, in Section 2.1 the underlying dynamical system associated with the Gauss map, together with the main properties of its branches (Section 2.2). We then recall well-known facts about irrational quadratic numbers and fundamental units in Section 2.3 and describe the costs of interest (of moderate growth) (Section 2.4). Then, Section 2.5 introduces various (bivariate) generating functions of interest. The initial generating function $P(s,w)$ involves quadratic numbers, and deals with a primitivity condition that is difficult to manage. We first “suppress” this primitivity condition, and then we make a change of size: we replace the initial size $\epsilon$ by another (closely related) size $\alpha$, which has multiplicativity properties. We finally obtain a series $E(s,w)$. The end of the section describes the first easy properties of those generating functions and explains why it suffices to deal with $E(s,w)$ throughout the rest.

2.1. The Euclidean dynamical system. The continued fraction expansion encodes the trajectories of the dynamical system associated with the Gauss map $T : I \to I$,

\begin{equation}
I := [0, 1], \quad T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \quad \text{for} \; x \neq 0, \quad T(0) = 0.
\end{equation}

Here, $\lfloor x \rfloor$ is the integer part of $x$. The trajectory

$\mathcal{T}(x) = (T(x), T^2(x), \ldots)$
of an irrational $x$ never meets 0 and is encoded by the infinite sequence
\[(m_1(x), m_2(x), \ldots) \quad \text{with} \quad m_k(x) := m(T^{k-1}(x)), \quad m(x) := \lfloor 1/x \rfloor.\]
The map $T$ is a piecewise complete interval map, with its set $\mathcal{H}$ of inverse branches,
\[\mathcal{H} = \left\{ h_m \big| x \mapsto \frac{1}{m + x}; m \geq 1 \right\}.\]
For any $k \geq 1$, the set $\mathcal{H}_k$ of inverse branches of $T^k$ consists of the LFTs (linear fractional transformations) of the form
\[h = h_{m_1} \circ \cdots \circ h_{m_k}.\]
By definition, the depth of a LFT $h \in \mathcal{H}_k$ (denoted by $|h|$) is equal to $k$. The set
\[\mathcal{H}^+ := \bigcup_{k \geq 1} \mathcal{H}_k\]
comprises the inverse branches of any strictly positive depth. For $h \in \mathcal{H}^+$, the interval $h(I)$ is called the fundamental interval relative to $h$. For a LFT $h$ described in (2.2), the interval $h(I)$ consists of the reals $x$ that have their $i$th partial quotients equal to $m_i$ for any $i \in [1, \ldots, k]$. For any $k \geq 1$, the fundamental intervals $h(I)$ for $h \in \mathcal{H}_k$ form a topological partition of the unit interval $I$.

2.2. First properties of the dynamical system. Here, we describe the main properties (of geometric flavor) satisfied by the branches of the set $\mathcal{H}^+$ defined in Section 2.1 and their derivatives. The intervals of interest are $I := [0, 1]$ and $J := [-1/4, 9/4]$; the latter intervenes because of Property (P3) associated with $r = 5/4$.

(P1) There is a constant $L$ (usually called the distortion constant, here equal to $\text{sup}(8/3, \exp(20/3))$) for which the following two inequalities hold for any inverse branch $h \in \mathcal{H}^+$:
\[|h''(x)| \leq L|h'(x)|, \quad \forall x \in J, \quad \frac{1}{L} \leq \frac{|h'(x)|}{|h'(y)|} \leq L, \quad \forall (x, y) \in J^2.\]

(P2) The map $T$ is piecewise strongly expanding. There is a contraction ratio $\rho$ equal to $\phi^{-2}$ for which the following inequalities hold and involve the fixed point $x_h$ of the inverse branch $h$ and the distortion constant $L$:
\[|h'(x_h)| \leq \rho|h|, \quad |h'(x)| \leq L\rho|h|, \quad \forall x \in J, \forall h \in \mathcal{H}^+.\]

(P3) Consider the disk $\mathcal{V}_r$ of center 1 and radius $r$. For $r \in ]1, (1 + \sqrt{5})/2[\}$, there exists $\tilde{r} < r$ for which the following hold, for any branch $h \in \mathcal{H}^+$:
(i) $h$ is analytic on $\mathcal{V}_r$, and $h(\mathcal{V}_r) \subset \mathcal{V}_{\tilde{r}}$.
(ii) The analytic extension $\hat{h}$ of the map $x \mapsto |h'(x)|$ is non-zero on $\mathcal{V}_r$ and the function $z \mapsto \log \hat{h}(z)$ is analytic on $\mathcal{V}_r$. 
Gaussian behavior of quadratic irrationals

In the following we consider the case \( r = 5/4 \) and denote \( V := V_{5/4} \).

(P4) The series
\[
\sum_{h \in H} h^s(z), \quad \sum_{h \in H} \log h(z) h^s(z)
\]
are absolutely convergent for any \( z \in V \) and on compact subsets of the half-plane \( \Re s > 1/2 \).

(P5) The UNI (Uniform Non-Integrability) condition holds. It is described by the “distance” \( \Delta(h, k) \) between two inverse branches \( h \) and \( k \) of the same depth (introduced in [I]),
\[
\Delta(h, k) := \inf_{x \in I} |\Psi_{h,k}'(x)| \quad \text{with} \quad \Psi_{h,k}(x) := \log \frac{|h'(x)|}{|k'(x)|},
\]
and the measure \( J(h, \eta) \) of the “ball” of center \( h \) and radius \( \eta > 0 \),
\[
J(h, \eta) := \sum_{\Delta(h,k) \leq \eta} |k(\mathcal{I})| \quad (\text{for } h \in H^n).
\]
The condition UNI states that there exists a constant \( K \) for which
\[
J(h, \rho^{an}) \leq K \rho^{an}, \quad \forall \eta \in ]0, 1[, \forall n, \forall h \in H^n.
\]
There is also a technical statement about the second derivatives \( \Psi_{h,k}''(x) \):
\[
Q := \sup \{ |\Psi_{h,k}''(x)| \mid n \geq 1, h, k \in H^n, x \in \mathcal{I} \} < \infty.
\]

2.3. Quadratic irrational numbers. Quadratic irrational numbers are the only numbers \( x \) for which the trajectory \( T(x) \) is eventually periodic, and a reduced quadratic irrational number (rqi for short) is associated with a purely periodic trajectory. An rqi \( x \) is then completely defined by its (least) period \( p(x) \) which defines a primitive cycle (of the same length).

The set of rqi numbers coincides with the set of fixed points \( x_h \) of LFTs \( h \in H^+ \). All the LFTs which determine a given rqi number \( x \) are the powers of a given LFT which is primitive. This defines a bijection between the set \( P \) of rqi numbers and the set of primitive LFTs, together with a decomposition of \( H^+ \) as a disjoint sum,

\[
H^+ = \bigcup_{k \geq 1} \mathcal{P}^{(k)}, \quad \mathcal{P}^{(k)} := \{ h^k \mid h \in \mathcal{P} \}.
\]

A fundamental quantity associated with a rqi number \( x \) is the product of all the shifted \( T^i(x) \) along its period. If \( h \) is the primitive LFT which defines \( x = x_h \), one considers
\[
\alpha(x) := \prod_{i=0}^{p(x)-1} T^i(x) = |h'(x_h)|^{1/2}.
\]
Here, the fundamental unit $\epsilon(x)$ is chosen as the size for the rqi $x$ (see Section 1.2). It is related to the product $\tilde{\alpha}(x)$ of all the shifted $T^i(x)$ along the *minimal even* period of length $e(x)$ (already mentioned in Section 1.2), and we have

$$\epsilon(x) = \tilde{\alpha}(x)^{-1}, \quad \tilde{\alpha}(x) := \prod_{i=0}^{e(x)-1} T^i(x),$$

which also translates as

$$\epsilon(x) = \alpha(x)^{-r(x)} \quad \text{with} \quad r(x) := \begin{cases} 1 & \text{for even } p(x), \\ 2 & \text{for odd } p(x). \end{cases}$$

It is then natural (and useful) to extend the definition of $\alpha, \epsilon, r$ from the set $P$ of rqi numbers to the set $\mathcal{H}^+$. For any $h \in \mathcal{H}^+$ (primitive or not), we define

$$\alpha(h) := \prod_{i=0}^{\lfloor |h| \rfloor - 1} T^i(x_h) = |h'(x_h)|^{1/2},$$

$$\epsilon(h) := \alpha(h)^{-r(h)} \quad \text{with} \quad r(h) := \begin{cases} 1 & \text{for even } |h|, \\ 2 & \text{for odd } |h|. \end{cases}$$

With this extension, $\alpha, \epsilon$ are now defined on $\mathcal{H}^+$ and $\alpha$ has the multiplicativity property $\alpha(h^k) = (\alpha(h))^k$. Now, the set $\mathcal{H}^+$ is endowed with the size $\epsilon$ (closely related to $\alpha$) and a cost $c$.

The rational trajectories are classically studied with the denominator size: one associates with $h \in \mathcal{H}^+$ the denominator $q(h) = |h'(0)|^{-1/2}$. The distortion constant $L$ introduced in Property (P1) of Section 3.1 now relates the two sizes $\epsilon$ and $q$, defined on $\mathcal{H}^+$, via the inequality

$$L^{-1/2} \epsilon(h)^{1/2} \leq L^{-1/2} \alpha(h)^{-1} \leq q(h) \leq L^{1/2} \alpha(h)^{-1} \leq L^{1/2} \epsilon(h),$$

which also proves that the set $\mathcal{P}_N$ of rqi numbers with size $\epsilon(x) \leq N$ is finite, and satisfies $|\mathcal{P}_N| = O(N^2)$.

**2.4. Cost of moderate growth.** A digit-cost $c : \mathbb{N}^* \to \mathbb{R}_{\geq 0}$ is also a mapping $c : \mathcal{H} \to \mathbb{R}_{\geq 0}$ defined by $c(h_m) := c(m)$. We next extend $c$ to $\mathcal{H}^+$ in an additive way and define

$$c : \mathcal{H}^+ \to \mathbb{R}_{\geq 0} \quad \text{by} \quad c(h) := \sum_{i=1}^{k} c(h_i) \quad \text{if} \quad h = h_1 \circ \cdots \circ h_k.$$
We first prove that the previous relation extends to any inverse branches will be central in the proof of Proposition 2.5.

**Definition 2.1.** A digit-cost \( c \) is of moderate growth if there exists \((A, B)\) with \( A, B \geq 0 \) and \( A + B > 0 \) for which

\[
(2.6) \quad c(m) \leq A \log m + B \quad \text{for all } m \geq 1.
\]

In the present rqi framework, a precise comparison between the various generating functions needs a precise study of these costs (which was not needed in previous studies, for instance in [1]). In particular, relation (2.7) below between costs and derivatives of inverse branches will be central in the proof of Proposition 2.5.

**Lemma 2.2.** For a digit-cost \( c \) of moderate growth, there exists a pair \((K, d)\) of strictly positive numbers for which

\[
(2.7) \quad e^{c(h)} \leq |h'(x_h)|^{-d}, \quad e^{c(h)} \leq K |h'(0)|^{-d} \quad \text{for any } h \in \mathcal{H}^+.
\]

The cost is then said to be of exponent \( d \).

**Proof.** For any \( h \) of depth 1, of the form \( h = h_m \), the equality \( 2 \log m = |\log |h'(0)|| \) holds and relation (2.6) reads

\[
c(h) \leq (A/2)|\log |h'(0)|| + B \quad \text{for all } h \in \mathcal{H}.
\]

We first prove that the previous relation extends to any \( h \in \mathcal{H}^+ \) as

\[
(2.8) \quad c(h) \leq (A/2)|\log |h'(0)|| + B|h| \quad \text{for all } h \in \mathcal{H}^+,
\]

where \( |h| \) is the depth of the LFT \( h \). One indeed performs an easy induction; when \( c = h_1 \circ h_2 \), one has

\[
c(h_1 \circ h_2) = c(h_1) + c(h_2) \leq \frac{A}{2}|\log |h_1'(0)|| + \frac{A}{2}|\log |h_2'(0)|| + B|h_1| + B|h_2|.
\]

Now, for any \( x \in [0, 1] \) we have \( |h'(0)| \geq |h'(x)| \), and

\[
|\log |h_1'(0)|| + |\log |h_2'(0)|| \leq |\log |h_1'(h_2(0))|| + |\log |h_2'(0)|| = |\log |h'(0)||.
\]

Relation (2.8) is now proven, and implies

\[
c(h) \leq (A/2)|\log |h'(x_h)|| + B|h|.
\]

Using Property (P2) with \( A := 2a, B := -b \log \rho \) now yields

\[
e^{c(h)} \leq \exp[-a \log |h'(x_h)||] \exp[B|h|] = |h'(x_h)|^{-a} \rho^{-b|h|} \leq |h'(x_h)|^{-(a+b)}.
\]

With the distortion constant \( L \), this entails the second inequality with \( K = L^{a+b} \).

**2.5. Generating functions of interest.** When the set \( \mathcal{P} \) of rqi numbers is endowed with the size \( \epsilon \) and an additive cost \( C \) that comes from a digit-cost \( c \), its basic generating function is a Dirichlet weighted generating function of the form

\[
(2.9) \quad P(s, w) := \sum_{x \in \mathcal{P}} \exp(wC(x))\epsilon(x)^{-s},
\]
where the variable $w$ “marks” the cost $C$ and the variable $s$ “marks” the size $\epsilon$.

We now define two variants of series $P$, namely series $Z$ and series $E$, both closely related to the initial series $P$ but easier to deal with. The series $Z(s, w)$ “avoids the primitivity condition”, whereas $E(s, w)$ deals with size $\alpha$ instead of size $\epsilon$.

**Series $Z$.** To avoid the primitivity condition, we “replace” the series $P(s, w)$ by the “total” series

$$Z(s, w) := \sum_{h \in H^+} \exp(wc(h))\epsilon(h)^{-s},$$

and the decomposition (2.3) leads to the decomposition

$$Z(s, w) = P(s, w) + A(s, w),$$

with

$$A(s, w) := \sum_{k \geq 2} P_k(s, w) \quad \text{and} \quad P_k(s, w) := \sum_{h \in p^{(k)}} \exp(wc(h)) \cdot \epsilon(h)^{-s}.$$  

Proposition 2.5 will prove that the “main” term in (2.11) comes from $P(s, w)$.

**Series $E$.** We wish to deal with size $\alpha$ (instead of $\epsilon$). Size $\alpha$ indeed fulfills multiplicative properties described in Section 2.3. This is why we introduce the series

$$Y(s, w) := \sum_{k \geq 1} Y_k(s, w)$$

with

$$Y_k(s, w) := \sum_{h \in H^k} \exp(wc(h))\alpha(h)^{2s} = \sum_{h \in H^k} \exp(wc(h))|h'(x_h)|^s.$$  

We also write $Z(s, w)$ as

$$Z(s, w) := \sum_{k \geq 1} Z_k(s, w), \quad Z_k(s, w) := \sum_{h \in H^k} \exp(wc(h))\epsilon(h)^{-s}.$$  

The notion of minimal even period leads to relation (2.5) between $\epsilon$ and $\alpha$. This entails the following relations between the series $Z_k(s, w)$ and $Y_k(s, w)$:

$$Z_k(s, w) = \begin{cases} Y_k(s/2, w) & \text{for } k \text{ even}, \\ Y_k(s, w) & \text{for } k \text{ odd}, \end{cases}$$

and leads to the decomposition

$$Z(s, w) = E(s, w) + O(s, w),$$

where the even and odd parts of $Z(s, w)$ are related to components $Y_k(s, w)$.  

E(s, w) = \sum_{k \text{ even}}^{k \geq 2} Y_k(s/2, w), \quad O(s, w) = \sum_{k \text{ odd}}^{k \geq 1} Y_k(s, w).}

Proposition 2.5 will prove that the “main” term in (2.13) comes from $E(s, w)$.

Sets $\Gamma(a)$. Each term of the generating functions involves a product $|h'(x)|\exp(wh(h))$. With a cost $c$ (of moderate growth) of exponent $d$ and a real $a > 0$, we associate the sets

\begin{equation}
C_a := \{(\sigma_0, \nu_0) \mid \sigma_0 > 1/2, \nu_0 > 0, \sigma_0 - d\nu_0 = a\},
\end{equation}

\begin{equation}
\Gamma(a) := \{(t, w) \mid \Re t > \sigma_0, |\Re w| < \nu_0, (\sigma_0, \nu_0) \in C_a\},
\end{equation}

that are well-adapted to the study of such products.

We will deal with $\Gamma(a)$ or some of its subsets of the form $S \times W$ where $W$ is a complex neighborhood of 0 of the form $\{|w| \leq \nu_0\}$ and $S$ is (part of a) vertical strip $\{s \mid \Re t > 1 - \delta_0\}$. Such a domain is contained in $\Gamma(a)$ if $1 - \delta_0 - d\nu_0 > a$ holds.

Series $Y$. We have introduced the three series $P, Z, E$, together with series $Y$. The series $Y$ will be a useful reference in this context, as we now explain:

**Lemma 2.3.** Consider the series $X(s, w)$ with $X \in \{P, Z, E\}$ and a pair $(s, w)$ with $\nu := \Re w$ and $\sigma := \Re s$.

(a) The following relations hold between $X$ and $Y$:

$|X(s, w)| \leq Y(\sigma/2, |\nu|)$, \quad $|P_k(s, w)| \leq Y((k/2)\sigma, k|\nu|)$.

(b) The following bound holds for each component $Y_k(s, w)$ on each $\Gamma(a)$, and involves the distortion constant $L$:

\begin{equation}
|Y_k(t, w)| \leq L\rho^{(a-1)k} \quad \text{for any } k \geq 1.
\end{equation}

**Proof.** (a) The first bound is due to the inequality $\epsilon(h)^{-\sigma} \leq \alpha(h)^\sigma = |h'(x_h)|^{\sigma/2}$ that holds for $\sigma > 0$. The second bound uses the multiplicativity of size $\alpha$, namely $\alpha(h^k) = \alpha(h)^k$.

(b) For a cost $c$ of exponent $d$, and a pair $(t, w)$ with $\sigma := \Re t$ and $\nu := \Re w$, Lemma 2.2 yields

\begin{equation}
|Y_k(s, w)| \leq Y_k(\sigma, |\nu|) = \sum_{h \in H^k} e^{\nu|c(h)|h'(x_h)|^{\sigma}} \leq \sum_{h \in H^k} |h'(x_h)|^{\sigma-d|\nu|}.
\end{equation}

Moreover, by Properties (P1) and (P2), one has

$\sum_{h \in H^k} |h(0) - h(1)| = 1, \quad \sum_{h \in H^k} |h'(0)| \leq L, \quad |h'(x_h)| \leq \rho^k \quad \text{for } h \in H^k$. 
This entails the bound
\[ \sum_{h \in \mathcal{H}^k} |h'(x_h)|^a \leq \rho^{(a-1)k} L \sum_{h \in \mathcal{H}^k} |h(0) - h(1)| \leq L \rho^{(a-1)k}, \]
and finally the bound (2.16) when \((t, w) \in \Gamma(a)\). \(\blacksquare\)

2.6. Analytical equivalence between generating functions. The inequality
\[ (2.18) \quad Y_k(1, 0) \geq \frac{1}{L} \sum_{h \in \mathcal{H}^k} |h(1) - h(0)| = \frac{1}{L} \quad \text{for any } k \geq 1 \]
entails that the series that defines \(E(s, w)\) is divergent at \((2, 0)\). The point \((2, 0)\) plays an important role, and we now wish to compare more precisely the three series of interest \(P, Z, E\) near the vertical line \(\Re s = 2\). We first introduce a definition:

**Definition 2.4.** Two Dirichlet series \(A(s, w)\) and \(B(s, w)\) are **analytically equivalent** if, for any neighborhood \(\mathcal{W}_0\) of \(w = 0\) small enough, there are reals \(\delta_0\) and \(\delta_1\) with \(\delta_1 > \delta_0 > 0\) for which the following hold:

(a) For any \(w \in \mathcal{W}\), the functions \(s \mapsto A(s, w)\), \(s \mapsto B(s, w)\) are analytic on the half-plane \(\mathcal{P}_0 := \{ s \mid \Re s > 2 + \delta_0 \}\) and bounded on \(\mathcal{P}_0 \times \mathcal{W}_0\).

(b) For any \(w \in \mathcal{W}\), the difference \(D(s, w) := A(s, w) - B(s, w)\) is analytic on the vertical strip \(\mathcal{S}_1 := \{ s \mid |\Re s - 2| < \delta_1 \}\) and bounded on \(\mathcal{S}_1 \times \mathcal{W}_0\).

The next proposition shows that the three series of interest are all analytically equivalent.

**Proposition 2.5.** For a cost of moderate growth, the three generating functions \(P(s, w), Z(s, w)\) and \(E(s, w)\) are analytically equivalent.

**Proof.** We use Lemma 2.3 and the series \(Y\) plays an important role in the proof.

**Condition (a).** Consider the neighborhood \(\mathcal{W}_0 := \{ w \mid |w| < \nu_0 \}\). The domain \(\mathcal{P}_0 \times \mathcal{W}_0\) is a subset of \(\Gamma(a)\) with \(a > 1\) if \(\delta_0 > 2d\nu_0\). Moreover, for \((s/2, w) \in \Gamma(a)\) with \(a > 1\), Lemma 2.3 entails that \(E(s, w), P(s, w)\) and \(Z(s, w)\) all satisfy condition (a).

**Condition (b) for \(Z(s, w) - P(s, w)\).** This difference is described in (2.11). For \(k \geq 2\), we use Lemma 2.3 together with (2.17) to get
\[ |P_k(s, w)| \leq P_k(\sigma, \nu) \leq Y((k/2)\sigma, k|\nu|) \leq \sum_{h \in \mathcal{H}^+} |h'(x_h)|^{(k/2)(\sigma - 2d|\nu|)}. \]
Assume now that \((s/2, w) \in \Gamma(a)\) with \(a > 1/2\). This means that \(\sigma - 2d|\nu|\) is at least \(\gamma > 1\) and let \(\gamma = 1 + 3\beta\) with \(\beta > 0\). Then, for any \(k \geq 2\), the inequality \((k/2)\gamma \geq (k - 1)\beta + 1\) holds and entails the bound
\[ |h'(x_h)|^{(k/2)(\sigma - 2d|\nu|)} \leq |h'(x_h)|^{(k-1)\beta}|h'(x_h)|, \]
and thus, in the same vein as in the proof of Lemma 2.3,
\[ Y((k/2)\sigma, k|\nu|) \leq L \sum_{n \geq 1} \rho^{n(k-1)\beta} \text{ for any } k \geq 2. \]

Finally, whenever \((s/2, w) \in \Gamma(a)\) with \(a > 1/2\), one obtains
\[ |Z(s, w) - P(s, w)| \leq \sum_{k \geq 2} Y((k/2)\sigma, k|\nu|) \leq L \sum_{n \geq 1} \sum_{k \geq 1} \rho^{nk\beta} \]
\[ = L \sum_{n \geq 1} \frac{\rho^{n\beta}}{1 - \rho^{n\beta}} \leq L \rho^\beta \frac{1}{(1 - \rho^\beta)^2}. \]

Now, set \(W_0 := \{ w | |w| < \nu_0 \}\). Then \(S_1 \times W_0 \subset \Gamma(a)\) for \(a > 1/2\) when \(\delta_1 < 1 - 2d\nu_0\). If furthermore \(\nu_0 < 1/(4d)\) then we can choose \(\delta_1 > \delta_0 > 2d\nu_0\), and condition (b) holds for \(Z(s, w) - P(s, w)\) on \(W_0\).

**Condition (b) for** \(Z(s, w) - E(s, w)\). This difference is equal to the odd series \(O(s, w)\). The bound \((2.16)\) implies that \(O(s, w)\) is absolutely convergent and uniformly bounded when \((s/2, w) \in \Gamma(a)\) with \(a > 1/2\). Then, for any neighborhood \(W_0\) with \(\nu_0 < 1/(4d)\), condition (b) holds for the same choice of \(\delta_1, \delta_0\) as before.

By Proposition 2.5, it will be sufficient to study \(E(s, w)\) around \(\Re s = 2\). The next section is devoted to this task.

3. Generating functions and weighted transfer operators. Beginning with the initial generating function \(P(s, w)\), we have related \(P(s, w)\) to \(E(s, w)\). In order to study the series \(E(s, w)\) more deeply, we introduce our main tool, the weighted transfer operator \(H_{t,w}\), and relate the series \(E(s, w)\) to the operator \(H_{t,w}\) and its iterates \(H^k_{t,w}\) with \(t = s/2\).

In Section 3.1, we define the transfer operators and make precise the two function spaces on which they act. We describe the first properties of these operators in Section 3.2. Then Section 3.3 explains the role of each function space. The space \(A_\infty(V)\) is a space of analytic functions where the traces of operators are well-defined; we obtain a precise alternative expression of the series \(E(s, w)\) in terms of traces in Section 3.4. We then deduce precise information on \(E(s, w)\) when \(|\Im s|\) is bounded (see Sections 3.5 and 3.6) and, in particular, we make a precise study of its poles in Section 3.5. However, the norms of the operators—on \(A_\infty(V)\)—grow exponentially with respect to \(|\Im s|\); this is why \(A_\infty(V)\) is not well-adapted to the study of \(E(s, w)\) for large \(|\Im s|\). In Section 3.7, we then deal with the space \(C^1(I)\), where we use only approximate relations between the series \(E(s, w)\) and the norms of transfer operators. Estimates à la Dolgopyat—of polynomial growth with respect to \(|\Im s|\)—are valid on the norms of the operators on \(C^1(I)\) and they entail that \(E(s, w)\) is of polynomial growth for \(|\Im s| \to \infty\). Finally, with
Proposition 3.3, we will return from $E(s, w)$ to $P(s, w)$, and obtain a precise knowledge of the initial series $P(s, w)$ for any complex $s$ (and $w$ near 0) in Section 3.8.

3.1. Weighted transfer operator and function spaces. We now introduce our main object, the weighted transfer operators (via the component transfer operators).

The component weighted transfer operator $H_{[h], t, w}$ relative to an inverse branch $h \in \mathcal{H}^+$ involves the map $\mathcal{h}$ defined in (P3),

$$H_{[h], t, w} f(x) := \exp(wc(h)) h^t(z) f \circ h(z).$$

The sum of the weighted component operators $H_{[h], t, w}$ taken over the set $\mathcal{H}$ defines the weighted transfer operator

$$H_{t, w} := \sum_{h \in \mathcal{H}} H_{[h], t, w}.$$  \hfill (3.2)

This is the main dynamical object of this study, as an extension of the density transformer of the system, obtained when $(t, w) = (1, 0)$.

Due to the additivity of the cost, and the multiplicativity of the derivative (or its associated mapping $h$), the $k$th iterate of the operator $H_{t, w}$ has exactly the same expression as the operator itself, with a sum that is now taken over $\mathcal{H}^k$,

$$H_{t, w}^k := \sum_{h \in \mathcal{H}^k} H_{[h], t, w}.$$  \hfill (3.3)

As in [1], the quasi-inverse plays a central role. Here, we will deal with the even quasi-inverse

$$E_{t, w} := \sum_{k \text{ even } \geq 2} H_{t, w}^k = H_{t, w}^2(I - H_{t, w}^2)^{-1}.$$  \hfill (3.4)

In the present study, there are two useful function spaces on which these operators act, which we now describe.

The space $A_{\infty}(V)$. Generating functions deal with the fixed point $x_h$ of each branch $h$. As we will see, traces of operators, whenever they are well-defined, provide a good tool to deal with fixed points. That is why we are led to study the operators when they act on spaces of analytic functions.

We consider the disk $V = \{z \in \mathbb{C} \mid |z - 1| < 5/4\}$ described in Property (P3) of Section 2.2 and we deal with the Banach space

$$A_{\infty}(V) := \{f : \overline{V} \to \mathbb{C} \mid f \text{ continuous, } f|_V \text{ analytic}\}$$  \hfill (3.4)

defined with the sup-norm $\|f\|_V := \sup\{|f(z)| \mid z \in \overline{V}\}$. This is a Banach space. In view of Property (P3), we associate with any $h \in \mathcal{H}^+$ the function $\mathcal{h}$ which is the analytic extension of the function $|h'(x)|$ to the disk $V$. 
Then, by (P3), the component weighted transfer operator $H_{[h],t,w}$ acts on the Banach space $A_\infty(V)$. Moreover, the equality
\begin{equation}
|H_{[h],t,w}[f](z)| = \exp(\nu c(h))|h(z)|^\sigma \exp(\tau \arg h(z))f \circ h(z)
\end{equation}
that holds for $(t, w)$ with $t = \sigma + i\tau$ and $\nu = \Re w$ shows that the norm $\|H_{[h],t,w}\|_V$ is of exponential growth with respect to $\tau := 3t$.

The space $C^1(\mathcal{I})$. That is why $A_\infty(V)$ is unsuitable when $\tau := 3t$ is large, where we expect polynomial growth with respect to $\tau$. However, the exponential term $\exp(\tau \arg h(z))$ disappears when $z$ is real. That is why we also deal with the Banach space $C^1(\mathcal{I})$ of continuously differentiable functions on the unit interval $\mathcal{I}$ with the usual norm $\|f\|_{(1, 1)} := \|f\|_0 + \|f'|_0$ where $\|f\|_0 := \sup\{|f(x)| : x \in \mathcal{I}\}$. In this case, the transfer operator $H_{[h],t,w}$ involves the absolute value of the derivative, satisfies
\begin{equation}
H_{[h],t,w}[f](x) = e^{wc(h)}|h'(x)|^\sigma f \circ h(x),
\end{equation}
and is proven to act on this space. We will adapt the Dolgopyat–Baladi–Vallée results which entail polynomial growth of the norm of $H_{t,w}$ when $|3t|$ becomes large.

3.2. First properties of $H_{t,w}$. We now study the main properties of the operator $H_{t,w}$ when the pair $(t, w)$ belongs to the sets defined in (2.15), and when it acts either on $A_\infty(V)$ or $C^1(\mathcal{I})$. We also study the analyticity of the mapping $(t, w) \mapsto H_{t,w}$ on each space.

Proposition 3.1. Consider a digit-cost $c$ of moderate growth and the associated sets $\Gamma(a)$ (defined in (2.15)). The following holds:

(a) [Case $a > 1/2$] For any $(t, w) \in \Gamma(a)$, the operator $H_{t,w}$ acts on $A_\infty(V)$ and on $C^1(\mathcal{I})$. The map $(t, w) \mapsto H_{t,w}$ is analytic on $\Gamma(a)$.

(b) [Case $a > 1$] For any $(t, w) \in \Gamma(a)$, the quasi-inverse $(I - H_{t,w})^{-1}$ acts on $A_\infty(V)$ and on $C^1(\mathcal{I})$. The map $(t, w) \mapsto (I - H_{t,w})^{-1}$ is analytic on $\Gamma(a)$.

(c) [Useful bounds] The bounds (3.7) and (3.8) hold on $\Gamma(a)$ for any $a > 1/2$.

Proof. By Lemma 2,
\begin{equation}
H_{\sigma,\nu}^n[1](-1/4) = \sum_{h \in \mathcal{H}^n} e^{\nu c(h)}|h'(-1/4)|^\sigma \leq K^{\nu}L^\sigma \sum_{h \in \mathcal{H}^n} |h'(0)|^{\sigma - d|\nu|}.
\end{equation}
For $z \in V$ and $f \in A_\infty(V)$, we have
\begin{equation}
|H_{t,w}^n[f](z)| \leq e^{\tau |\tau|}H_{\sigma,\nu}^n[1](-1/4)\|f\|_V,
\end{equation}
and by (3.7) this ensures that, for any $(t, w) \in \Gamma(a)$ with $a > 1/2$, the operator $H_{t,w}$ acts on $A_\infty(V)$; similar arguments, together with Lasota–Yorke bounds (see [1]), show that $H_{t,w}$ acts on $C^1(\mathcal{I})$. 

Gaussian behavior of quadratic irrationals
Moreover, the bound obtained in (3.7) entails
\begin{equation}
H_{\sigma,\nu}^n[1](-1/4) \leq K|\nu|L^{a+\sigma}\rho^{(a-1)n} \quad \text{for any } (t, w) \in \Gamma(a) \text{ and } n \geq 1.
\end{equation}
Then, by Property (P1), for \((t, w) \in \Gamma(a)\) with \(a > 1\), the quasi-inverse \((I - H_{t,w})^{-1}\) and its variants act on \(A_{\infty}(V)\) and on \(C^1(I)\).

We now explain why we restrict ourselves to costs of moderate growth. Consider a cost \(c\) that is not of moderate growth. In this case, the ratio \(c(m)/\log m\) is not bounded, and, for any \(\nu, \epsilon > 0\), the inequality \(\nu c(m) > (1 + \epsilon) \log m\) holds for a subsequence of integers \(m\); the series \(H_{1,\nu}[1](0)\) is thus divergent for any \(\nu > 0\). Then there is no neighborhood \(S \times W\) of \((1, 0)\) on which \(H_{t,w}\) is well-defined. However, the existence of such a neighborhood is crucial in all our study, notably for the Implicit Function Theorem used in Theorem 1.2(iii), Dolgopyat’s bounds (Theorem A) and Quasi-Powers Theorem (Theorem C).

3.3. Role of each function space. The expression of \(\alpha(h)\) given in (2.4) in terms of derivatives entails an alternative expression for \(Y_k(t, w)\) (introduced in (2.12)), which involves the component operators \(H_{[h],t,w}\) defined in (3.6):
\begin{equation}
Y_k(t, w) = \sum_{h \in \mathcal{H}^k} H_{[h],t,w}[1](x_h).
\end{equation}
Here, each component operator is evaluated at a point that depends on \(h\), namely the fixed point \(x_h\) of \(h\). This is the main difference from previous studies (rational trajectories for instance) where the same “\(x\)” is used for each component operator: there, the analog of \(Y_k(t, w)\) is expressed in terms of the \(k\)th iterate of \(H_{t,w}\).

Far from the real axis. In Section 3.7, we deal with the space \(C^1(I)\), where there are interesting Dolgopyat–Baladi–Vallée bounds on the \(k\)-iterates of the operator for large \(|\Im s|\) (established by Dolgopyat [6], and extended by Baladi and Vallée [1]). We thus need to relate (but not exactly) \(Y_k(s, w)\) and various \(H_{s,w}^\ell[f](x)\) for \(\ell \leq k\), and some pairs \((f, x)\), when the operator acts on the function space \(C^1(I)\).

Near the real axis, or in the intermediate region. Here, we need exact alternative expressions, and we deal with the space \(A_{\infty}(V)\) defined in (3.4). In this space, as we will show in the next section, the trace of each weighted component transfer operator is well-defined and admits an expression in terms of \(\alpha(h)\). Due to Proposition 2.5 that relates \(P(s, w)\) and \(E(s, w)\), we are then led to introduce the even quasi-inverse \(E_{t,w}\) (with \(t = s/2\), already described in (3.3), together with its trace. We then obtain an alternative expression for the series \(E(s, w)\), given in Proposition 3.2. This will be useful to find the main singularities of \(E(s, w)\) near the real
axis, described in Proposition 3.3 of Section 3.5 and study $E(s, w)$ in the intermediary region (see Proposition 3.4).

3.4. Trace of the even quasi-inverse. In [9], Grothendieck proved that the usual formulae that hold for the trace of a matrix are also valid in a precise framework that is described in the Appendix. As proven there, the space $(\mathcal{A}_\infty(\mathcal{V}), \|\|_\mathcal{V})$, defined in (3.4) provides an instance of such a framework.

First (see Proposition 6.4(a)), the trace of each component operator $H_{[h], t, w}$ is well-defined there, and is expressed in terms of $\alpha(h) = |h'(x_h)|^{1/2}$ and depth $|h|$ of the branch $h$,

$$\text{Tr } H_{[h], t, w} = \alpha(h)^{2t} \frac{\exp(wc(h))}{1 - (-1)^{|h|} \alpha(h)^2}.$$  

Then, each component of the series $E(s, w)$, relative to a branch $h$ of even depth, can be written as

$$\alpha(h)^{2t} \exp(wc(h)) = \text{Tr } H_{[h], t, w} - \text{Tr } H_{[h], t+1, w}.$$  

Second, for $(t, w) \in \Gamma(a)$ with $a > 1$, and as proven in Proposition 6.4(c), the trace of the even quasi-inverse $E_{t, w}$ defined in (3.3) is also well-defined and equals the sum of the traces $\text{Tr } H_{[h], t, w}$ taken over all the inverse branches $h \in \mathcal{H}$ with even depth. Then, the Dirichlet series (3.9)

$$E(s, w) := \text{Tr } E(s/2, w)$$  

satisfies (3.10)

$$E(s, w) = E(s, w) - E(s + 2, w).$$  

The next proposition relates the two series $E(s, w)$ and $E(s, w)$.

**Proposition 3.2.** The Dirichlet series $E(s, w)$ and $E(s, w)$ are analytically equivalent.

**Proof.** Proposition 6.4 provides in (6.6) an explicit expression of $E(s, w)$. The maximal value for $h \mapsto \alpha(h)$ over $\mathcal{H}^+$ is attained for $h = h_1$ and equals $\rho^2$ with $\rho$ defined in Property (P2). With $K := 1 - \rho^2$, this entails

$$|E(s, w)| \leq (1/K) Y(\sigma/2, |\nu|)$$  

for $\sigma = \Re s$, $\nu = \Re w$.

**Condition** (a). By (2.16), the series $E(s, w)$ satisfies (a) when $(s/2, w)$ belongs to $\Gamma(a)$ with $a > 1$.

**Condition** (b). The difference $E(s, w) - E(s, w)$ is equal to $E(s + 2, w)$, is a shift to the left of $E(s, w)$. This proves, by (2.16), that $E(s + 2, w)$ is analytic and uniformly bounded as long as $((s + 2)/2, w) \in \Gamma(a)$ with $a > 1$. This means that $(s/2, w) \in \Gamma(a)$ with $a > 1/2$. Thus, condition (b) holds under the same conditions as in Proposition 2.5.
3.5. Study of $E(s,w)$ near the real axis. With Proposition $3.2$ at hand, we are led to study the trace $E(s,w)$ of the even quasi-inverse. Proposition $6.4$ proves that the map $(t,w) \mapsto \text{Tr}E_{t,w}$ is analytic on $\Gamma(a)$ with $a > 1$. We now study an analytic continuation of this map on a neighborhood of $(1,0)$. This will be done in the Appendix, in Sections $6.4$–$6.6$. Section $6.4$ describes the spectral decomposition of the operator $H_{t,w}$ near the real axis, and relates it to the spectral decomposition of the even quasi-inverse $E_{t,w}$. We recall here the main steps:

For real pairs $(t,w) \in \Gamma(a)$ (with $a > 1/2$), the operator $H_{t,w}$ is compact and has strong positivity properties that entail the existence of dominant spectral objects, in the same vein as the Perron–Frobenius properties. For $(t,w) \in \Gamma(a)$ with $a > 1/2$, the dominant spectral objects defined for real pairs may be extended by analytic perturbation of the dominant part of the spectrum when the pair $(t,w)$ is close to a real pair. More precisely, the Appendix introduces in (6.8) the spectral decomposition $H_{t,w} = \lambda(t,w)P_{t,w} + N_{t,w}$, that holds when $(t,w)$ is in a neighborhood of $(1,0)$, where $\lambda(t,w)$ is the dominant eigenvalue, and the spectral radius of $N_{t,w}$ is less than $1$. This decomposition gives rise to the spectral decomposition, $E_{t,w} = H_{t,w}^2(I - H_{t,w}^2)^{-1} = \frac{\lambda(t,w)^2}{1 - \lambda(t,w)^2}P_{t,w} + N_{t,w}^2(I - N_{t,w}^2)^{-1}$.

Now, Section $6.5$ studies the trace of this quasi-inverse. It explains how the previous decomposition gives rise to the spectral decomposition $\text{Tr}E_{t,w} = \frac{\lambda(t,w)^2}{1 - \lambda(t,w)^2} + \text{Tr}N_{t,w}^2(I - N_{t,w}^2)^{-1}$, and leads to the analytic continuation of $(t,w) \mapsto E_{t,w}$. Moreover, Section $6.6$ describes the constants that are involved in this decomposition.

Using Proposition $3.2$, we now transfer these results to the generating function $E(s,w)$ when $(s,w)$ is close to the point $(2,0)$.

**Proposition 3.3.** There is a domain $\mathcal{T}_2 := S_2 \times \mathcal{W}_2$ formed by a neighborhood $\mathcal{W}_2$ of $0$ and a rectangle $S_2 := \{ s = \sigma + i\tau \mid |\Re s - 2| < \delta_2, |\tau| < \tau_2 \}$ where the following hold:

(a) For each $w \in \mathcal{W}_2$, the mapping $s \mapsto E(s,w)$ is meromorphic on $S_2$.
(b) This mapping has a unique (simple) pole on $S_2$, located at $s_w := 2\sigma(w)$ where $\sigma(w)$ is defined by the equation $\lambda(\sigma(w),w) = 1, \sigma(0) = 1$.
(c) The residue $v(w)$ of $s \mapsto E(s,w)$ at $s = s_w$ equals $-1/\lambda'(\sigma(w),w)$. The product $(s - s_w)E(s,w)$ is bounded on $\overline{\mathcal{T}_2}$.

**Proof.** Sections $6.5$ and $6.6$ describe the analogs of properties (a)–(c) for $\text{Tr}E_{t,w}$ and introduce the domains $S_1$ and $\mathcal{W}_1$ and their product $\mathcal{T}_1 :=$
\(S_1 \times \mathcal{W}_1\). We now return from \(t = s/2\) to \(s\) and the change \(s/2 \mapsto s\) brings the factor 2 in the residue. It transforms the domain \(S_1\) into its homothetic image \(S_2 := 2S_1\) and we let \(\mathcal{T}_2 := S_2 \times \mathcal{W}_1\). We now use the property of Section 6.6 which entails that the product \((s-s_w)\text{Tr}\mathbf{E}_{s/2,w} = (s-s_w)E(s,w)\) is bounded on \(\overline{\mathcal{T}}_2\). Thus, properties (a)–(c) are proven for \(E(s,w)\). Finally, by means of Proposition 3.2, we return to \(E(s,w)\) and conclude the proof. \(\blacksquare\)

3.6. In the middle. This section studies the behavior of \(E(s,w)\) when \(s\) is in the intermediary region, and belongs to the union of two rectangles \(\{s \mid |\Re s - 2| < \delta, \tau_2 \leq |\tau| \leq \tau_4\}\). The bound \(\tau_2\) comes from the domain \(\mathcal{T}_2\) of Proposition 3.3 and the bound \(\tau_4\) will be provided later on by the domain \(\mathcal{T}_4\) of Theorem 3.5 in the next section. The behavior of the mapping \((t,w) \mapsto \text{Tr}\mathbf{E}_{t,w}\) in the intermediary region is related to the behavior of the even quasi-inverse \(\mathbf{E}_{t,w}\) there, which is itself based on the following spectral property of the plain operator \(\mathbf{H}_t := \mathbf{H}_{t,0}\) when \(t\) is on the vertical line \(\Re s = 1\) (see, for instance, Faivre [7] or Vallée [27]).

**Proposition A.** On the punctured vertical line \(\{\Re t = 1, t \neq 0\}\), the spectral radius \(R(t)\) of the operator \(\mathbf{H}_t\) (when acting on \(\mathcal{A}_\infty(\mathcal{V})\)) is strictly less than 1.

Using analytic perturbation of parameters \((t,w)\), the following result proves that \(E(s,w)\), and thus \(E(s,w)\), remains bounded in the intermediary region.

**Proposition 3.4.** For any pair \((\tau_2, \tau_4)\) with \(0 < \tau_2 < \tau_4\), there exists a domain \(\mathcal{T}_3 := S_3 \times \mathcal{W}_3\) formed by a neighborhood \(\mathcal{W}_3\) of \(w = 0\) and a union \(S_3\) of two rectangles, \(S_3 := \{s \mid |\Re s - 2| \leq \delta_3, \tau_2 \leq |\Im s| \leq \tau_4\}\), and there exists \(M_3 > 0\), for which the series \(E(s,w)\) satisfies \(|E(s,w)| \leq M_3\) on the domain \(\overline{\mathcal{T}}_3\).

**Proof.** We first study the function \((t,w) \mapsto \text{Tr}\mathbf{E}_{t,w}\). Proposition A together with perturbation theory of finite parts of the spectrum implies that, for any two fixed positive numbers, say \(0 < \tau_2 < \tau_4\), there is a domain \(\mathcal{T}_3 := \mathcal{R}_3 \times \mathcal{W}_3\), built with a neighborhood \(\mathcal{W}_3\) of \(w = 0\) and a union \(\mathcal{R}_3\) of two rectangles,

\[\mathcal{R}_3 := \{t \mid |\Re t - 1| \leq \delta_3/2, \tau_2/2 \leq |\tau| \leq \tau_4/2\},\]

such that the spectral radius \(R(t,w)\) of the operator \(\mathbf{H}_{t,w}\) is strictly less than 1 on \(\overline{\mathcal{T}}_3\).

We may suppose that \(\overline{\mathcal{T}}_3 \subset \Gamma(a)\) for \(a > 3/4\). Then, on \(\overline{\mathcal{T}}_3\), Proposition 6.4(c) proves that \(\text{Tr}\mathbf{E}_{t,w}\) is given by the absolutely convergent series of general term \(\text{Tr}\mathbf{H}^k_{t,w}\). Now, in Proposition 6.4(b), it is proved that each map \((t,w) \mapsto \text{Tr}\mathbf{H}^k_{t,w}\) is analytic on \(\Gamma(a)\) with \(a > 3/4\). Then, on the open domain \(\mathcal{T}_3\), the map \((t,w) \mapsto \text{Tr}\mathbf{E}_{t,w}\) is analytic, and it is bounded on \(\overline{\mathcal{T}}_3\).
If $S_3 := 2\mathcal{R}_3$ then $E(s, w)$ is bounded on $\mathcal{T}_3 := S_3 \times \mathcal{W}_3$ and analytic on its interior. This neighborhood $\mathcal{T}_3$ can be chosen so that $\mathcal{T}_3 \subset S_0 \times \mathcal{W}_0$ from Proposition 3.2, and then $E(s, w)$ is also bounded and analytic on $\mathcal{T}_3$. ■

3.7. Far from the real axis. We now study the series $E(s, w)$ when $w$ is close to 0, and $s$ is in a vertical strip close to the vertical line $\Re s = 2$ but now with large values of $|\Im s|$. We wish to exhibit polynomial growth of the mapping $s \mapsto E(s, w)$ for large $|\Im s|$. We have already explained in Section 3.1 why the space $A_\infty(V)$ is not well-adapted to this task. We thus do not use the decomposition (3.10) and directly deal with $E(s, w)$. We will then compare directly the series $E(s, w)$ and iterates $H_{t,w}^k$ of the transfer operator when acting on $C^1(I)$, for $t = s/2$ with a real part close to 1.

The behavior of the iterates $H_t^k$ of the (unweighted) transfer operator $H_t$ in a vertical strip near $\Re t = 1$ was studied by Dolgopyat for a Markovian dynamical system with a finite number of branches. When such a system satisfies the UNI Property (recalled as Property (P5) in Section 2.2), Dolgopyat gives bounds on the iterates $H_t^k$ which prove that the (plain) quasi-inverse $(I - H_t)^{-1}$ is analytic with polynomial growth. In [1], Baladi and Vallée extended his work to dynamical systems with an infinite number of branches, and to weighted operators (associated with a cost of moderate growth). Dolgopyat dealt with the Banach space $C^1(I)$, introduced a family of norms $\| \cdot \|_{(1,\tau)}$ on $C^1(I)$ indexed by the real parameter $\tau \neq 0$, and obtained estimates on $\|H_{t,w}^k\|_{(1,\tau)}$ for $\tau = \Im t$. The next theorem describes the Dolgopyat–Baladi–Vallée estimates, as stated in [1].

**Theorem A** (Dolgopyat–Baladi–Vallée estimates). There exist a domain $\mathcal{T}_5 := S_5 \times \mathcal{W}_5$ formed by a neighborhood $\mathcal{W}_5$ of $w = 0$ and an unbounded rectangle $S_5 := \{s = \sigma + i\tau \mid |\sigma - 1| \leq \delta_5, |\tau| > \tau_5\}$, an exponent $\xi_5 > 0$, a contraction constant $\gamma_5 < 1$, and a bound $M_5$, such that, for any $(s, w) \in \mathcal{T}_5$, the norm $(1, \tau)$ of the $k$th iterate of the operator $H_{t,w}^k$ satisfies, for $\tau = \Im t$,

$$\|H_{t,w}^k\|_{(1,\tau)} \leq M_5 |\tau|^{|\xi_5 \gamma_5|^k} \quad \text{for } k \geq 1. \tag{3.11}$$

For any $w \in \mathcal{W}_5$, the quasi-inverse $t \mapsto (I - H_{t,w})^{-1}$ is analytic on $S_5$ and has polynomial growth for $|\Im t| \to \infty$.

We now study similar bounds for the components $Y_k(t, w)$ of the series $E(s, w)$ (with $t = s/2$),

$$Y_k(t, w) = \sum_{h \in \mathcal{H}^k} \exp(wc(h))|h'(x_h)|^t.$$

We precisely relate $Y_k(t, w)$ to the transfer operator $H_{t,w}^k$, and prove, in the
next Theorem 3.5 that $E(s, w)$ is analytic and of polynomial growth in a vertical strip $|\Re s - 2| \leq \delta_0$ and $|\Im s| \geq \tau_0$ for some small enough $\delta_0 > 0$ and large enough $\tau_0 > 0$.

**Theorem 3.5.** There exist a domain $\mathcal{T}_4 := S_4 \times \mathcal{W}_4$ formed by a neighborhood $\mathcal{W}_4$ of $w = 0$ and an unbounded rectangle $S_4 := \{s = \sigma + i\tau \mid |\sigma - 2| \leq \delta_4, |	au| > \tau_4 \}$ with $\tau_4 \geq 1$, an exponent $\xi_4 > 0$, a contraction constant $\gamma_4 < 1$, and a bound $M_4$, such that, for any $w \in \mathcal{W}_4$, $s \mapsto E(s, w)$ is analytic on $S_4$, and

$$|E(s, w)| \leq M_4|\tau|^{\xi_4} \quad \text{for any } (s, w) \in \mathcal{T}_4.$$

**Proof.** The proof is an extension of the proof given by Pollicott and Sharp [22]. The authors of [22] use an idea of Ruelle and relate the zeta series to the transfer operator, with two restrictions: they only consider the unweighted case where the dynamical system has a finite number of branches. We here show that their proof extends to a weighted transfer operator with an infinite denumerable set of branches and then relate the weighted zeta series to the weighted transfer operator.

We consider the Dolgopyat domain $\mathcal{T}_5 = S_5 \times \mathcal{W}_5$ of Theorem A together with a pair $(s, w)$ with $w \in \mathcal{W}_5$ and $s \in 2S_5$. If $s \in 2S_5$, then $t = s/2 \in S_5$, and we may apply the Dolgopyat bounds to the iterates $H^k_{t, w}$.

**General strategy.** Due to the definition of $E(s, w)$ given in (2.14), we deal with $Y_k(t, w)$ with $t = s/2$. We let $\sigma := \Re t = \frac{1}{2}\Re s$, $\tau := \frac{1}{2}\Im s$ and $\nu := \Re w$.

We associate with each pair $(h, \ell)$ with $|h| \leq \ell$ the function

$$F_h^{(\ell)} := H_{t, w}[1_{I_h}] = \sum_{g \in H^{(\ell)}} H_{[g], t, w}[1_{I_h}].$$  

(3.12)

Denote by $m$ the depth $|h|$ of $h$. As $H_{[g], t, w}$ is a component operator which deals with the LFT $g$, the function $H_{[g], t, w}[1_{I_h}]$ involves the function $1_{I_h} \circ g$ and there are two cases according as $g$ begins with $h$ or not:

(i) if there exists $u$ for which $g = h \circ u$, then $H_{[g], t, w}[1_{I_h}] = H_{[h \circ u], t, w}[1]$;

(ii) if $g$ cannot be written as $g = h \circ u$ with $u \in H^{(\ell-m)}$, then $H_{[g], t, w}[1_{I_h}] = 0$.

Finally, in both cases, this leads to another expression for any function $F_h^{(\ell)}$ defined in (3.12) relative to $h \in H^m$:

$$F_h^{(\ell)} := H_{t, w}[1_{I_h}] = \sum_{u \in H^{(\ell-m)}} H_{[h \circ u], t, w}[1] = H_{t, w}^{(\ell-m)}[F_h^{(m)}].$$  

(3.13)

This entails, with the distortion property and Lasota–Yorke bounds (see [1]), the useful inequality

$$\sum_{h \in H^m} \|F_h^{(\ell)}\|_{(1,1)} \leq H_{\sigma, \nu}^m[1](-1/4).$$  

(3.14)
Our object of interest $Y_k(t, w)$ is written in terms of these $F_h^{(k)}$,

$$Y_k(t, w) = \sum_{h \in \mathcal{H}} F_h^{(k)}(x_h).$$

The main idea (which extends an idea due to Ruelle) is to write $Y_k(t, w)$ as a sum

$$Y_k(t, w) = \sum_{h \in \mathcal{H}} F_h^{(k)}(x_h) + \sum_{m=2}^{k} \Delta_m, \quad \Delta_m := \sum_{h \in \mathcal{H}^m} F_h^{(k)}(x_h) - \sum_{h \in \mathcal{H}^{m-1}} F_h^{(k)}(x_h).$$

We first study the first term $T_k(t, w) := \sum_{h \in \mathcal{H}} F_h^{(k)}(x_h)$, and then each difference $\Delta_m$ contained in the second term.

**First term.** In this case, $F_h^{(k)} = H_{t, w}^{k-1}[F_h^{(1)}]$. Then

$$|T_k(t, w)| \leq \sum_{h \in \mathcal{H}} \|F_h^{(k)}\|_0 \leq \|H_{t, w}^{k-1}\|_0 \sum_{h \in \mathcal{H}} \|F_h^{(1)}\|_0.$$

Using the relation between the 0-norm and the Dolgopyat norm, together with the Dolgopyat bound (3.11) and inequality (3.14), we obtain

$$|T_k(t, w)| \leq \sum_{h \in \mathcal{H}} \|F_h^{(k)}\|_0 \ll |\tau|^5 \gamma^k \ll |\tau|^5 \gamma^k.$$

**Second term.** We first prove that

(3.16) $F_{b(h)}^{(k)} = \sum_{h \mid h = b(h) \circ g} F_h^{(k)}$ for any $b(h) \in \mathcal{H}^{m-1}$.

The fundamental interval $I_{b(h)}$ relative to the LFT $b(h)$ is the disjoint union of the fundamental intervals $I_{b(h) \circ g}$ for $g \in \mathcal{H}$ and, using (3.13) twice, we get

$$F_{b(h)}^{(k)} = H_{t, w}^k[I_{b(h)}] = \sum_{g \in \mathcal{H}} H_{t, w}^k[I_{b(h) \circ g}] = \sum_{g \in \mathcal{H}} \sum_{u \in \mathcal{H}^{k-m}} H_{[b(h) \circ g u], t, w}[1]$$

$$= \sum_{u \in \mathcal{H}^{k-m}} \sum_{g \in \mathcal{H}} H_{[b(h) \circ g u], t, w}[1]$$

$$= \sum_{h \mid h = b(h) \circ g} H_{[h \circ u], t, w}[1] = \sum_{h \mid h = b(h) \circ g} H_{t, w}^k[I_{b(h)}] = \sum_{h \mid h = b(h) \circ g} F_h^{(k)},$$

proving (3.16). Now, as any $h \in \mathcal{H}^{m-1}$ is the beginning $b(h)$ of some $h \in \mathcal{H}^m$, it easily entails the equality

$$\Delta_m = \sum_{h \in \mathcal{H}^m} F_h^{(k)}(x_h) - \sum_{h \in \mathcal{H}^{m-1}} F_h^{(k)}(x_h)$$

$$= \sum_{h \in \mathcal{H}^m} [F_h^{(k)}(x_h) - F_h^{(k)}(xb(h))].$$
As the two points \(x_h\) and \(x_{b(h)}\) belong to the same fundamental interval \(I_{b(h)}\) of depth \(m - 1\), whose length is less than \(\rho^{m-1}\), we have

\[
|\Delta_m| \leq \rho^{m-1} \left[ \sum_{h \in \mathcal{H}^m} \|F_h^{(k)}\|_1 \right].
\]

Using the expression of \(F_h^{(k)}\) in (3.13), we are then led to evaluate the \((1, 1)\) norm of each function

\[
F_h^{(k)} = H_{t,w}^{k-m} \circ H_{[h],t,w}[1] \quad \text{for } h \in \mathcal{H}^m.
\]

Using the relation (7) between the \((1, 1)\) norm and the \((1, \tau)\) norm, together with the Dolgopyat bound (3.11) and inequality (3.14), we get the following bound, for any \(\gamma_4 \geq \gamma_5\):

\[
|\Delta_m| \ll |\tau| \cdot \|H_{t,w}^{k-m}\|_{(1,\tau)}H_{\sigma,\nu}[1](-1/4)\rho^m
\ll |\tau|^{1+\xi_5} \gamma_4^{k-m}[\rho^mH_{\sigma,\nu}[1](-1/4)]
\ll |\tau|^{1+\xi_5} \gamma_4^{k}[\rho^mH_{\sigma,\nu}[1](-1/4)\gamma_4^{-m}].
\]

The bound (3.8) leads to the inequality

\[
\rho^mH_{\sigma,\nu}[1](-1/4) \ll \rho^m\rho^{m(\sigma-1-d|\nu|)} = \rho^{m(\sigma-d|\nu|)}.
\]

The exponent satisfies \(\sigma - d|\nu| \geq 1 - \delta_5 - d\nu_5\). Then, choosing \(\gamma_4 < 1\) and \(\gamma_4 \geq \gamma_5\) entails the final bound

\[
\sum_{m=2}^k |\Delta_m| \ll k|\tau|^{\xi_4} \gamma_4^k \quad \text{with } \xi_4 = 1 + \xi_5.
\]

Now, we use the bound \(|Y_k(t, w)| \leq |T_k(t, w)| + \sum_{m=2}^k |\Delta_m|\), we change \(\tau\) to \(2\tau\), and find that the series \(Y_k(t, w)\) itself satisfies, for some constant \(\tilde{M}_4\),

\[
|Y_k(t, w)| \ll \tilde{M}_4 k|\tau|^{\xi_4} \gamma_4^k.
\]

This yields the final bound with \(M_4 := \frac{\tilde{M}_4}{\gamma_4^4} \left( \frac{1}{1-\gamma_4} \right)^2\).

### 3.8. Conclusion: main properties of \(P(s, w)\)

We now adjust various neighborhoods and bounds from Propositions 2.5, 3.3, 3.4 and Theorem 3.5.

Each result provides a domain \(T_i := S_i \times W_i\): The index \(i\) equals 0 for Proposition 2.5, \(i = 2\) for Proposition 3.3, \(i = 3\) for Proposition 3.4, and \(i = 4\) for Theorem 3.5. Each \(W_i\) is a neighborhood of 0, and we denote by \(W'\) the intersection of the four neighborhoods. Each \(S_i\) is a vertical strip or a part of a vertical strip of the form \(\{s \mid |\Re s - 2| < \delta_i\}\). We thus let \(\delta := \min(\delta_0, \delta_2, \delta_3, \delta_4)\) and consider the strip \(S := \{s \mid |\Re s - 2| < \delta\}\).

Then the horizontal lines \(|\tau| = \tau_2\) from Proposition 3.3 and \(|\tau| = \tau_4\) from

(7) In [22], the authors seem to have forgotten the factor \(|\tau|\)....
Theorem 3.5 separate the strip $S$ into three regions, where the three results of Proposition 2.5, Proposition 3.4 and Theorem 3.5 hold when $w \in \mathcal{W}'$.

We need a last adjustment of $\mathcal{W}'$ to better deal with the rectangle “near the real axis”, namely $\{s \mid |\Re s - 2| < \delta, |\tau| \leq \tau_2\}$; we want indeed the set of poles $\{s_w \mid w \in \mathcal{W}'\}$ to be not too close to the left vertical line $\Re s = 2 - \delta$. As the mapping $\sigma$ is analytic on $\mathcal{W}'$, the derivative $|\sigma'(w)|$ is bounded on $\mathcal{W}'$ (say by some $B > 0$). Then, with the new (and final) neighborhood $\mathcal{W}$ defined as

$$\mathcal{W} := \mathcal{W}' \cap \left\{ |w| \leq \frac{\delta}{4B} \right\},$$

the following holds for $w \in \mathcal{W}$:

$$|\Re s_w - 2| = 2|\Re \sigma(w) - 1| \leq 2|w|B \leq \frac{1}{2}\delta, \quad |\Re s_w - (2 - \delta)| \geq \frac{1}{2}\delta.$$  

We let $\tau_0 := \tau_4 > 1$ and $\xi := \xi_4$. We then obtain (after adjustment of the various constants $M_i$) the final result of this section, which is clearer when using the following notations.

**Notations.** A width $\delta > 0$ and two horizontal lines $|\tau| = \tau_0$ (with $\tau_0 > 1$) divide the strip $\{s \mid |\Re s - 2| < \delta\}$ into three domains: a bounded rectangle $\mathcal{R} := \{s \mid |\Re s - 2| < \delta, |\tau| \leq \tau_0\}$ and the union $\mathcal{U} := \{s \mid |\Re s - 2| < \delta, |\tau| \geq \tau_0\}$ of two unbounded rectangles.

**Theorem 3.6.** Consider the bivariate generating series $P(s, w)$ associated with a digit-cost $c$ of moderate growth and defined in (2.9). There exists a neighborhood $\mathcal{W}$ of $w = 0$, a width $\delta > 0$, two horizontal lines $|\tau| = \tau_0$ (with $\tau_0 > 1$), an exponent $\xi > 0$ and a bound $M$, for which the following hold:

(a) For $w \in \mathcal{W}$, the mapping $s \mapsto P(s, w)$ is meromorphic on the half-plane $\{s \mid \Re s > 2 - \delta\}$.

(b) The mapping has a unique (simple) pole, located inside the rectangle $\mathcal{R}$, at the point $s_w = 2\sigma(w)$ defined by $\lambda(\sigma(w), w) = 1$, $\sigma(0) = 1$. The residue $v(w)$ of $s \mapsto P(s, w)$ at $s = s_w$ equals $-1/\lambda_w'(\sigma(w), w)$. The distance between $s_w$ and the vertical line $\Re s = 2 - \delta$ is at least $\frac{1}{2}\delta$. The function $v(s, w) := (1/s)(s - s_w)P(s, w)$ satisfies $|v(s, w)| \leq M$ on $\overline{\mathcal{R}} \times \overline{\mathcal{W}}$.

(c) $|P(s, w)| \leq M|3s|^\xi$ on $\overline{\mathcal{U}} \times \overline{\mathcal{W}}$.

(d) $|P(s, w)| \leq M$ on $\{s \mid \Re s \geq 2 + \delta\} \times \overline{\mathcal{W}}$.

**4. Extraction of coefficients and the Gaussian law.** The previous section provides a precise description of the analytic properties of the Dirichlet bivariate generating function $P(s, w)$. We now return to our probabilistic setting. Section 4.1 recalls the role played by the moment generating function $M_N(w)$ of the cost $C$ on the subset $\mathcal{P}_N$. Then, Section 4.2 gives asymptotic
estimates of this moment generating function that are obtained by “extracting” the coefficients of the generating function \(P(s, w)\): applying a uniform version of the Landau Theorem leads to uniform estimates when \(w\) belongs to the neighborhood \(W\) defined in Theorem 3.6. With this uniform estimate of \(M_N(w)\) at hand, Section 4.3 applies the Quasi-Powers Theorem, which “transfers” this (uniform) asymptotic estimate into an asymptotic Gaussian law, with the speed of convergence stated in Theorem 1.1. The precise study of the constants of interest, performed in Section 4.4, gives rise to Theorem 1.2.

4.1. Using the moment generating function. In this paper, we study (when \(N \to \infty\)) the asymptotic distribution of the cost \(C\) restricted to \(\mathcal{P}_N\) via the sequence of moment generating functions \(M_N(w) := \mathbb{E}_N[\exp(wC)]\) where \(w\) is a complex number close to 0. From its definition, the moment generating function \(M_N(w)\) is written as a quotient,

\[
M_N(w) := \mathbb{E}_N[\exp(wC)] = \frac{S_w^{[C]}(N)}{S_0^{[C]}(N)},
\]

and involves the cumulative sum \(S_w^{[C]}(N)\) of the cost \(\exp(wC)\) over \(\mathcal{P}_N\),

\[
S_w^{[C]}(N) := \sum_{x|e(x) \leq N} \exp(wC(x)) \quad \text{with} \quad S_0^{[C]}(N) = |\mathcal{P}_N|.
\]

4.2. Landau Theorem. The Landau Theorem [15] was proved by Landau in 1924. This is a strong tool which gives estimates on the sum of coefficients of a Dirichlet series, provided the series satisfies (not too) strong analytical properties. This theorem is not so well-known, and many works (like [1] or [22]) that use it, do not deal with the strong version (existing in the original work of Landau [15], but perhaps somewhat hidden). These authors begin with a weak version and prove the strong version by hand (8). What we mean by weak or strong is related to the exponent \(\xi\): the weak version only deals with \(\xi < 1\), and very often (as is the case here) one would need a strong version with \(\xi\) that may be larger than 1. A proof of the strong version is precisely given in Roux’s thesis [23] and also available in [2]. We here need a “uniform” version (with respect to the parameter \(w\)) of this strong version:

**Theorem B** (Landau Theorem: uniform version). If there is a complex neighborhood \(W\) of 0 where the series \(P(s, w)\) satisfies the “uniform” conclusions of Theorem 3.6 with an exponent \(\xi > 0\), and a width \(\delta\), the following

\[\text{(8)}\] In [1], the authors use the weak version of the Landau Theorem, and they have to introduce the so-called smoothed probabilistic model. With the strong version at hand, part of their work may be changed and shortened.
holds for any \( N \geq 1 \):

\[
S_{w}^{[c]}(N) = \frac{v(w)}{2\sigma(w)} N^{2\sigma(w)}[1 + O(N^{-\beta})], \quad \beta := \frac{\delta}{2([\xi] + 3)},
\]

where the constant in the \( O \)-term is uniform for \( w \in \mathcal{W} \).

This leads to a first result, easily deduced from Theorems 3.6 and B:

**Proposition 4.1.** Consider the set \( \mathcal{P} \) of rqi numbers, and a cost \( C \) on \( \mathcal{P} \) associated with a digit-cost \( c \) of moderate growth. Denote by \( C_N \) the restriction of \( C \) to the set \( \mathcal{P}_N \) of rqi numbers \( x \) with \( \epsilon(x) \leq N \). Then there exists a neighborhood \( \mathcal{W} \) of 0 on which the moment generating function \( M_N(w) \) admits for any \( N \geq 1 \) a “uniform quasi-powers” form

\[
M_N(w) = \frac{v(w)}{\sigma(w)v(0)} N^{2(\sigma(w)-1)}[1 + O(N^{-\beta})],
\]

which involves the exponent \( \beta \) defined in Theorem B and two analytic functions: the function \( w \mapsto \sigma(w) \) defined implicitly by \( \lambda(\sigma(w), w) = 1 \) with \( \sigma(0) = 1 \), and the residue function \( w \mapsto v(w) \), defined in Theorem 3.6. Moreover, the constant in the \( O \)-term is uniform for \( w \in \mathcal{W} \).

4.3. Quasi-Powers Theorem. The following result, known as the Quasi-Powers Theorem, due to Hwang [12], shows that a “uniform quasi-powers” expression for the moment generating functions \( M_N(w) := E_N[\exp(wC_N)] \) entails an asymptotic Gaussian law for the cost \( C_N \).

**Theorem C (Quasi-Powers Theorem).** Consider a sequence \( C_N \) of variables defined on probability spaces \( (\mathcal{P}_N, \mathbb{P}_N) \), and their moment generating functions \( M_N(w) := \mathbb{E}_N[\exp(wC_N)] \). Suppose that the functions \( M_N(w) \) are analytic on a complex neighborhood \( \mathcal{W} \) of zero, and each satisfies there

\[
M_N(w) = \exp[\beta_N U(w) + V(w)](1 + O(\kappa_N^{-1})),
\]

where the \( O \)-term is uniform on \( \mathcal{W} \). Moreover, suppose the two sequences \( \beta_N \) and \( \kappa_N \) tend to \( \infty \) as \( N \to \infty \), and \( U(w), V(w) \) are analytic on \( \mathcal{W} \). Then:

(i) The mean and the variance satisfy

\[
\mathbb{E}_N[C_N] = \beta_N U'(0) + V'(0) + O(\kappa_N^{-1}), \quad \mathbb{V}_N[C_N] = \beta_N U''(0) + V''(0) + O(\kappa_N^{-1}).
\]

(ii) If \( U''(0) \neq 0 \), then distribution of \( C_N \) on \( \mathcal{P}_N \) is asymptotically Gaussian, with speed of convergence \( O(\alpha_N) \) with \( \alpha_N = (\kappa_N^{-1} + \beta_N^{-1}/2) \),

\[
\mathbb{P}_N\left[ x \left| \frac{C_N(x) - U'(0) \log N}{\sqrt{U''(0) \log N}} \leq t \right. \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2} \, du + O(\alpha_N).
\]
Proposition 4.1 shows that Theorem C may be applied to our setting with 
\[ \beta_N = \log N, \quad \kappa_N = N^{\beta}, \] 
and thus 
\[ \alpha_N = \(\log N\)^{-1/2}, \]
together with the functions \(U, V\) defined as 
\[ U(w) = 2(\sigma(w) - 1) \quad \text{and} \quad V(w) = \log \left( \frac{v(w)}{\sigma(w) \cdot v(0)} \right), \]
that involve the two mappings, \(w \mapsto \sigma(w)\) and \(w \mapsto v(w)\) defined in Proposition 4.1. The next section brings a deeper study of these functions near \(w = 0\), and proves that \(U, V\) will be analytic on \(W\) with \(U''(0) \neq 0\). This finishes the proof of our main Theorem 1.1.

4.4. Constants. The function \(U\) exactly coincides with its analog which intervenes in the Gaussian law of the cost \(C\) on rational trajectories and was studied in [1]. One has 
\[ U(0) = 0, \quad U'(0) = 2\sigma'(0) = -2 \frac{\lambda'_t(1,0)}{\lambda'_{t}(1,0)}, \]
and Baladi and Vallée prove \(U''(0) \neq 0\).

With the expression of \(v(w)\) given in Proposition 4.1 one obtains 
\[ V(w) = \log \left( \frac{\lambda'_t(1,0)}{\lambda'_{t}(\sigma(w), w)} \right) - \log \sigma(w). \]
The first two derivatives \(V'(0), V''(0)\) then involve the first three derivatives of the mapping \((t, w) \mapsto \lambda(t, w)\) at \((t, w) = (1, 0)\).

The present function \(V\) does not coincide with its analog \(V_r\) that is associated with rational trajectories and described in [1]. The function \(V_r\) is expressed in terms of the analog of the residual function \(v_r\) as 
\[ V_r(w) = \log \left( \frac{v_r(w)}{\sigma(w) \cdot v_r(0)} \right), \]
\[ v_r(w) := \frac{-1}{\lambda'_t(\sigma(w), w)} \frac{1}{\sigma(w)} F_{\sigma(w), w} \circ P_{\sigma(w), w}[1](0). \]
This expression involves the projector \(P_{t,w}\) on the dominant eigensubspace together with the “final” transfer operator \(F_{t,w}\) which only deals with digits \(m \geq 2\), and is defined as 
\[ F_{t,w}[f](x) := H_{t,w}[f](x) - e^{wc(1)} \frac{1}{(1 + x)^{2t}} f \left( \frac{1}{1 + x} \right). \]
The equality \(F_{t,w} \circ P_{t,w}|1](0) = f_{t,w}(0) - e^{wc(1)} f_{t,w}(1)\) holds and involves the dominant eigenfunction \(f_{t,w}\). Finally, 
\[ v_r(w) = \frac{-1}{\sigma(w) \lambda'_t(\sigma(w), w)} (f_{\sigma(w), w}(0) - e^{wc(1)} f_{\sigma(w), w}(1)). \]
Theorems 1.1 and 1.2 are now proven, and this ends the main part of our study.

5. Possible extensions and open problems.. This section describes possible (easy) extensions of the present work. First, Section 5.1 studies the probabilistic behavior of the “cost” \( \log \epsilon(x) \). Sections 5.2 and 5.3 are devoted to the “constrained” cases, where the numbers have bounded digits (in their continued fraction expansion). Finally, Sections 5.4 and 5.5 discuss possible local limit theorems, and the speed of convergence towards the local limit law.

5.1. The cost \( C(x) := \log \epsilon(x) \). Even though the cost \( x \mapsto \log \epsilon(x) \) is not additive, it can be easily studied via our methods. The bivariate generating functions (primitive or general) involve the general term

\[
e^{w \log \epsilon(x)} \epsilon(x)^{-s} = \epsilon(x)^{-(s-w)}
\]

and exactly coincide with the (univariate) generating functions \( P(s-w), Z(s-w) \). The pole function \( \sigma \) is defined as \( \sigma(w) := 1+w \) and the residue function \( v \) associates with \( w \) the real \(-1/\lambda'(1+w)\). Then \( U'(w) = 2 \) for any \( w \), and the second derivative \( U''(0) \) equals 0. Thus, assertion (i) of Theorem C applies and shows that the variance is very small, of order \( O(1) \), whereas the mean is of order \( \log N \). This means that the cost \( C := \log \epsilon \) is very concentrated around its mean. Assertion (ii) of Theorem C does not apply, and there does not exist a Gaussian limit law. Always with assertion (i) of Theorem C we obtain precise estimates of the expectation and the variance of the cost \( x \mapsto \log \epsilon(x) \) on the set \( \mathcal{P}_N \) of rqi numbers with \( \log \epsilon(x) \leq N \).

These estimates involve the first derivatives of the dominant eigenvalue \( t \mapsto \lambda(t) \) at \( t = 1 \),

\[
\mathbb{E}_N[C] = 2 \log N - \frac{\lambda''(1)}{\lambda'(1)} + O(N^{-\beta}),
\]

\[
\mathbb{V}_N[C] = \frac{\lambda'''(1)\lambda'(1) - \lambda''(1)^2}{\lambda'(1)^2} + O(N^{-\beta}).
\]

The dominant term of the mean value is well-known, but the other terms of the estimates appear to be new, notably the important fact that the variance is of order \( O(1) \).

5.2. Quadratic irrational numbers with bounded digits (I). Here, a given bound \( M \geq 2 \) is fixed and possibly equal to \( \infty \). This section is devoted to studying the “constrained” set \( \mathcal{P}[M] \) of rqi numbers whose continued fraction expansion only contains digits \( \leq M \). The analogous set \( \mathcal{R}[M] \) of reals whose continued fraction expansion only contains digits \( \leq M \) has been extensively studied, and there are many works that deal with its Hausdorff
dimension $\sigma_M$. The constrained (unweighted) operator defined as

\begin{equation}
H_{M,t}[f](x) := \sum_{m \leq M} \frac{1}{(m+x)^2} f\left(\frac{1}{m+x}\right)
\end{equation}

directly appears here, together with its dominant eigenvalue $t \mapsto \lambda_M(t)$, as $\sigma_M$ coincides with the value $t$ for which $\lambda_M(t) = 1$. The strict inequality $\sigma_M < 1$ holds if $M < \infty$.

It is natural to study the rational numbers and the rqi numbers whose continued fraction expansion follows the same constraints, namely the sets $Q[M] := Q \cap R[M]$ and $P[M] := P \cap R[M]$. The rational case was studied by Cusick and Hensley; then Vallée \cite{27} provided a unified framework based on analytic combinatorics, where she analysed, for a fixed bound $M$, the probability $\pi(N,M)$ that a rational number of denominator at most $N$ (or a quadratic number $x$ with $\epsilon(x) \leq N$) has all its digits $\leq M$. In each case, the estimate

$$\pi(N,M) = N^{2(\sigma_M-1)}[1 + \eta_M(N)] \quad \text{where } \eta_M(N) \to 0 \text{ as } N \to \infty,$$

was proven to hold. Moreover, she also studied the mean value of the depth $p$ (in the rational case) or of the period length $p$ (in the quadratic irrational case), and proved the asymptotic estimates, for any fixed $M$,

$$\mathbb{E}_N[p] = \mu_M \log N[1 + \eta_M(N)] \text{ with } \mu_M := -\frac{2}{\lambda'_M(\sigma_M)}, \eta_M(N) \to 0.$$

We note that the constant $\mu_M$ is the same in the two cases.

She dealt in the two cases with the “constrained” analogs $Q_M(s)$ and $P_M(s)$ of univariate “unconstrained” generating functions $Q(s)$ and $P(s)$,

$$Q_M(s) := \sum_{x \in Q[M]} q(x)^{-s}, \quad P_M(s) := \sum_{x \in P[M]} \epsilon(x)^{-s},$$

which she related to the constrained (plain) transfer operator defined in (5.1). Her results were based on spectral properties of this constrained operator on the half-plane $\Re s > \sigma_M$, further transferred to the coefficients of the univariate series $P_M(s), Q_M(s)$. As she was (only) interested in average-case results, and did not know (at the moment) the existence of bounds à la Dolgopyat, she applied Tauberian theorems, which do not provide estimates for the remainder term $\eta_M(N)$ for $N \to \infty$.

We now mix her previous approach with our present methods, and exhibit a general “constrained” framework which combines the present rqi framework and the rational framework of \cite{1}. We introduce the bivariate constrained generating functions associated with some additive cost $C$,

$$Q_M(s, w) := \sum_{x \in Q[M]} e^{wC(x)} \cdot q(x)^{-s}, \quad P_M(s, w) := \sum_{x \in P[M]} e^{wC(x)} \cdot \epsilon(x)^{-s},$$
together with the weighted constrained operator
\[ H_{M,t,w}[f](x) = \sum_{m \leq M} e^{wc(m)} \left( \frac{1}{(m+x)^2} f \left( \frac{1}{m+x} \right) \right). \]

We (easily) obtain the following result:

**Theorem 5.1.** With a bound \( M \geq 2 \), and a cost of moderate growth, associate the additive cost \( C \) on each set \( \mathcal{Q}[M] \) or \( \mathcal{P}[M] \). If \( M \geq M_0 \) (for some \( M_0 \geq 3 \)), then on each of the two sets,
- the set \( \mathcal{Q}_N[M] \) of rationals \( x \in \mathcal{Q}[M] \) with denominator \( q(x) \leq N \),
- the set \( \mathcal{P}_N[M] \) of quadratic irrationals \( x \in \mathcal{P}[M] \) with \( \epsilon(x) \leq N \),

the distribution of \( C \) is asymptotically Gaussian (for \( N \to \infty \)), and there are constants \( \mu_M(c), \nu_M(c) \) for which
\[
\mathbb{P}_{N,M} \left[ \left| C(x) - \mu_M(c) \log N \right| \leq u \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-u^2/2} du + O \left( \frac{1}{\sqrt{\log N}} \right).
\]

**Sketch of proof.** We adapt the general description given in the introduction. Now the Dirichlet series \( Q_M(s,w), P_M(s,w) \) have both a pole on the curve \( s = 2\sigma_M(w) \) where \( \sigma_M(w) \) is implicitly defined by \( \lambda_M(\sigma_M(w),w) = 1 \), \( \sigma_M(0) = \sigma_M \).

Moreover, if \( M \) is large enough (say \( M \geq M_0 \)), the real number \( \sigma_M \) is close enough to 1, and the vertical line \( \Re s = \sigma_M \) is contained in the Dolgopyat strip of Theorem \( \Box \). Then, bounds à la Dolgopyat are available for the constrained transfer operator (which can be viewed as a perturbation of the unconstrained one, as in \[4\]). This entails the polynomial growth of the quasi-inverse \( (I - H_{M,t,w})^{-1} \), and thus the polynomial growth of \( Q_M(s,w), P_M(s,w) \) for \( |3s| \to \infty \) and \( w \) close to 0. The constants \( \mu_M(c) \) and \( \nu_M(c) \) are equal to the first and second derivatives of \( w \mapsto 2\sigma_M(w) \) at \( w = 0 \).

**5.3. Quadratic irrational numbers with bounded digits (II).**
Now, the bound \( M \) is no longer fixed, and we wish to describe the behavior of the probability \( \pi(N,M) \) (both in the rational and in the rqi case) when \( N \) and \( M \) both tend to \( \infty \). First, the behavior of the Hausdorff dimension \( \sigma_M \) of the set \( \mathcal{R}[M] \) has been studied by Hensley \[11\], who gave an asymptotic estimate for the speed of convergence of \( \sigma_M \) towards 1,
\[
2(\sigma_M - 1) = -\frac{2}{\zeta(2)} \frac{1}{M} - \frac{4}{\zeta(2)^2} \frac{\log M}{M^2} + O \left( \frac{1}{M^2} \right) \quad (M \to \infty).
\]

Then, the present authors \[4\] studied the probability \( \pi(N,M) \) in the rational case. Besides some specifications and extensions of a first result obtained by Hensley \[10\], they developed a methodology (based on principles
of analytic combinatorics) which can be easily transferred from the rational case to the rqi case and implies the following general result:

**Theorem 5.2.** There are an integer $M_0 = M_0 \geq 3$ and a real $\beta$ with $0 < \beta < 1/2$ such that, for any $N \geq 1, M \geq M_0$, the two probabilities $\pi(N, M)$ that

- a reduced quadratic irrational $x$ with size $\epsilon(x) \leq N$, or
- a rational number $x$ with denominator $q(x) \leq N$

has all its digits less than $M$, satisfy

$$\pi(N, M) = N^{2(\sigma_M - 1)}[1 + O(N^{-\alpha})] \cdot \left[1 + O\left(\frac{\log M}{M}\right)\right].$$

Here, $\sigma_M$ is the Hausdorff dimension of the set $\mathcal{R}[M]$, and $\beta$ is the width of the Dolgopyat strip.

This result exhibits a threshold phenomenon, already obtained by Hensley and precisely described in [4] in the rational case, depending on the relative order of $\sigma_M - 1$ (of order $O(1/M)$) with respect to $n := \log N$:

(a) If $M/n \to \infty$, then almost any number of size at most $N$ has all its digits less than $M$.

(b) If $M/n \to 0$, then almost any number of size at most $N$ has at least one of its digits greater than $M$.

### 5.4. Local limit laws: case of a lattice cost.

For any additive cost $C$ associated with a digit-cost $c$ of moderate growth, Baladi and Vallée [1] also obtained a Local Limit Theorem on rational trajectories. Moreover, when $c$ is a lattice cost, they proved an optimal speed of convergence. We recall that a cost is *lattice* if it is non-zero and there exists $L \in \mathbb{R}^+$ such that $c/L$ is integer-valued. The largest such $L$ is then called the *span* of $c$. For instance, the three costs of interest are integer-valued and thus lattice (with span 1).

Local limit theorems also deal with the moment generating function $M_N(w) := \mathbb{E}_N[\exp(wC)]$, but now with a complex number $w$ on the vertical line $\Re w = 0$. When the cost $c$ is lattice with span $L$, it is enough to study the generating function $P(s, w)$ or its analog $Z(s, w)$ for a pair $(s, w)$ with $s$ in a vertical strip near $\Re s = 2$ and $w \in ]-i\pi/L, +i\pi/L[$. It is straightforward to adapt the proof of [1] to the present framework and derive the result below which is expressed with the following scale:

$$Q(y, N) := \mu(c) \log N + y\sqrt{\nu(c) \log N} \quad (y \in \mathbb{R}).$$

**Theorem 5.3 (Local Limit Theorem for lattice costs).** For any lattice digit-cost $c$ of moderate growth and of span $L$, letting $\mu(c), \nu(c) > 0$ be the constants from Theorems 1.1 and 1.2, the following holds for the cost $C$
associated with $c$ on the set $\mathcal{P}_N$:
\[
P_N\left[ x \in \mathcal{P} \mid |C(x) - Q(y, N)| \leq L/2 \right] = \frac{1}{\sqrt{\nu(c) \log N}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} + O\left( \frac{1}{\log N} \right)
\]
with an $O$ uniform for $y \in \mathbb{R}$.

5.5. Local limit laws: case of a non-lattice cost. In [1], the authors did not succeed in estimating the speed of convergence when the cost is non-lattice; this problem is known to be difficult, even in the simpler case of a dynamical system with affine branches. The problem was later considered by Baladi and Hachemi, and then by Vallée [30] (always for rational trajectories) who obtained a local limit theorem with speed of convergence, which depended on arithmetical properties of the cost $c$. More precisely, when the cost satisfies a diophantine condition, then the speed depends on the irrationality exponent that intervenes in the diophantine condition, as we now explain.

We first recall that a real number $y$ is diophantine of exponent $\mu$ if there exists $C > 0$ such that the inequality $|x - p/q| > Cq^{-(2+\mu)}$ is satisfied for any pair $(p, q) \in \mathbb{N}^2$. We now state the diophantine condition on the cost $c$.

**Definition 5.4.** For any pair $(h, k) \in \mathcal{H}^+$ let $c(h, k) := |h|c(k) - |k|c(h)$. The cost $c$ is diophantine of exponent $\mu$ if there exists a triple $(h, k, \ell) \in (\mathcal{H}^+)^3$ for which

(i) $c(h, k)$ and $c(h, \ell)$ are non-zero,

(ii) the ratio $c(h, k)/c(h, \ell)$ is diophantine of exponent $\mu$.

Vallée [30] proved that the cost $C$ associated with a digit-cost $c$ (diophantine of exponent $\mu$) and restricted to the rational numbers of denominator at most $N$ asymptotically follows a local limit law, where the speed of convergence depends of the exponent $\mu$.

**Theorem D.** Consider a cost $c$ of moderate growth, diophantine of exponent $\mu$, and let $\chi := 6(\mu + 1)$. Then the following holds for the cost $C$ associated with $c$ on the set $\mathcal{P}_N$: for any $\epsilon$ with $\epsilon < 1/\chi$ and any compact interval $J \subset \mathbb{R}$, there exists a constant $M_J$ such that, for every $y \in \mathbb{R}$ and all integers $N \geq 1$,
\[
\left| \sqrt{\log N} P_N[C - Q(y, N) \in J] - |J| e^{-y^2/2} \sqrt{2\nu(c)\pi} \right| \leq M_J \frac{\log^\epsilon N}{N}.
\]

The proof of [30] is based on fine diophantine arguments, and is closely related to the generating function $Q(s, w)$ of the rational case: it is not clear if this result may be adapted to the generating function $P(s, w)$ of the rqi case.
6. Appendix: Nuclearity and traces of quasi-inverses. This Appendix is devoted to providing detailed proofs of Propositions 3.2 and 3.3. The main object here is the trace of various transfer operators. In which sense the usual notion of trace of matrices may be extended to operators that act on Banach spaces was mostly clarified by Grothendieck, for instance, in [9]. Later on, Ruelle and Mayer applied Grothendieck’s theory to the study of periodic points in dynamical systems, as explained in Mayer’s survey [18].

Most of the classical studies in the dynamical system context indeed deal with the Fredholm determinant $F(u)$ associated with a family of operators $G_t$, which satisfies the trace formula,

$$F(u,t) := \det(I - uG_t) = \exp\left[-\sum_{k \geq 1} \frac{u^k}{k} \text{Tr} G_t^k\right],$$

and its logarithmic derivative $t \mapsto \eta(t)$,

$$\eta(t) := \frac{\partial}{\partial u} \log F(u,t) \bigg|_{u=1} = \sum_{k \geq 1} \text{Tr} G_t^k.$$

Here, in an approach which appears to be new, we directly deal with this object, which is proven to exactly coincide with the trace of the quasi-inverse $(I - G_t)^{-1}$, without studying the Fredholm determinant.

6.1. Basic facts

**Definition 6.1 (Grothendieck).** Consider a Banach space $B$ that admits a Schauder basis. A linear operator $H : B \to B$ is nuclear if it admits a representation

$$H[f] = \sum_{n \geq 1} \lambda_n f_n^*[f] f_n$$

(6.1)

where $\{f_n\}$ and $\{f_n^*\}$ are families, respectively in $B$ and $B^*$, with $\|f_n\| \leq 1$ and $\|f_n^*\| \leq 1$, and $\lambda_n$ is a sequence of complex numbers that belongs to $\ell^1$. The sum $\sum_{n \geq 1} \lambda_n f_n^*[f_n]$ does not depend on the choice of the representation in (6.1), and defines the functional trace of the operator. If $q := \inf\{p \mid (\lambda_n) \in \ell^p\}$ the operator $H$ is said to be nuclear of order $q$.

Note that such a nuclear operator belongs to $\mathcal{L}(B)$ and is also compact. Its spectral trace is defined as the sum of its eigenvalues counted with algebraic multiplicities, if it exists. The following result obtained by Grothendieck [9] and used in Mayer’s [18] and Momeni–Venkov’s [19] works exhibits important properties of nuclear operators of order $q \leq 2/3$:

**Theorem B (Grothendieck).** The functional trace of a nuclear operator of order $q \leq 2/3$ coincides with its spectral trace.
6.2. Two useful lemmas. The following two lemmas describe (sufficient) conditions for dealing with the trace of a series of operators.

**Lemma 6.2.** Consider a Banach space \((B, \| \cdot \|)\) and a sequence of operators \(H_{[j]} \in \mathcal{L}(B)\) for \(j \in \mathcal{J}\). Assume there exists \(q_0 > 0\) for which the two following conditions hold:

(H1) Each operator \(H_{[j]}\) is nuclear of order \(q < q_0\), associated with a sequence of coefficients \(m \mapsto \lambda_{m,[j]}\).

(H2) The series of general term \(\sigma_j := \sum_{m=0}^{\infty} |\lambda_{m,[j]}|^{q_0}\) is convergent.

Then the following holds for the operator \(H := \sum_{j \in \mathcal{J}} H_{[j]}:\)

(i) \(H\) is nuclear of order \(q \leq q_0\).

(ii) If \(q_0 \leq 2/3\), then \(\text{Tr} H = \sum_{j \in \mathcal{J}} \text{Tr} H_{[j]}\).

**Proof.** Each operator \(H_{[j]}\) admits a decomposition of type (6.1) which involves a sequence \(m \mapsto \lambda_{m,[j]}\) that belongs to \(\ell^{q_0}\) and the inequalities \(\|f_{m,[j]}\|, \|f^*_{m,[j]}\| \leq 1\) hold. Hypothesis (H2) entails that the sequence \((m, j) \mapsto \lambda_{m,[j]}\) belongs to \(\ell^{q_0}\). Then the decomposition

\[
H[f] = \sum_{(m,j) \in \mathbb{N} \times \mathcal{J}} \lambda_{m,[j]} f_{m,[j]} f^*_{m,[j]} [f]
\]

(with \(\|f_{m,[j]}\|, \|f^*_{m,[j]}\| \leq 1\)) shows that the operator \(H\) is nuclear of order \(q \leq q_0\). Moreover, for \(q_0 \leq 2/3\), the spectral traces satisfy

\[
\text{Tr} H = \sum_{m,j} \lambda_{m,[j]} f^*_{m,[j]} [f_{m,[j]}] = \sum_{j \in \mathcal{J}} \sum_{m=0}^{\infty} \lambda_{m,[j]} f^*_{m,[j]} [f_{m,[j]}] = \sum_{j \in \mathcal{J}} \text{Tr} H_{[j]}.
\]

The next lemma is useful for dealing with operators which involve quasi-inverses.

**Lemma 6.3.** Consider an operator \(T \in \mathcal{L}(B)\) and a nuclear operator \(A\) of order \(q \leq 2/3\) together with its sequence \((\sigma_n)\) of coefficients with \(\ell^1\) norm \(\sigma\). Then:

(i) The operator \(AT\) is nuclear of order \(q \leq 2/3\), and

\[
|\text{Tr} AT| \leq \sigma \|T\|.
\]

(ii) Assume that the spectral radius of \(T\) is strictly less than 1. Then the operator \((I - T)^{-1}\) exists in \(\mathcal{L}(B)\), the operator \(A(I - T)^{-1}\) is nuclear of order \(q \leq 2/3\), and

\[
\text{Tr} A(I - T)^{-1} = \text{Tr} \sum_{k=1}^{\infty} AT^k = \sum_{k=1}^{\infty} \text{Tr} AT^k.
\]
Proof. Assertion (i) is clear. For (ii), the operator $A(I - T)^{-1}$ is nuclear of order $q \leq 2/3$. Now, for two operators $V$ and $U$, (6.2) yields

$$|\text{Tr} AV - \text{Tr} AU| = |\text{Tr} A(V - U)| \leq \sigma\|V - U\|.$$  

Applying this with $V_K := \sum_{k=1}^K T^k$, $U := (I - T)^{-1}$, together with the bound $\|V_K - U\| \ll \|T\|^K$ ends the proof. ■

6.3. Nuclearity and trace properties of $H_{t,w}$ on $A_\infty(V)$. We now return to the present framework, where the Banach space $A_\infty(V)$, defined in (3.4), admits the Taylor basis $\{(z - 1)^m \mid m \in \mathbb{Z}_{\geq 0}\}$ as a Schauder basis.

The next result summarizes the main properties of the various operators of interest, relative to their nuclearity and their traces.

**Proposition 6.4.**

(a) For any $h \in H^+$, the weighted transfer operator $H_{[h],t,w}$ is nuclear of order 0, with

$$\text{Tr} H_{[h],t,w} = \alpha(h)^{2t} \frac{\exp(wc(h))}{1 - (-1)^{|h|}\alpha(h)^2}. \quad (6.3)$$

(Recall that $|h|$ is the depth of $h$.)

(b) Let $(s, w) \in \Gamma(a)$ (see (2.15)) with $a > 3/4$. Then the complete operator $H_{t,w}$ and its iterates $H_{t,w}^k$ (for $k \geq 1$) are nuclear of order $q \leq 2/3$. For $(s, w) \in \Gamma(a)$,

$$\text{Tr} H_{t,w}^k = \sum_{h \in H^k} \text{Tr} H_{[h],t,w} = \sum_{h \in H^k} \alpha(h)^{2t} \frac{\exp(wc(h))}{1 - (-1)^k\alpha(h)^2} \quad \text{for } k \geq 1. \quad (6.4)$$

For any $k \geq 1$, the mapping $(t, w) \mapsto \text{Tr} H_{t,w}^k$ is analytic on $\Gamma(a)$.

(c) Let $(t, w) \in \Gamma(a)$ with $a > 1$. Suppose that the spectral radius of $H_{t,w}$ is less than 1. Then the quasi-inverse $H_{t,w}(I - H_{t,w})^{-1}$ and the even quasi-inverse $E_{t,w} := H_{t,w}^2(I - H_{t,w}^2)^{-1}$ are nuclear of order $q \leq 2/3$, and

$$\text{Tr} H_{t,w}(I - H_{t,w})^{-1} = \sum_{k \geq 1} \text{Tr} H_{t,w}^k, \quad \text{Tr} E_{t,w} = \sum_{k \text{ even}} \text{Tr} H_{t,w}^k. \quad (6.5)$$

Moreover, the mappings $(t, w) \mapsto \text{Tr} H_{t,w}(I - H_{t,w})^{-1}$ and $(t, w) \mapsto \text{Tr} E_{t,w}$ are analytic on $\Gamma(a)$ with $a > 1$, and

$$\text{Tr} H_{t,w}(I - H_{t,w})^{-1} = \sum_{h \in H^+} \alpha(h)^{2t} \frac{\exp(wc(h))}{1 - (-1)^{|h|}\alpha(h)^2}, \quad (6.6)$$

$$\text{Tr} E_{t,w} = \sum_{h \in H^+, |h| \text{ even}} \alpha(h)^{2t} \frac{\exp(wc(h))}{1 - \alpha(h)^2}. \quad (6.6)$$
Proof. (a) Mayer and Momeni–Venkov have studied the nuclearity of the unweighted transfer operator. We follow the approach of [19 Lemma 2.7], where the authors use the Schauder basis of \( A_\infty(V) \) formed by the polynomials \((z-1)^m\) with \(m \geq 0\). We adapt their method to the case of the weighted operator \( H_{[h],t,w} \) for any \( h \in \mathcal{H}^+ \).

Using the Taylor expansion of \( f \) at \( z = 1 \) in the expression of \( H_{[h],t,w}[f](z) \) provides a new expression

\[
H_{[h],t,w}[f](z) = e^{wc(h)} \left[ \sum_{m=0}^{\infty} \frac{f(m)(1)}{m!} (h(z) - 1)^m \right] h^t(z),
\]

which involves the analytic extensions \( h \) of \( |h'| \). We consider here the case \( r = 5/4 \) in Property (P3), and the related radii \( r, \tilde{r} \). For any \( h \in \mathcal{H}^+ \), and \( t = \sigma + i\tau \), we introduce the bound

\[
a_h(t) := \sup_{z \in V} |h^t(z)| \leq e^{\pi|\tau|} |h'(-1/4)|^\sigma.
\]

For \( m \geq 0 \) and \( h \in \mathcal{H}^+ \), define

\[
f_{m,[h]}(z) = \frac{h^t(z)}{a_h(t)} \left( \frac{h(z) - 1}{\tilde{r}} \right)^m, \quad f_{m,[h]}^*[f] = r^m \frac{f(m)(1)}{m!}.
\]

Then by (P3) and Cauchy estimates,

\[
\|f_{m,[h]}\| \leq 1, \quad \|f_{m,[h]}^*\| \leq 1,
\]

and we have the decomposition

\[
H_{[h],t,w}[f] = \sum_{m=0}^{\infty} \lambda_{m,[h]} f_{m,[h]} f_{m,[h]}^*[f] \quad \text{with} \quad \lambda_{m,[h]} := \left( \frac{r}{\tilde{r}} \right)^{-m} e^{wc(h)} a_h(t).
\]

As the sequence \( m \mapsto \lambda_{m,[h]} \) is geometric and thus belongs to \( \ell^p \) for any \( 0 < p \leq 1 \), this proves the nuclearity of order 0. Each component operator \( H_{[h],t,w} \) is a composition operator of the form \( f \mapsto g \cdot f \circ h \) where \( g := \exp(wc(h)) \cdot h^t \).

By (P3), it acts on \( A_\infty(V) \). Moreover, due to (P3)(i), any branch \( h \in \mathcal{H}^+ \) maps the domain \( V \) strictly inside itself. According to [18 Lemma 7.10], the spectrum of such a composition operator acting on \( A_\infty(V) \) is the set

\[
\{ \mu_n := g(x_n) \cdot (h'(x_n))^n | n \in \mathbb{Z}_{\geq 0} \},
\]

which involves the unique fixed point \( x_n \) of \( h \). Every eigenvalue is simple. This result applies to \( H_{[h],t,w} \) and provides the expression given in (6.3) for the trace of \( H_{[h],t,w} \) which involves \( \alpha(h) \) via (2.4).

(b) For any fixed \( h \in \mathcal{H}^+ \), any \((t,w)\), and any real \( p > 0 \), the sequence \( m \mapsto \lambda_{m,[h]} \) satisfies, by Lemma 1,

\[
\sum_{m \geq 0} |\lambda_{m,[h]}|^p \leq \left( \frac{\tilde{r}}{r} \right)^{pm} e^{\nu p + c(h)} |a_h(t)|^p \ll \left( \frac{\tilde{r}}{r} \right)^{pm} e^{\pi p|\tau|} |h'(-1/4)|^\sigma |\nu - d| p.
\]
Consider now a pair \((t, w) \in \Gamma(a)\) and an exponent \(p\) for which \(ap := (\sigma - d|\nu|)p > 1/2\). Then using the bound (3.7) we obtain, for any \((t, w) \in \Gamma(a)\),

\[
\sum_{(m, h) \in \mathbb{N} \times H_k} |\lambda_{m, [h]}|^p \leq \frac{e^{\pi |\tau|}}{1 - (\tilde{r}/r)^p} H_{ap}^k[1](-1/4).
\]

Thus, for \((t, w) \in \Gamma(a)\), the operator \(H_{t, w}^k\) is nuclear of order \(q < 1/(2a)\), and is nuclear of order \(q < 2/3\) whenever \((t, w) \in \Gamma(a)\) with \(a > 3/4\). Then, for \(a > 3/4\), we apply Lemma 6.2 with \(q_0 = 2/3\), to get

\[
(6.7) \quad \text{Tr} \ H_{t, w}^k = \sum_{h \in H_k} \text{Tr} \ H_{[h], t, w}.
\]

The explicit expression in (6.3) shows that the mapping \((t, w) \mapsto \text{Tr} \ H_{t, w}^k\) is analytic on \(\Gamma(a)\).

(c) (6.5) is a direct application of Lemma 6.3(ii) with \(A = T = H_{t, w}\) or \(A = T = H_{t, w}^2\). When \((t, w) \in \Gamma(a)\) with \(a > 1\) the spectral radii of \(H_{t, w}\) and of \(H_{t, w}^2\) are less than 1 (see Proposition 3.1), hence (6.5) holds. On \(\Gamma(a)\), the expression of \(\text{Tr} \ H_{t, w}^k\) given in (6.4) is the general term of a series which is absolutely convergent. This yields (6.6).

6.4. Spectral dominant properties of \(H_{t, w}\) on \(A_\infty(V)\). For real pairs \((t, w) \in \Gamma(a)\) (with \(a > 1/2\)), the operator \(H_{t, w}\) has strong positivity properties that entail the existence of dominant spectral objects, in the same vein as the Perron–Frobenius properties. Also, for \((t, w) \in \Gamma(a)\) (with \(a > 1/2\)), the map \((t, w) \mapsto H_{t, w}\) is analytic in the sense of Kato [13, Chapter 3]. As \(H_{t, w}\) is compact, the dominant spectral objects defined for real pairs may be extended (with analytic perturbation of the dominant part of the spectrum) when \((t, w)\) is close to a real pair.

We are mainly interested in a neighborhood of the pair \((1, 0)\), and we begin with a complex neighborhood \(T = S \times W\) of \((1, 0)\). For \((t, w) \in T\), a spectral decomposition for the operator \(H_{t, w}\) holds and there exist operators \(P_{t, w}\) and \(N_{t, w}\) for which

\[
(6.8) \quad H_{t, w} = \lambda(t, w)P_{t, w} + N_{t, w}
\]

with \(P_{t, w} \circ N_{t, w} = N_{t, w} \circ P_{t, w} = 0\). As \(H_{t, w}\) is compact, and due to the equality \(\lambda(1, 0) = 1\), the spectral radius \(R(t, w)\) of \(N_{t, w}\) is \(\leq R < 1\). The operator \(P_{t, w}\) is rank one, of the form

\[
(6.9) \quad P_{t, w}[f](x) = f_{t, w}(x)\mu_{t, w}[f]
\]

and involves the dominant eigenfunction \(f_{t, w}\) of \(H_{t, w}\) and the eigenmeasure \(\mu_{t, w}\) of \(H^*_{t, w}\). Moreover, with analytic perturbation of the dominant part of the spectrum, all the spectral objects depend analytically on \(T\).
The spectral decomposition extends to the iterates,
\begin{equation}
H_{t,w}^k = \lambda(t,w)^k P_{t,w} + N_{t,w}^k \quad \text{for any } k \geq 1.
\end{equation}
As the “remainder quasi-inverse” \( N_{t,w}(I - N_{t,w})^{-1} \) is analytic on \( T \), and the quotient \( \lambda(t,w)/(1 - \lambda(t,w)) \) is meromorphic on \( T \), the decomposition of the quasi-inverse,
\begin{equation}
(I - H_{t,w})^{-1} = \frac{\lambda(t,w)}{1 - \lambda(t,w)} P_{t,w} + (I - N_{t,w})^{-1},
\end{equation}
proves that \( (I - H_{t,w})^{-1} \) is meromorphic on \( T \). There is an analogous proof for the “even quasi-inverse” \( E_{t,w} \) defined in (3.3):
\begin{equation}
E_{t,w} := H_{t,w}^2(I - H_{t,w}^2)^{-1} = \frac{\lambda(t,w)^2}{1 - \lambda(t,w)^2} P_{t,w} + N_{t,w}^2(I - N_{t,w}^2)^{-1}.
\end{equation}

6.5. Spectral decomposition and traces. The projector \( P_{t,w} \) is clearly nuclear of order 0 (as any operator of finite rank), with trace 1. Then, as \( T \) is a subset of \( \Gamma(a) \) for some \( a > 3/4 \), and due to Proposition 6.4, all the operators involved are nuclear of order \( q < 2/3 \). Moreover, as the dominant eigenvalue \( \lambda(t,w) \) and the traces \( \text{Tr} H_{t,w}^k \) are well-defined and analytic, it is the same for the operators \( N_{t,w}^k \): their traces are well-defined and analytic on \( T \). As the norm of the operator \( N_{t,w} \) is strictly less than 1 on \( T \), Lemma 6.3 applies and entails that the sequence
\[
\left[ \sum_{k=1}^{K} \text{Tr} H_{t,w}^{2k} \right] - \sum_{k=1}^{K} \lambda(t,w)^{2k}
\]
has a limit (for \( K \to \infty \)) that is analytic on \( T \). As
\[
\lim_{K \to \infty} \sum_{k=1}^{K} \lambda(t,w)^{2k} = \frac{\lambda(t,w)^2}{1 - \lambda(t,w)^2}
\]
is meromorphic on \( T \), we have shown the equality
\[
\text{Tr} H_{t,w}^2(I - H_{t,w}^2)^{-1} = \sum_{k=1}^{\infty} \text{Tr} H_{t,w}^{2k}
\]
that defines a meromorphic function on \( T \). In particular,
\begin{equation}
\text{Tr} E_{t,w} = \frac{\lambda(t,w)^2}{1 - \lambda(t,w)^2} + \text{Tr} N_{t,w}^2(I - N_{t,w}^2)^{-1}.
\end{equation}
defines a meromorphic function on \( T \), whose poles are located on the curve \( \{(t,w) \in T \mid \lambda(t,w)^2 = 1\} \) that contains the point \( (1,0) \) and thus the curve \( \{(t,w) \in T \mid \lambda(t,w) = 1\} \).
6.6. Properties near \((t, w) = (1, 0)\). The operator \(H_{1,0}\) is a density transformer, and thus \(\lambda(1,0) = 1\). Furthermore, the dominant eigenfunction \(f_{1,0}\) coincides with the Gauss density \(\psi(x) := (1/\log 2) \cdot 1/(1+x)\). There are simple formulae for the two partial derivatives of \(\lambda(t, w)\) at \((1,0)\). They are easy consequences of the spectral equality \(H_{t,w}[f_{t,w}] = \lambda(t,w)f_{t,w}\), the fact that \(H_{1,0}\) is a density transformer, and the explicit forms of the derivatives of the operator. One obtains

\[
-\lambda'_t(1,0) = \int \log |T'(x)| \psi(x) \, dx = \frac{\pi^2}{6 \log 2},
\]

\[
\lambda'_w(1,0) = \sum_{h \in \mathcal{H}} c(h) \int_{\mathcal{H}(x)} \psi(x) \, dx.
\]

The constant \(-\lambda'_t(1,0) > 0\) is the Kolmogorov entropy (denoted here by \(E\)) and \(\lambda'_w(1,0)\) is the average \(\mathbb{E}[c]\) of the cost \(c\) with respect to the Gauss density \(\psi\). They are both non-zero.

Since both derivatives are non-zero at \((t, w) = (1, 0)\), there is a complex neighborhood \(S_1\) of \(t = 1\), a neighborhood \(W_1\) of \(0\) and a unique analytic function \(\sigma : W_1 \to \mathbb{C}\), with \(\sigma(0) = 1\), for which

\[
\{(t, w) \in S_1 \times W_1 \mid \lambda(t, w) = 1\} = \{(t, w) \in S_1 \times W_1 \mid t = \sigma(w), w \in W_1\},
\]

\[
\sigma'(w) = -\frac{\lambda'_w(\sigma(w), w)}{\lambda'_t(\sigma(w), w)}.
\]

First, we may choose the neighborhood \(S_1\) to be a rectangle of the form \(S_1 = \{s = \sigma + i\tau \mid |\tau| < \tau_1, |\sigma-1| \leq \delta_1\}\). Second, “perturbing” the inequality \(\lambda'_t(1,0) \neq 0\), and taking a (possibly) smaller neighborhood \(T_1 := S_1 \times W_1\), yields a constant \(A > 0\) for which the ratio \(|\lambda(t, w) - 1|/|s - \sigma(w)|\) is at least \(A > 0\) on \(\overline{T}_1\) (and its inverse is then bounded).

Then, for each \(w \in W_1\), the map \(t \mapsto \text{Tr} \mathbf{E}_{t,w}\) has a unique (simple) pole on \(S_1\), located at \(s = \sigma(w)\), with residue

\[
\text{Res}[\text{Tr} \mathbf{E}_{t,w}; s = \sigma(w)] = \frac{1}{2} \left( \frac{-1}{\lambda'_t(\sigma(w), w)} \right).
\]

**Acknowledgements.** We would like to thank Corinna Ulcigrai for her suggestion to point out here the relation between continued fractions and the Farey tessellation of the hyperbolic plane that we have described in Section 1.2. We are also grateful to Valérie Berthé, Cristian Conde and Pablo Rotondo for interesting discussions about rqi numbers and traces of operators.

This work was partially funded by the ANR projects Dyna3S (ANR-13-BS02-0033) and CoDys (ANR-18-CE40-0007), the STIC-AmSud Project AleaEnAmSud, and the French-Argentinian Laboratories in Computer Science INFINIS (www.infinis.org) and then SINFIN (www.lia-sinfin.org). Eda
Cesaratto was partially supported by the grant PIO CONICET-UNGS 14420140100027.

References

[1] V. Baladi and B. Vallée, *Euclidean algorithms are Gaussian*, J. Number Theory 110 (2005), 331–386.
[2] V. Berthé, L. Lhote, and B. Vallée, *Probabilistic analyses of the plain multiple gcd algorithm*, J. Symbolic Comput. 74 (2016), 425–474.
[3] F. Boca, *Products of matrices $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and the distribution of reduced quadratic irrationals*, J. Reine Angew. Math. 606 (2007), 149–165.
[4] E. Cesaratto and B. Vallée, *Small quotients in Euclidean algorithms*, Ramanujan J. 24 (2011), 183–218.
[5] H. Daudé, Ph. Flajolet, and B. Vallée, *An average-case analysis of the Gaussian algorithm for lattice reduction*, Combin. Probab. Comput. 6 (1997), 397–433.
[6] D. Dolgopyat, *On decay of correlations in Anosov flows*, Ann. of Math. 147 (1998), 357–390.
[7] Ch. Faivre, *Distribution of Lévy constants for quadratic numbers*, Acta Arith. 61 (1994), 1–34.
[8] Ph. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, 2009.
[9] A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires*, Mem. Amer. Math. Soc. 16 (1955).
[10] D. Hensley, *The largest digit in the continued fraction expansion of a rational number*, Pacific J. Math. 151 (1991), 237–255.
[11] D. Hensley, *Continued fraction Cantor sets, Hausdorff dimension, and functional analysis*, J. Number Theory 40 (1992), 336–358.
[12] H.-K. Hwang, *Large deviations for combinatorial distributions. I. Central limit theorems*, Ann. Appl. Probab. 6 (1996), 297–319.
[13] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1980.
[14] D. Kelmer, *Quadratic irrational and linking numbers of modular knots*, J. Modern Dynam. 6 (2012), 539–561.
[15] E. Landau, *Über die Anzahl der Gitterpunkte in gewissen Bereichen*, IV. Nachr. Ges. Wiss. Gött. Math.-Phys. Kl. 1924, 137–150.
[16] L. Lhote, *Computation of a class of continued fraction constants*, in: Proceedings of the Sixth Workshop on Algorithm Engineering and Experiments and the First Workshop on Analytic Algorithmics and Combinatorics, L. Arge et al. (eds.), 2004, 199–210.
[17] D. H. Mayer, *The thermodynamic formalism approach to Selberg’s zeta function for PSL(2,Z)*, Bull. Amer. Math. Soc. 25 (1991), 55–70.
[18] D. H. Mayer, *Continued fractions and related transformations*, in: Ergodic Theory, Symbolic Dynamics, and Hyperbolic Spaces, T. Bedford et al. (eds.), Oxford Sci. Publ., Oxford Univ. Press, Oxford, 1991, 175–222.
[19] A. Momeni and A. Venkov, *Mayer’s transfer operator approach to Selberg’s zeta function*, St. Petersburg Math. J. 24 (2013), 529–553.
[20] W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque 187-188 (1990).
[21] M. Pollicott, *Distribution of closed geodesics on the modular surface and quadratic irrationals*, Bull. Soc. Math. France 114 (1986), 431–446.
M. Pollicott and R. Sharp, *Exponential error terms for growth functions on negatively curved surfaces*, Amer. J. Math. 120 (1998), 1019–1042.

M. Roux, *Théorie de l’information, séries de Dirichlet, et analyse d’algorithmes*, PhD thesis, Univ. of Caen, 2011.

D. Ruelle, *Zeta-functions for expanding maps and Anosov flows*, Invent. Math. 34 (1976), 231–242.

C. Series, *The modular surface and continued fractions*, J. London Math. Soc. (2) 31 (1985), 69–80.

A. Ustinov, *Spin chains and Arnold’s problem on the Gauss–Kuz’min statistics for quadratic irrationals*, Sb. Math. 204 (2013), 762–779.

B. Vallée, *Dynamique des fractions continues à contraintes périodiques*, J. Number Theory 2 (1998), 183–235.

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

B. Vallée, *Euclidean dynamics*, Discrete Contin. Dynam. Systems 15 (2006), 281–352.

B. Vallée, *The Euclid algorithm is “totally” Gaussian*, in: Analysis of Algorithms (Montreal, 2012), N. Broutin and L. Devroye (eds.), Discrete Math. Theoret. Computer Sci. AQ, Nancy, 2012, 283–301.