

STATISTICAL DISTRIBUTION OF THE STERN SEQUENCE

S. BETTIN, S. DRAPPEAU, AND L. SPIEGELHOFER

ABSTRACT. We prove that the Stern diatomic sequence is asymptotically distributed according to a normal law, on a logarithmic scale. This is obtained by studying complex moments, and the analytic properties of a transfer operator.

1. INTRODUCTION

The Stern diatomic sequence [Ste58] is the sequence defined by the particularly simple recurrence relation

$$(1.1) \quad s(0) = 0, \quad s(1) = 1, \quad s(2n) = s(n), \quad s(2n + 1) = s(n) + s(n + 1)$$

for all $n \geq 1$. The first few terms are

$$(s(n))_{n \geq 0} = (0, 1, 1, 2, 1, 3, 2, 3, 1, 4, 3, 5, 2, 5, 3, 4, 1, \dots).$$

It is an example of a 2-regular sequence [AS03, Chapter 16, Exercise 32], and enjoys various connections with mathematical objects. As important examples, the Stern sequence is related to the Stern-Brocot and the Calkin-Wilf tree.

Starting from the set $F_0 = \{\frac{0}{1}, \frac{1}{1}\}$, for each $N \in \mathbb{N}_{\geq 0}$, let F_{N+1} be built from F_N by inserting, between any two consecutive fractions $\frac{a}{b}$ and $\frac{c}{d}$, its median $\frac{a+c}{b+d}$.

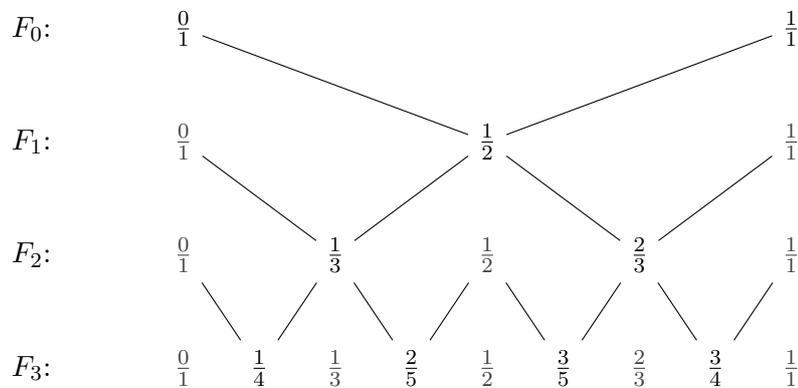


FIGURE 1. The first four rows of the Stern-Brocot tree

We may present the resulting construction as an infinite “tree”, labelled by rationals in $[0, 1]$, which is known as the Stern-Brocot tree¹: see for instance chapter 4.5 of [GKP94] or section 1.5.1 of [KMS16]. The numerators and denominators appearing in F_N , ordered by size of the fraction, are respectively the values of $s(m)$ and $s(m + 2^N)$, for $m \in [0, 2^N]$.

Let $\mathcal{I}_N = \mathbb{Z} \cap [2^N, 2^{N+1})$. In this paper, we will study properties of the values $s(n)$ for $n \in \mathcal{I}_N$: as we mentioned above, these are the denominators of the elements of the N -th row of the Farey tree described above. Several properties of $(s(n))_{n \in \mathcal{I}_N}$ were recorded by

Date: April 10, 2018.

2010 Mathematics Subject Classification. Primary 11B57; Secondary 37C30, 37A45.

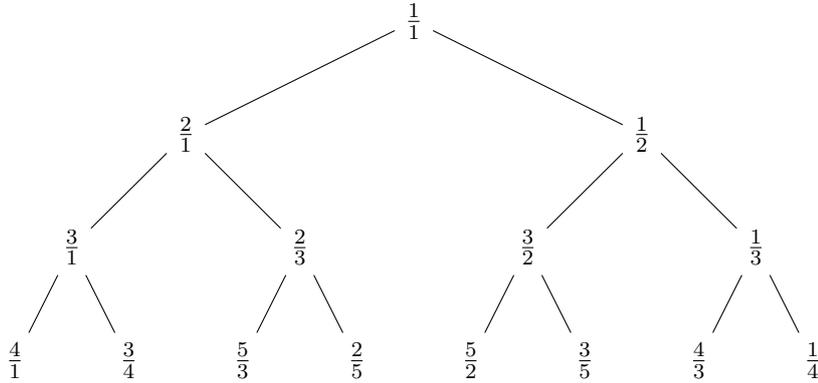
Key words and phrases. Stern diatomic sequence, transfer operator, central limit theorem.

¹Classical constructions start from $\{\frac{0}{1}, \frac{1}{0}\}$, which makes little difference for our purposes.

Stern [Ste58] and Lehmer [Leh29]. There has been much interest in understanding the structure of the largest values of $s(n)$ [CT14, Def16, Lan14, Pau17, CS17]: as Lehmer showed, we have $\max_{n \in \mathcal{I}_N} s(n) = F_{N+2}$, where $(F_r)_r$ is the Fibonacci sequence. Recently, Paulin [Pau17] gave a complete description of the $\lfloor N/2 \rfloor$ largest values taken by $(s(n))_{n \in \mathcal{I}_N}$: they are given by various combinations of Fibonacci numbers.

Another interpretation of the Stern sequence can be obtained from the *Calkin-Wilf tree* [CW00]: it is the infinite binary tree, labelled by positive rationals in reduced form, starting from $\frac{1}{1}$, and where each node $\frac{a}{b}$ has children $\frac{a+b}{b}$ and $\frac{a}{a+b}$. Each positive rational appears exactly once. In this case the denominators appearing at level N are the values of $s(m+2^N)$ for $m \in [0, 2^N)$.

FIGURE 2. First four rows of the Calkin-Wilf tree



In the present paper, we are interested in the question of the *statistical* distribution of values of $s(n)$. A relevant setting consists in endowing, for each $N \geq 0$, the finite set $\mathcal{I}_N = \mathbb{Z} \cap [2^N, 2^{N+1})$ with the uniform probability measure; let \mathbf{S}_N be the random variable

$$\mathbf{S}_N = s(n)$$

where n is taken uniformly randomly in \mathcal{I}_N . Another way to look at this is the following: start from the root of the Calkin-Wilf, and follow a walk of N steps down the tree, choosing the left or right child with equal probability. Then the random variable \mathbf{S}_N is the denominator of the fraction eventually encountered.

Our main result is the following effective central limit theorem for $\log \mathbf{S}_N$.

Theorem 1.1. *For some constants $\alpha, \sigma > 0$, as N tends to infinity, the values $(\log s(n))$ are asymptotically distributed according to a Gaussian law, with mean αN and variance $\sigma^2 N$: for $t \in \mathbb{R}$ satisfying $t = O(N^{1/6})$, we have*

$$(1.2) \quad \mathbb{P}_N \left[\frac{\log \mathbf{S}_N - \alpha N}{\sigma \sqrt{N}} \leq t \right] = \int_{-\infty}^t \frac{e^{-v^2/2}}{\sqrt{2\pi}} dv + O\left(\frac{(1+t^2)e^{-t^2/2}}{\sqrt{N}} \right).$$

Moreover, for some $(\nu_1, \nu_2) \in \mathbb{R}^2$ and $\theta \in [0, 1)$, we have

$$(1.3) \quad \begin{aligned} \mathbb{E}_N[\log \mathbf{S}_N] &= \alpha N + \nu_1 + O(\theta^N), & ([\text{Bac17}]) \\ \mathbb{V}_N[\log \mathbf{S}_N] &= \sigma^2 N + \nu_2 + O(N\theta^N). \end{aligned}$$

Remark. – This answers a question posed in [Lan14]; the error term in (1.2) is relevant and optimal in the whole central limit range $t = O(N^{1/6})$.

– The constants α and σ are numerically close to

$$\alpha \approx 0.396212\dots, \quad \sigma \approx 0.148905\dots$$

Both constants admit expressions as integrals of elementary functions with respect to a certain singular measure, the Minkowski measure (see formulas (2.4) and (2.7)). They are studied in more detail in Section 2.1 below.

- Formula (1.3) was proved very recently by Bacher [Bac17, Theorem 12.1] with a strong quantitative error term $O(2^{-N})$. The approach used there does not require complex analysis, but is ineffective for other moments than the first.

We will in fact prove that small complex moments of \mathbf{S}_N admit a quasi-powers expansion in the sense of Hwang [Hwa98, Hwa96], meaning that $\mathbb{E}_N[(\mathbf{S}_N)^\tau]$ behaves asymptotically as $A(\tau)^N B(\tau)$, for some holomorphic functions A, B in the neighborhood of the origin.

Theorem 1.2. *For some $\eta > 0$ and $\theta \in [0, 1)$, there exist holomorphic functions U, V on the disc $\{\tau \in \mathbb{C} : |\tau| \leq \eta\}$ such that*

$$(1.4) \quad \mathbb{E}_N[(\mathbf{S}_N)^\tau] = \exp\{NU(\tau) + V(\tau)\}(1 + O(\theta^N)),$$

uniformly in $N \geq 1$ and $|\tau| \leq \eta$. Moreover, $U''(0) \neq 0$.

This kind of estimates gives rise to the asymptotic normal law described in Theorem 1.1, with speed of convergence, but also to bounds on probabilities of large deviations, and asymptotic formulae for $\mathbb{E}[(\log \mathbf{S}_N)^k]$, with error term $O_k(N^{k-1}\theta^N)$, for any fixed $k \geq 1$. We refer to the above-quoted papers of Hwang [Hwa98, Hwa96], and to Chapter IX.5 of [FS09], for more explanations.

In principle, explicit estimates for θ could be obtained numerically. Experiments seem to suggest that any fixed $\theta > 1/2$ is admissible, at the cost of reducing the value of η accordingly.

2. OVERVIEW

To place Theorem 1.1 into context, we quote from [CW00] the fact that $(s(n+1)/s(n))_{n \in \mathcal{I}_N}$ is the N -th row of the Calkin-Wilf tree (Figure 2). It is easy to deduce from this the expression

$$(2.1) \quad \begin{pmatrix} s(n+1) \\ s(n) \end{pmatrix} = A_{\varepsilon_0} \cdots A_{\varepsilon_{N-1}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where $n = 2^N + \sum_{j=0}^{N-1} \varepsilon_j 2^j$, $\varepsilon_j \in \{0, 1\}$, $A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Picking $n \in \mathcal{I}_N$ at random means choosing $\varepsilon_j \in \{0, 1\}$ independently with equal probability. This is an instance of a problem about random products of matrices, which is a vast and active area of research; we quote the seminal papers [Bel54, Fur63], and we direct to the recent monograph [BQ16b] for more references. A general result such as Theorem 1.1 of [BQ16a] would indeed yield a slightly weaker version of our estimate (1.2). We also refer to [EH18] for a recent work concerned with a related situation (a “law of large numbers” for measures satisfying a recurrence relation similar to (1.1)).

In order to obtain our precise statements, in particular Theorem 1.2, we rely on similar pools of ideas; however we will take a different approach, and cast the arguments in a more direct form, by exploiting a connection between the Stern sequence and the Minkowski question-mark function [Min05, Fig.7]. As we will see shortly, this function arises naturally in our problem as a conjugacy between the two dynamical systems underlying the recursion formulas (1.1): the binary map and the Farey map, describing respectively the transformation rules of n and $s(n)$. Let us first describe it following [Den38], as the function mapping an irrational number $x \in (0, 1]$ to

$$(2.2) \quad ?(x) := \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2^{a_1 + \dots + a_n}}, \quad \text{if } x = [0; a_1, a_2, \dots] = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots}}}.$$

This function is extended as a strictly increasing, continuous bijection from $[0, 1]$ to itself. Let us define for notational convenience

$$\Psi(x) := ?(x), \quad \Phi(x) := ?^{-1}(x).$$

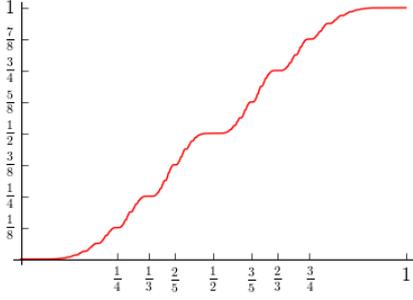


FIGURE 3A. Graph of the Minkowski function $\Psi(x) = ?(x)$.

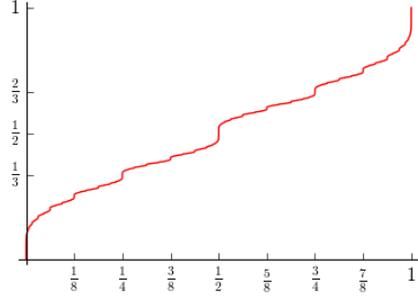


FIGURE 3B. Graph of Conway's box function $\Phi(x) = ?^{-1}(x)$.

These functions are drawn in Figures 3a and 3b.

The question-mark function was introduced by Minkowski for its properties of mapping rational and quadratic irrational numbers into, respectively, dyadic and non-dyadic rational numbers. By construction it is increasing and continuous, but it is not absolutely continuous with respect to the Lebesgue measure, and in fact it is singular [Sal43]: it has zero derivative for Lebesgue-almost every $x \in [0, 1]$. It is however Hölder continuous with Lipschitz exponent $\frac{\log 2}{\sqrt{5}+1}$.

An important property for us is the fact that the Minkowski function is a topological conjugation between two maps from $[0, 1]$ to itself, the Farey map and the binary map, which are implicitly at play in the recurrence relation (1.1). We will detail this further below in Section 2.2; a very practical form of this fact is given by the identity [Bac17, Proposition 2.1]

$$\frac{s(m)}{s(2^N + m)} = \Phi\left(\frac{m}{2^N}\right)$$

valid for all $N \geq 0$ and $m \in \{0, 1, \dots, 2^N\}$. This formula will be our starting point: it provides a fast algorithm to express $s(n)$ for $n \in \mathcal{I}_N$ in terms of product of values of Φ , and brings the problem into the framework of dynamical analysis of Euclidean algorithms [Val03, BV05].

In the rest of this section, we will first gather some facts about the mean-value α (Section 2.1), then we provide a naive heuristic towards Theorem 1.1 (Section 2.2), and finally we outline our proof (Section 2.3).

2.1. The mean-value α and the variance σ^2 . The constants α and σ appearing in the mean-value estimate (1.3) admit explicit expressions in terms of the Minkowski function. Let μ denote the *Minkowski measure* on the interval $[0, 1]$,

$$(2.3) \quad d\mu(x) = d\Psi(x) = d(\Phi^{-1}(x)).$$

Then the constant α is given by

$$(2.4) \quad \alpha = -\frac{1}{2} \int_0^1 (\log x) d\mu(x).$$

The next theorem records and links several different expressions of the constant α and σ .

Theorem 2.1. (i) Let $m_k := \int_0^1 x^k d\mu(x)$. Then

$$(2.5) \quad \alpha = \log 2 - \sum_{k=1}^{\infty} \frac{m_k}{k2^k};$$

this agrees with the definition given in [Bac17, Théorème 12.1].

(ii) We have

$$\alpha = \int_0^1 \log(1+x) d\mu(x);$$

therefore the μ -almost-sure Lipschitz exponent β of $x \mapsto \Psi(x)$, given in [Kin60], is related to α by $\beta = \frac{\log 2}{2\alpha}$.

(iii) Consider the system $([0, 1], \mathcal{F}, \mu)$ given by the Farey map $\mathcal{F}(x) = \min\{\frac{x}{1-x}, \frac{1-x}{x}\}$. Then we have

$$(2.6) \quad 2\alpha = \int_0^1 \log |\mathcal{F}'(x)| d\mu(x);$$

in this form the constant 2α has an interpretation as a relative metric entropy.

(iv) The constant α is the maximal Lyapunov exponent λ_1 of the measure η on $GL_2(\mathbb{R})$ defined by $\eta = \frac{1}{2}\delta_{A_0} + \frac{1}{2}\delta_{A_1}$, where $A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. An η -stationary measure ξ on $\mathbb{P}(\mathbb{R}^2) = \mathbb{R} \cup \{\infty\}$ is given by

$$d\xi(x) = \frac{1}{2}\mu + \frac{1}{2}T^*\mu \quad (T(x) = 1/x),$$

where μ is extended trivially to a measure on $\mathbb{R} \cup \{\infty\}$ with support in $[0, 1]$.

(v) The variance σ^2 can be expressed as

$$(2.7) \quad \sigma^2 = \frac{1}{2} \int_0^1 \left(\log x + \alpha \lfloor \frac{1}{x} \rfloor + \int_0^1 \log \left(\frac{1 + y \lfloor \frac{1}{x} \rfloor}{1 + yx} \right) d\mu(y) \right)^2 d\mu(x).$$

Remark. The above facts and comments are detailed and proven in Section 7, where we will also show that the results of [Kin60] can be easily deduced from Theorem 1.1. The main point in (iv) is the explicit expression for the stationary measure, which allows one to compute α using a formula of Furstenberg.

Bacher [Bac17] showed that the moments m_k can be computed very accurately, allowing him to obtain a 50-digits approximation for α .

2.2. Iterates of the binary map, and a heuristic toward Theorem 1.1. In this section we derive a formula for $s(n)$ in terms of the function Φ , and then state a naive heuristic towards the fact that $N^{-1} \log \mathbf{S}_N \rightarrow \alpha$ in law as $N \rightarrow \infty$.

Consider the function Φ defined on the dyadic rationals of $[0, 1]$ by the formula

$$(2.8) \quad \Phi\left(\frac{m}{2^N}\right) = \frac{s(m)}{s(m + 2^N)}, \quad \left(\frac{m}{2^N} \in [0, 1]\right).$$

As we have mentioned earlier, the function Φ is given by the inversion bijection of Minkowski's function; but let us ignore this for a moment. The recurrence relations (1.1) are equivalent to the facts that:

- (1) The definition (2.8) is well-posed, *i.e.*, it genuinely only depends on the ratio $m/2^N$,
- (2) Whenever $N \geq 0$ and $0 \leq m < 2^N$, we have

$$(2.9) \quad \Phi\left(\frac{1}{2}\left(\frac{m}{2^N} + \frac{m+1}{2^N}\right)\right) = \frac{s(m) + s(m+1)}{s(m + 2^N) + s(m+1 + 2^N)}.$$

Additionally, we recall the fact that the Farey tree described in Figure 2 enumerates all rational numbers on $[0, 1]$. Since, by (2.9), the function Φ is increasing, and $(\Phi(0), \Phi(1)) = (0, 1)$, we obtain that Φ may be extended to a homeomorphism from $[0, 1]$ to itself. By these facts, a result of Panti [Pan08, Proposition 1.1] guarantees that Φ must be the inverse of the Minkowski function.

We now iterate the relation (2.8), in order to express $s(2^N + m)$ in terms of values of Φ . Given $1 \leq m < 2^N$, there is a unique $N' \in \mathbb{N}_{\geq 0}$, with $N' < N$, such that $2^{N'} \leq m < 2^{N'+1}$. Write $m = 2^{N'+1} - m'$, so that $1 \leq m' \leq 2^{N'}$. It was noted by Stern [Bac17] (see [Leh29] for an account) that $s(m) = s(2^{N'+1} - m') = s(2^{N'} + m')$. Therefore, by (2.8), we deduce

$$(2.10) \quad s(2^N + m) = \Phi\left(\frac{m}{2^N}\right)^{-1} s(2^{N'} + m').$$

It is easily checked that the map $\mathcal{B}_\triangleright : m/2^N \mapsto m'/2^{N'}$ is given by

$$(2.11) \quad \mathcal{B}_\triangleright(x) = \begin{cases} 2 - 2^k x & \text{for } x \in [2^{-k}, 2^{-k+1}), k \geq 1, \\ 0 & \text{for } x \in \{0, 1\}. \end{cases}$$

This map is shown on the right, in Figure 4b. It is the jump transformation of the binary (or “tent”) map $\mathcal{B}(x) = 2 \min(x, 1 - x)$ on the interval $[\frac{1}{2}, 1]$ (see [KMS16], fig. 1.8, page 46). The map \mathcal{B} is shown in Figure 4a.

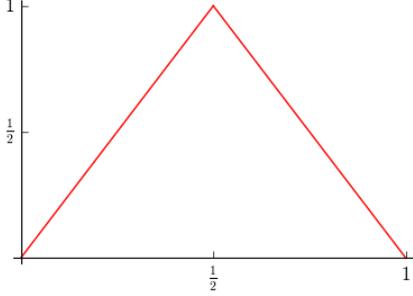


FIGURE 4A. Graph of the binary map $x \mapsto \mathcal{B}(x)$.

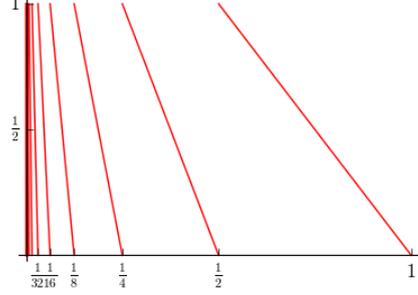


FIGURE 4B. Graph of the jump transformation $x \mapsto \mathcal{B}_\triangleright(x)$.

Therefore, for any dyadic number $m/2^N \in (0, 1]$, if we denote by $K = K(m/2^N) \geq 0$ the least integer such that $\mathcal{B}_\triangleright^K(m/2^N) = 1$, then by iterating (2.10), we obtain

$$(2.12) \quad s(2^N + m) = \left(\Phi\left(\frac{m}{2^N}\right) \Phi\left(\mathcal{B}_\triangleright\left(\frac{m}{2^N}\right)\right) \cdots \Phi\left(\mathcal{B}_\triangleright^{K-1}\left(\frac{m}{2^N}\right)\right) \right)^{-1}.$$

The formula (2.12) has the advantage that the points $m/2^N$ are easily described, however, the right-hand side involves the function Φ . At this point we may use the property that the function Φ conjugates the binary map with another simple map: by Proposition 1.1 of [Pan08], we have the conjugacy relation

$$\mathcal{B} = \Phi^{-1} \circ \mathcal{F} \circ \Phi$$

between the binary map \mathcal{B} , which we have already mentioned, and the *Farey map* on the interval $[0, 1]$, defined by

$$(2.13) \quad \mathcal{F}(x) = \min\left(\frac{x}{1-x}, \frac{1-x}{x}\right).$$

By induction, we obtain the relation

$$(2.14) \quad \mathcal{B}_\triangleright = \Phi^{-1} \circ \mathcal{F}_\triangleright \circ \Phi$$

between the map $\mathcal{B}_\triangleright$, defined at (2.11), and the “jump transformation” of the Farey map on the interval $[\frac{1}{2}, 1]$, which is known (see *e.g.* [PS92, Iso02]) to be precisely the *Gauss map*

$$\mathcal{F}_\triangleright(x) = \mathcal{G}(x) := \begin{cases} 1/x - n & \text{for } x \in [\frac{1}{n+1}, \frac{1}{n}), n \geq 1, \\ 0 & \text{for } x \in \{0, 1\}. \end{cases}$$

We mention at this point that the Lebesgue measure is invariant for the binary maps \mathcal{B} and $\mathcal{B}_\triangleright$. Upon conjugating, we deduce that the Minkowski measure μ is invariant for the Farey and the Gauss maps; this fact will be used repeatedly in our arguments. Moreover, the measure μ has maximal entropy $\log 2$ for the Farey map; by contrast, the Farey map also admits dx/x as a unique invariant measure absolutely continuous with respect to the Lebesgue measure, but it is not finite, and has entropy zero. We refer to [Lag92] and to Chapter 1.2 of [KMS16] for more details.

By (2.14), the relation (2.12) becomes the following.

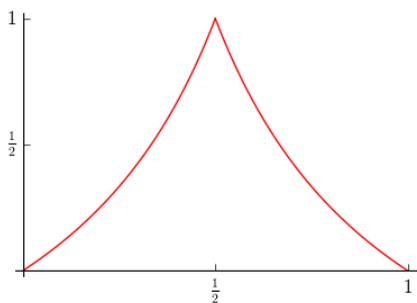


FIGURE 5A. Graph of the Farey map $x \mapsto \mathcal{F}(x)$.

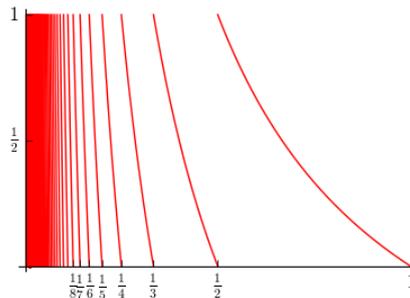


FIGURE 5B. Graph of the Gauss map $x \mapsto \mathcal{G}(x) = \mathcal{F}_{>}(x)$.

Proposition 2.2. *Let $N \geq 0$, $1 \leq m \leq 2^N$ and denote by $K = K(m/2^N) \geq 0$ the least integer such that $(\mathcal{G}^K \circ \Phi)(m/2^N) = 1$. Then*

$$(2.15) \quad s(2^N + m) = \left(\Phi\left(\frac{m}{2^N}\right) \mathcal{G}\left(\Phi\left(\frac{m}{2^N}\right)\right) \cdots \mathcal{G}^{K-1}\left(\Phi\left(\frac{m}{2^N}\right)\right) \right)^{-1}.$$

Equivalently, we have

$$\log s(2^N + m) = - \sum_{i=0}^{K(m/2^N)-1} \log \left((\mathcal{G}^i \circ \Phi)\left(\frac{m}{2^N}\right) \right)$$

Our situation at this point is formally similar to the work of Baladi and Vallée [BV05], concerned with statistical properties of orbits of rationals under the Gauss map. Our “cost function” here is $x \mapsto -\log x$. This analogy provides a heuristic explanation for Theorem 1.1 as follows.

It is easily seen that for $r \geq 0$ and $N \geq 0$, $|\{m \in \{1, \dots, 2^N\} : K(\frac{m}{2^N}) = r\}| = \binom{N}{r}$. It follows that, as $N \rightarrow \infty$, we have

$$K(m/2^N) \sim \frac{N}{2}$$

for a proportion $1 + o(1)$ of integers $m \in \{1, \dots, 2^N\}$ as $N \rightarrow \infty$ (which we abbreviate by “generic m ”).

On the other hand, by Theorem 5.12 of [MR17], we have that for a generic $m \in \{1, \dots, 2^N\}$, the pre-images $(\mathcal{B}^i(\frac{m}{2^N}))_{i=0}^N$ by the binary map will equidistribute as $N \rightarrow \infty$ according to the Lebesgue measure on $[0, 1]$. We may then guess that the same is true for the pre-images $(\mathcal{B}_{>}^i(\frac{m}{2^N}))_{i=0}^{K(m/2^N)}$ by the *jump transformation* $\mathcal{B}_{>}$. After conjugation by Φ , the Lebesgue measure is sent to the Minkowski measure $\mu = d(\Phi^{-1})$: in particular, we are led to expect that for generic $m \in \{1, \dots, 2^N\}$, we have

$$\frac{1}{K(m/2^N)} \sum_{i=0}^{K(m/2^N)-1} \log \left((\mathcal{G}^i \circ \Phi)\left(\frac{m}{2^N}\right) \right) \sim \int_0^1 \log x \, d\mu(x),$$

and therefore for generic $m \in \{1, \dots, 2^N\}$,

$$\log s(2^N + m) \sim -\frac{N}{2} \int_0^1 \log x \, d\mu(x).$$

This is indeed a consequence of Theorem 1.1 by our definition (2.4).

This guess, which might seem naive at this stage, echoes a similar phenomenon for real, *resp.* rational trajectories under the Gauss map [Hen94, BV05]: see in particular the parallel between Theorems 1 and 3 of [BV05], and the factorization “ $\mu(c) = \mu \times \hat{\mu}(c)$ ” stated there on page

350. For us, “ μ ” plays the role the typical ratio of the length of rational trajectories $K(m/2^N)$ by N ; and “ $\hat{\mu}(c)$ ” is the generic average of the cost function over real trajectories.

2.3. Outline of the proof. Having at hand the expression (2.2) for the Stern sequence in terms of iterates of the Gauss map, the first step towards the proof of Theorem 1.2 is to construct a generating series for the moment on the left-hand side of (1.4). The precise form has to be amenable to analytic tools. For $(\tau, z) \in \mathbb{C}^2$ of small enough moduli, let

$$(2.16) \quad \begin{aligned} S_\tau(z) &= \sum_{N \geq 0} (2z)^N \mathbb{E}_N[(\mathbf{S}_N)^{-\tau}] \\ &= \sum_{N \geq 0} z^N \sum_{n \in \mathcal{I}_N} s(n)^{-\tau}. \end{aligned}$$

The framework of analytic combinatorics [FS09] relates the properties of $S_\tau(z)$ (analytic continuation, meromorphy, location of poles, spectral gap) with the asymptotic behavior of $\mathbb{E}_N[(\mathbf{S}_N)^{-\tau}]$.

The next step is to obtain the analytic information required on $S_\tau(z)$. To this end, we adapt methods of “dynamical analysis”, introduced by Vallée and described for instance in [Val03, BV05]. The main point, which is behind the choice of the generating function $S_\tau(z)$, is the expression

$$(2.17) \quad S_\tau(z) = \frac{1}{1-z} \sum_{k \geq 0} \left(\mathbb{H}_{\tau,z}^k[\mathbf{1}] \right) (1).$$

This involves the iterates of an operator $\mathbb{H}_{\tau,z}$, which acts on bounded functions $f : [0, 1] \rightarrow \mathbb{C}$ by

$$\mathbb{H}_{\tau,z}[f] : t \mapsto \sum_{n \geq 0} \frac{z^n}{(n+t)^\tau} f\left(\frac{1}{n+t}\right).$$

The operator $\mathbb{H}_{\tau,z}$ is a particular case of a weighted Ruelle-Perron-Frobenius transfer operator [Rue94, Bal00]. Our situation is similar to the work of Baladi-Vallée [BV05] on the Gaussian behavior of Euclidean algorithms. The pole of smallest modulus of $S_\tau(z)$ will occur at a point $z = \rho(\tau)$, where $\mathbb{H}_{\tau,z}$ has dominant eigenvalue 1, and the Cauchy formula transfers this information into the estimate (1.4) with $U(\tau) = -\log(2\rho(-\tau))$.

The deduction of Theorem 1.1 from Theorem 1.2, which is standard in probability theory, will be made by appealing to Hwang’s Quasi-Powers theorem [Hwa96]. We remark that the uniformity of Theorem 1.2 is crucial in this deduction.

As in [Mor15], we have chosen to include details of the arguments from spectral theory, rather than quote them as a black box, so that readers unacquainted with these topics, but who are still interested in the arithmetic application, may follow through.

The plan is the following: in Section 3, we derive the expression (2.17). In Section 4, we study the operator $\mathbb{H}_{\tau,z}$, with a particular emphasis on the reference pair $(\tau, z) = (0, \frac{1}{2})$. In Section 5, we carry out the analysis of $S_\tau(z)$. In Section 6, we complete the proof of Theorems 1.1 and 1.2. Finally in Section 7, we return to the mean-value α and the variance σ^2 , and prove Theorem 2.1.

3. EXPRESSING THE MOMENT-GENERATING FUNCTION

In this section we express the moment-generating function on the left-hand side of (1.4) in terms of a weighted transfer operator for the Gauss map, continuing the arguments of Section 2.2.

We recall Proposition 2.2. Note that the Gauss map has the property that $\mathcal{G}'(x) = -1/x^2$ for $x > 0$. It is therefore very convenient to use it in conjunction with (2.15) to obtain the

“product of cocycles”

$$(3.1) \quad \begin{aligned} s(2^N + m)^2 &= \prod_{i=0}^{K-1} |\mathcal{G}'| \circ \mathcal{G}^i \left(\Phi \left(\frac{m}{2^N} \right) \right) \\ &= |(\mathcal{G}^K)'| \left(\Phi \left(\frac{m}{2^N} \right) \right). \end{aligned}$$

The inverse branches of the Gauss map \mathcal{G} form the set

$$\mathcal{H} = \{h_n, n \geq 1\}, \quad h_n(x) = \frac{1}{n+x} \quad (x \in [0, 1]).$$

For any given $K \in \mathbb{N}_{\geq 0}$, the set \mathcal{H}^K of inverse branches of the function $\mathcal{G}^K : [0, 1] \rightarrow [0, 1]$ are then given by

$$\mathcal{H}^K = \{h_{n_1} \circ \cdots \circ h_{n_K} : n_j \geq 1\},$$

where it is understood that $\mathcal{H}^0 = \{\text{id}\}$. The decomposition $h = h_{n_1} \circ \cdots \circ h_{n_K}$ is unique, therefore, it makes sense to define

$$w(h) := n_1 + \cdots + n_K.$$

An immediate verification shows that

$$\left\{ \frac{m}{2^N} \mid 0 \leq m \leq 2^N, m \text{ odd} \right\} = \mathcal{B}^{-N}(\{1\})$$

and thus, letting $\mathcal{H}^* = \cup_{K \geq 0} \mathcal{H}^K$,

$$(3.2) \quad \begin{aligned} \left\{ \Phi \left(\frac{m}{2^N} \right) \mid 0 \leq m \leq 2^N, m \text{ odd} \right\} &= \mathcal{F}^{-N}(\{1\}) \\ &= \{h(1) \mid h \in \mathcal{H}^*, w(h) = N\}. \end{aligned}$$

Gathering the above, we obtain

$$(3.3) \quad \begin{aligned} S_\tau(z) &= \sum_{N \geq 0} z^N \sum_{m=1}^{2^N} s(m + 2^N)^{-\tau} \\ &= \sum_{N \geq 0} z^N \sum_{0 \leq r \leq N} \sum_{\substack{m=1 \\ m \text{ odd}}}^{2^r} s(m + 2^r)^{-\tau} \\ &= \frac{1}{1-z} \sum_{r \geq 0} z^r \sum_{K \geq 0} \sum_{\substack{h \in \mathcal{H}^K \\ w(h)=r}} |(\mathcal{G}^K)' \circ h(1)|^{-\tau/2} \\ &= \frac{1}{1-z} \sum_{h \in \mathcal{H}^*} z^{w(h)} |h'(1)|^{\tau/2}, \end{aligned}$$

where we have extracted the largest power of 2 dividing m in the first line; the third line used (3.1) and (3.2), and the last line followed by the derivative formula for the inverse.

The sum over $h \in \mathcal{H}^*$ is now recognized as a sum of iterates of a “density transformer”, or transfer operator, also called Ruelle-Perron-Frobenius operator. These objects, and their spectral properties, have a long history and have been extensively studied, notably in connection with continued fraction, and thermodynamic formalism. We refer to the surveys [Rue94], the lecture notes [May91] and the monographs [Bal00, Rue04]. More explanations and references relevant to our case are found in Section 2.2 of [BV05]. For all $(\tau, z) \in \mathbb{C}^2$ with $|z| < 1$, let the operator $\mathbb{H}_{\tau, z}$ act on continuous functions $f \in \mathcal{C}([0, 1])$ by

$$(3.4) \quad \begin{aligned} \mathbb{H}_{\tau, z}[f](t) &= \sum_{h \in \mathcal{H}} z^n |h'(t)|^{\tau/2} (f \circ h)(t) \\ &= \sum_{n \geq 1} \frac{z^n}{(n+t)^\tau} f \left(\frac{1}{n+t} \right). \end{aligned}$$

Then from formula (2.7) of [BV05], we have that for all $K \geq 0$

$$(3.5) \quad \mathbb{H}_{\tau,z}^K[f](t) = \sum_{h \in \mathcal{H}^K} z^{w(h)} |h'(t)|^{\tau/2} (f \circ h)(t).$$

Let now

$$(3.6) \quad \mathcal{W} = \{(\tau, z) \in \mathbb{C}^2 : |z| < \min(1, 2^{-\operatorname{Re} \tau})\}.$$

By equations (3.3), (3.4) and (3.5) above, we conclude the following.

Proposition 3.1. *For $(\tau, z) \in \mathcal{W}$, we have*

$$S_\tau(z) = \frac{1}{1-z} \sum_{k \geq 0} \mathbb{H}_{\tau,z}^k[\mathbf{1}](1) = \frac{1}{1-z} (\operatorname{Id} - \mathbb{H}_{\tau,z})^{-1}[\mathbf{1}](1).$$

4. PROPERTIES OF THE TRANSFER OPERATOR

As we have mentioned, spectral properties of transfer operators have been extensively studied; we refer again to the lecture notes [May91], sections 7.1 and 7.4, and the references therein. The actual operator $\mathbb{H}_{\tau,z}$ has been defined and studied at many occurrences in the literature: see [PS92, formula (10)], [Dod17], [Iso02, formula (3.39)]. Indeed, many of the forthcoming properties of $\mathbb{H}_{\tau,z}$ for real z can be found in [PS92]. Nonetheless, to make the arguments as clear as possible for readers unacquainted with these topics, we will provide full proofs, following the presentation of [Mor15], apart from perturbation theory of operators, which we will quote from [Kat95].

4.1. Definitions. A critical first step is to define an appropriate functional space in which to study $\mathbb{H}_{\tau,z}$. For our arithmetic application, two constraints must be satisfied: the interval $[0, 1]$ should be contained in their domain, and it must include the constant function $\mathbf{1}$ and all its iterates under $\mathbb{H}_{\tau,z}$ for all (τ, z) in a neighborhood of the origin.

Consider the domain $\mathbb{D} = \{t \in \mathbb{C} : |t - \frac{2}{3}| < 1\}$, and the set of functions

$$H^\infty(\mathbb{D}) = \{f : \mathbb{D} \rightarrow \mathbb{C} : f \text{ is holomorphic and bounded}\}.$$

For $(\tau, z) \in \mathcal{W}$ and $f \in H^\infty(\mathbb{D})$, we define $\mathbb{H}_{\tau,z}$ as in (3.4), taking the principal determination of the logarithm.

For \mathbb{T} an operator, we denote by $\operatorname{srd}(\mathbb{T})$ the spectral radius of \mathbb{T} . We further recall the definition of the Gauss map and its inverse branches,

$$\mathcal{G}(x) := \begin{cases} 1/x - n, & x \in [\frac{1}{n+1}, \frac{1}{n}), \\ 0, & x \in \{0, 1\}, \end{cases} \quad h_n(x) := \frac{1}{n+x}.$$

4.2. Decomposition of the transfer operator. In this section, we obtain the basic properties of the operator $\mathbb{H}_{\tau,z}$. The arguments involved have a long history [May91]; we will mostly follow the presentation found in [Mor15], which is well adapted to our setting. We are interested in the behavior of $\mathbb{E}_N[(\mathbf{S}_N)^\tau]$ when τ is in a neighborhood of $\tau = 0$. At this value the power series S_τ has radius of convergence $\frac{1}{2}$, and thus of particular importance for us is the value $(\tau, z) = (0, \frac{1}{2})$, around which $\mathbb{H}_{\tau,z}$ will have 1 as an eigenvalue. We denote

$$\mathbb{H} := \mathbb{H}_{0, \frac{1}{2}}.$$

We will first focus on \mathbb{H} , then on the case $\tau = 0$, $|z| = \frac{1}{2}$, and finally we study arbitrary (τ, z) using the theory of analytic perturbation (as in [Kat95, Chapter IV]).

Proposition 4.1. *(i) For $|z| < 1$ and $\tau \in \mathbb{C}$, the operator $\mathbb{H}_{\tau,z}$ is compact and depends holomorphically on (τ, z) in the sense of [Kat95, p.366].*

(ii) The eigenmeasure of the adjoint operator \mathbb{H}^* is μ , that is, for all $f \in L^1(\mu)$ we have $\mathbb{H}[f] \in L^1(\mu)$ and

$$(4.1) \quad \int \mathbb{H}[f] d\mu = \int f d\mu.$$

Moreover, the operator \mathbb{H} has a simple isolated eigenvalue equal to 1; the corresponding eigenspace is generated by the constant function $\mathbf{1}$. Finally, this eigenvalue is the only element of modulus 1 in the spectrum of \mathbb{H} .

Furthermore, for all small enough $\delta_1 > 0$, we may find $\delta_2 > 0$ and $\theta \in [0, 1)$ such that, writing

$$\begin{aligned} \mathcal{V}_1 &:= \{(\tau, z) \in \mathbb{C}^2 : |\tau| \leq \delta_1, |z - \frac{1}{2}| \leq \delta_1\}, \\ \mathcal{V}_2 &:= \{(\tau, z) \in \mathbb{C}^2 : |\tau| \leq \delta_2, |z| \leq \frac{1}{2} + \delta_2, |z - \frac{1}{2}| \geq \delta_1\}, \end{aligned}$$

the following holds.

(iii) For $(\tau, z) \in \mathcal{V}_1$, $\mathbb{H}_{\tau, z}$ has a simple and isolated dominant eigenvalue at $\lambda(\tau, z) \in \mathbb{C}$ with eigenfunction $f_{\tau, z} \in H^\infty(\mathbb{D}) \setminus \{0\}$; also, $\lambda(0, \frac{1}{2}) = 1$ and $f_{0, \frac{1}{2}} = \mathbf{1}$. In particular, we have the decomposition

$$(4.2) \quad \mathbb{H}_{\tau, z} = \lambda(\tau, z)\mathbb{P}_{\tau, z} + \mathbb{N}_{\tau, z},$$

with $\mathbb{P}_{\tau, z}$ and $\mathbb{N}_{\tau, z}$ two compact operators on $H^\infty(\mathbb{D})$. Moreover, the image of $\mathbb{P}_{\tau, z}$ is one-dimensional: $\text{Im } \mathbb{P}_{\tau, z} = f_{\tau, z}\mathbb{C}$ and $f_{\tau, z} = \mathbb{P}_{\tau, z}[\mathbf{1}]$. Finally, we have $\text{srd}(\mathbb{N}_{\tau, z}) \leq \theta$ and $\mathbb{P}_{\tau, z}\mathbb{N}_{\tau, z} = \mathbb{N}_{\tau, z}\mathbb{P}_{\tau, z} = \mathbf{0}$.

(iv) For $(\tau, z) \in \mathcal{V}_1$, $f_{\tau, z}$ and $\lambda(\tau, z)$ depend holomorphically on both variables.

(v) For $(\tau, z) \in \mathcal{V}_2$, we have $\text{srd}(\mathbb{H}_{\tau, z}) \leq \theta$.

Proof. (i) Let $\tilde{\mathbb{D}} := \{t \in \mathbb{C} : |t - \frac{2}{3}| < \frac{17}{16}\}$. We have that h_n maps $\tilde{\mathbb{D}}$ into \mathbb{D} for all $n \geq 1$. In particular, all the elements of the sets $S := \{f \circ h_n \mid f \in H^\infty(\mathbb{D}), \|f\|_\infty \leq 1\}$ can be extended to functions in the unit disk of $H^\infty(\tilde{\mathbb{D}})$. By Montel's theorem the set obtained by these extended functions is pre-compact with respect to the compact-open topology on $H^\infty(\tilde{\mathbb{D}})$ and thus S is pre-compact with respect to the uniform topology of $H^\infty(\mathbb{D})$. In particular, the operator $f \mapsto f \circ h_n$ on $H^\infty(\mathbb{D})$ is compact and thus $f \mapsto \frac{z^n}{(n+)^{\tau}} f(\frac{1}{n+})$ is also compact (and holomorphic in (τ, z)). The same then holds for $\mathbb{H}_{\tau, z}$ since its defining series converges locally uniformly in (τ, z) .

(ii) One has that $\mathbb{H}[f](x) = \mathbb{U}[f \circ \Phi](\Phi^{-1}(x))$, where \mathbb{U} is the Ruelle operator relative to the binary map $\mathcal{B}_>$ defined at (2.11); since the Lebesgue measure is an eigenmeasure for \mathbb{U}^* [KMS16, Prop. 2.3.21] with eigenvalue 1, it follows that μ is an eigenmeasure for \mathbb{H}^* with the same eigenvalue. One could also verify this directly from the definition (2.2). Indeed, since $\Psi(h_n(x)) = 2^{-n}(2 - \Psi(x))$ for all $n \in \mathbb{N}^*$ then one has $\mu(A) = 2^{-n}\mu(h_n^{-1}(A))$ for any interval $A = [\frac{1}{n+1}, \frac{1}{n+x}]$, with $x \in [0, 1]$. By the monotone class lemma one also has $\mu(A \cap [\frac{1}{n+1}, \frac{1}{n}]) = 2^{-n}\mu(h_n^{-1}(A))$ for all μ -measurable set A , and finally $2^{-n} \int_{[0, 1]} (f \circ h_n) d\mu = \int_{[1/(n+1), 1/n]} f d\mu$ for all $f \in L^1(\mu)$. Summing over $n \geq 1$ yields (4.1).

Since \mathbb{H} is a compact operator, the non-zero elements of its spectrum are isolated eigenvalues of finite multiplicity; also, notice that $\mathbb{H}[\mathbf{1}] = \mathbf{1}$. Let $f \in H^\infty(\mathbb{D})$ be an eigenfunction with eigenvalue λ normalized so that $\max_{t \in [0, 1]} |f(t)| = |f(t_0)| = 1$ for some $t_0 \in [0, 1]$. Then, by the definition of $\mathbb{H}[f]$ and the triangle inequality we have

$$|\lambda| = |\mathbb{H}[f](t_0)| = \left| \sum_{n \geq 1} 2^{-n} f\left(\frac{1}{n+t_0}\right) \right| \leq \sum_{n \geq 1} 2^{-n} = 1.$$

Thus, $|\lambda| \leq 1$. Also, if $|\lambda| = 1$ then the equality holds everywhere and so there exists $c \in \mathbb{C}$ of modulus 1 such that $f(\frac{1}{n+t_0}) = c$ for all $n \geq 1$. Since f is holomorphic and $\frac{1}{n+t_0}$ has $0 \in \mathbb{D}$ as an accumulation point, we must have $f = c\mathbf{1}$. Thus, $\lambda = 1$ is the only eigenvalue of modulus 1 and $\ker(\mathbb{H} - \text{Id})$ is 1-dimensional. Moreover, if $(\mathbb{H} - \text{Id})^2[f] = \mathbf{0}$ then $(\mathbb{H} - \text{Id})[f]$ is

an eigenfunction of \mathbb{H} with eigenvalue 1 and so $(\mathbb{H} - \text{Id})[f] = c\mathbf{1}$ for some $c \in \mathbb{C}$. Integrating this equation with respect to $d\mu$ by (4.1) we find $c = 0$. Thus, f is itself a multiple of $\mathbf{1}$ and so 1 is a simple isolated eigenvalue of \mathbb{H} .

(iii) Now, let C be a small circle centered at 1 which doesn't enclose any other eigenvalue of \mathbb{H} and assume $(\tau, z) \in \mathcal{V}_1$. If δ_1 is small enough, then by [Kat95, Thm IV.3.16] we have that C doesn't intersect the spectrum of $\text{srd}(\mathbb{H}_{\tau,z})$. It follows that we can write $\mathbb{H}_{\tau,z}$ as a sum of compact operators $\mathbb{H}_{\tau,z} = \mathbb{P}_{\tau,z} + \mathbb{N}_{\tau,z}$ with $\mathbb{P}_{\tau,z}\mathbb{N}_{\tau,z} = \mathbb{N}_{\tau,z}\mathbb{P}_{\tau,z} = 0$ and $\mathbb{P}_{\tau,z}^2 = \mathbb{P}_{\tau,z}$, where $\mathbb{P}_{\tau,z}$ is the Riesz projection associated to C [Kat95, Thm III.6.17]. Moreover, since 1 is a simple eigenvalue, the image of $\mathbb{P}_{0,\frac{1}{2}}$ is one dimensional [Kat95, pp.180-181] and thus the same holds for $\mathbb{P}_{\tau,z}$ if δ_1 is small enough [Kat95, Thm IV.3.16]. Also, the spectrum of $\mathbb{H}_{\tau,z}$ restricted to the image of $\mathbb{P}_{\tau,z}$ consists of a unique eigenvalue $\lambda(\tau, z)$, with $\lambda(0, \frac{1}{2}) = 1$, corresponding to the eigenfunction $f_{\tau,z} := \mathbb{P}_{\tau,z}\mathbf{1}$, whereas the spectrum of $\mathbb{N}_{\tau,z}$ consists of that of $\mathbb{H}_{\tau,z}$ with $\lambda(\tau, z)$ removed [Kat95, Thm IV.3.16]. In particular since $\text{srd}(\mathbb{N}_{0,\frac{1}{2}}) \leq \theta' < 1$ for some $\theta' \in [0, 1)$, then by the upper-semicontinuity of the spectral radius [Kat95, Thm IV.3.16] there exists $\theta \in [0, 1)$ such that $\text{srd}(\mathbb{N}_{\tau,z}) \leq \theta$ for all $(\tau, z) \in \mathcal{V}_1$ with δ_1 small enough.

(iv) By [Kat95, Thm VII.1.7] $\mathbb{P}_{\tau,z}$ and $\mathbb{N}_{\tau,z}$ depend homomorphically on (τ, z) and thus so does $f_{\tau,z} = \mathbb{P}_{\tau,z}[\mathbf{1}]$. Moreover, since $\mathbb{P}_{0,\frac{1}{2}}[\mathbf{1}](0) = 1$ we have $f_{\tau,z}(0) \neq 0$ for $(\tau, z) \in \mathcal{V}_1$ and δ_1 small enough. Thus the holomorphicity of $\lambda(\tau, z)$ follows since $\mathbb{P}_{\tau,z}[f_{\tau,z}](0) = \lambda(\tau, z)f_{\tau,z}(0)$.

(v) First, consider the case $\tau = 0$ and $z = \frac{1}{2}e^{2\pi i\phi}$ with $\phi \in [0, 1)$. By the triangle inequality we have $\|\mathbb{H}_{0,z}[f]\|_\infty \leq \|f\|_\infty$ for all $f \in H^\infty(\mathbb{D})$; in particular $\text{srd}(\mathbb{H}_{0,z}) \leq 1$ for $|z| = \frac{1}{2}$. Suppose now $\phi \neq 0$. Then since $\mathbb{H}_{0,z}$ is compact, it has an eigenfunction $f \in H^\infty(\mathbb{D})$ with eigenvalue λ of maximum modulus $\text{srd}(\mathbb{H}_{0,z})$. Up to re-scaling f , we can assume $\max_{t \in [0,1]} |f(t)| = |f(t_0)| = 1$ for some $t_0 \in [0, 1]$, whence

$$|\lambda| = |\mathbb{H}_{0,z}[f](t_0)| = \left| \sum_{n \geq 1} 2^{-n} e^{2\pi i n \phi} f\left(\frac{1}{n+t_0}\right) \right| \leq \sum_{n \geq 1} 2^{-n} = 1.$$

If we had $|\lambda| = 1$, then this would mean equality holds everywhere. This implies that all the summands in the first series have the same argument and the n -summand has modulus 2^{-n} , that is

$$e^{2\pi i n \phi} f\left(\frac{1}{n+t_0}\right) = e^{2\pi i \alpha}$$

for some $\alpha \in [0, 1)$. Letting $n \rightarrow \infty$, we deduce $e^{2\pi i n \phi} \rightarrow f(0)$, which implies $\phi = 0$. We conclude that $\text{srd}(\mathbb{H}_{0,z}) < 1$ whenever $|z| = \frac{1}{2}$ and $z \neq \frac{1}{2}$. Then, by the upper semi-continuity of $\text{srd}(\mathbb{H}_{\tau,z})$ [Kat95, Thm IV.3.16], we have that there exists $0 < \delta_2 < \frac{1}{2}$ such that for τ, z satisfying

$$\frac{1}{2} - \delta_2 \leq |z| \leq \frac{1}{2} + \delta_2, \quad |z - \frac{1}{2}| \geq \delta_1, \quad |\tau| \leq \delta_2$$

we also have $\text{srd}(\mathbb{H}_{\tau,z}) \leq \theta < 1$. Finally, assume $|z| \leq \frac{1}{2} - \delta_2$ and $|\tau| \leq \delta_2 < -\frac{1}{2} \log_2(1 - 2\delta_2)$. Then, by the triangle inequality for any $f \in H^\infty(\mathbb{D})$, we have

$$\|\mathbb{H}_{\tau,z}[f]\|_\infty \leq \|f\|_\infty \sum_{n \geq 1} |z|^n (n+1)^{\delta_2}.$$

Since $n+1 \leq 2^n$, computing the series proves that $\text{srd}(\mathbb{H}_{\tau,z}) < 1$. In particular, $\text{srd}(\mathbb{H}_{\tau,z}) < 1$ for all $(\tau, z) \in \mathcal{V}_2$ and (iv) follows. \square

4.3. Specific properties at $(\tau, z) = (0, \frac{1}{2})$. We have the following properties linking \mathbb{H} and μ . We recall that the value $\lambda(0, \frac{1}{2}) = 1$ was proved at Proposition 4.1.(ii).

Proposition 4.2. (i) We have $\mathbb{P}_{0,\frac{1}{2}}[f] = (\int f d\mu)\mathbf{1}$ for all $f \in H^\infty(\mathbb{D})$.

(ii) The derivatives of λ satisfy

$$\frac{\partial}{\partial \tau} \lambda(0, \frac{1}{2}) = \int_{[0,1]} \log d\mu, \quad \frac{\partial}{\partial z} \lambda(0, \frac{1}{2}) = 4.$$

Proof. (i) First we recall that $f_{0, \frac{1}{2}} = \mathbf{1}$ and thus $\text{Im } \mathbb{P}_{0, \frac{1}{2}} = \mathbb{C}\mathbf{1}$. Given $f \in H^\infty(\mathbb{D})$, we have $f \in L^1(\mu)$ and so, by (4.1), $\int f d\mu = \int \mathbb{H}^k[f] d\mu$ for all $k \in \mathbb{N}$. Letting $k \rightarrow \infty$, from the decomposition (4.2) we have $\mathbb{H}^k[f] \rightarrow \mathbb{P}_{0, \frac{1}{2}}[f]$ uniformly on $[0, 1]$, and so $\int \mathbb{H}^k[f] d\mu \rightarrow \int \mathbb{P}_{0, \frac{1}{2}}[f] d\mu$. Since $\mathbb{P}_{0, \frac{1}{2}}[f] \in \mathbb{C}\mathbf{1}$ is constant, this proves our claim.

(ii) We have, following the notations of Proposition 4.1, $\lambda(\tau, z)f_{\tau, z} = \mathbb{H}_{\tau, z}[f_{\tau, z}]$ for (τ, z) in a neighborhood of $(0, \frac{1}{2})$. By uniform convergence of the series defining $\mathbb{H}_{\tau, z}$, we may differentiate term-wise and obtain

$$\begin{cases} \frac{\partial}{\partial \tau} \lambda(\tau, z)f_{\tau, z} + \lambda(\tau, z) \frac{\partial}{\partial \tau} f_{\tau, z} = \mathbb{H}_{\tau, z}[f_{\tau, z} \log + \frac{\partial}{\partial \tau} f_{\tau, z}], \\ \frac{\partial}{\partial z} \lambda(\tau, z)f_{\tau, z} + \lambda(\tau, z) \frac{\partial}{\partial z} f_{\tau, z} = \sum_{n \geq 1} \frac{nz^{n-1}}{(n + \cdot)^\tau} f_{\tau, z} \circ h_n + \mathbb{H}_{\tau, z}[\frac{\partial}{\partial z} f_{\tau, z}]. \end{cases}$$

We evaluate each line at $(\tau, z) = (0, \frac{1}{2})$ and apply the operator $\mathbb{P}_{0, \frac{1}{2}}$, which amounts to integrating against $d\mu$ by point (i). We obtain the claimed identities $\frac{\partial}{\partial \tau} \lambda(0, \frac{1}{2}) = \int \log d\mu$ and $\frac{\partial}{\partial z} \lambda(0, \frac{1}{2}) = \sum_{n \geq 1} n2^{1-n} = 4$. \square

4.4. The quasi-inverse $(\text{Id} - \mathbb{H}_{\tau, z})^{-1}$. We conclude the section by deducing the main properties of the quasi-inverse $(\text{Id} - \mathbb{H}_{\tau, z})^{-1}$.

Proposition 4.3. *For $\delta_1 > 0$ small enough there exist $\delta_2 > 0$, $0 \leq \theta < 1$ such that, for $\mathcal{V}_1, \mathcal{V}_2$ as in Proposition 4.1, one has that:*

(i) *The operator $(\text{Id} - \mathbb{H}_{\tau, z})^{-1}$ can be analytically continued to $(\tau, z) \in \mathcal{V}_1$, $\lambda(\tau, z) \neq 1$, via the identity*

$$(4.3) \quad (\text{Id} - \mathbb{H}_{\tau, z})^{-1} = \frac{\lambda(\tau, z)}{(1 - \lambda(\tau, z))} \mathbb{P}_{\tau, z} + (\text{Id} - \mathbb{N}_{\tau, z})^{-1}.$$

Moreover, $(\text{Id} - \mathbb{N}_{\tau, z})^{-1}$ is holomorphic for $(\tau, z) \in \mathcal{V}_1$.

(ii) *$(\text{Id} - \mathbb{H}_{\tau, z})^{-1}$ is holomorphic for $(\tau, z) \in \mathcal{V}_2$.*

Proof. We start with (ii). For $(\tau, z) \in \mathcal{V}_2$, by Proposition 4.1 (v) we have $\text{srd}(\mathbb{H}_{\tau, z}) \leq \theta < 1$. Therefore, for such (τ, z) and all $f \in H^\infty(\mathbb{D})$ the series

$$\sum_{k=0}^{\infty} \mathbb{H}_{\tau, z}^k[f] = (\text{Id} - \mathbb{H}_{\tau, z})^{-1}[f]$$

converges uniformly and defines an analytic function which depends on $(\tau, z) \in \mathcal{V}_2$ holomorphically.

To prove (i) first we notice that by Proposition 4.2 we can find a non-empty open set $\mathcal{U} \subset \mathcal{V}_1$ such that $|\lambda(\tau, z)| < 1$ for all $(\tau, z) \in \mathcal{U}$. Then, for $(\tau, z) \in \mathcal{U}$ by Proposition 4.1 (iii) one has

$$(\text{Id} - \mathbb{H}_{\tau, z})^{-1} = \sum_{k=0}^{\infty} \mathbb{H}_{\tau, z}^k = \sum_{k=1}^{\infty} \lambda(\tau, z)^k \mathbb{P}_{\tau, z} + \sum_{k=0}^{\infty} \mathbb{N}_{\tau, z}^k = \frac{\lambda(\tau, z)}{(1 - \lambda(\tau, z))} \mathbb{P}_{\tau, z} + (\text{Id} - \mathbb{N}_{\tau, z})^{-1}.$$

Moreover, $\mathbb{P}_{\tau, z}$ and $(\text{Id} - \mathbb{N}_{\tau, z})^{-1}$ are holomorphic for $(\tau, z) \in \mathcal{V}_1$ and depend analytically on τ, z . By Proposition 4.2 (ii) the set $\{(\tau, z) \in \mathcal{V}_1 \mid \lambda(\tau, z) \neq 1\}$ is connected and so we can conclude by the identity principle. \square

5. THE MEROMORPHIC CONTINUATION FOR $S_\tau(z)$

In this section we study the meromorphic continuation for $S_\tau(z)$, using Propositions 3.1 and 4.3. As can be guessed from Proposition 3.1, the function $S_\tau(z)$ will have a pole whenever $\mathbb{H}_{\tau, z}$ has 1 as an eigenvalue. Therefore an important role is played by the quantity $\rho(\tau)$, defined by the implicit relation $\lambda(\tau, \rho(\tau)) = 1$. As will appear clearly in Section 6, this produces a pole of $S_\tau(z)$ at $z = \rho(\tau)$ which will eventually give the dominant contribution in the estimate (1.4).

5.1. Continuation and poles. We start with the continuation and location of poles of $S_\tau(z)$.

Proposition 5.1. *Given $\eta > 0$ sufficiently small, there exists $0 < c < \frac{1}{2}$ such that $S_\tau(z)$ is meromorphic for*

$$(\tau, z) \in \{(\tau, z) \in \mathbb{C}^2 \mid |\tau| \leq \eta, |z| \leq \frac{1}{2} + c\}.$$

with a simple pole at $z = \rho(\tau)$ only, where ρ is an analytic function in $|\tau| \leq \eta$ such that $\lambda(\tau, \rho(\tau)) = 1$ and $\rho(0) = \frac{1}{2}$. Finally, for $|\tau| \leq \eta$ we have

$$\operatorname{Res}_{z=\rho(\tau)} S_\tau(z) = \frac{\mathbb{P}_{\tau, \rho(\tau)}[\mathbf{1}](1)}{(1 - \rho(\tau)) \frac{\partial}{\partial z} \lambda(\tau, \rho(\tau))}.$$

Proof. The well-definedness and holomorphicity of ρ in a neighborhood of $\tau = 0$ follows from the implicit function theorem since λ is holomorphic on \mathcal{V}_1 and $\frac{\partial}{\partial z} \lambda(0, \frac{1}{2}) \neq 0$. Since $\lambda(0, \frac{1}{2}) = 1$, this also yields $\rho(0) = \frac{1}{2}$. The rest of the statement is then an immediate consequence of Proposition 3.1 and Proposition 4.3. \square

5.2. The derivatives of $\log(\rho(\tau))$. We now derive additional informations on the first two derivatives of $\tau \mapsto \rho(\tau)$ at $\tau = 0$. This will lead to the expression (2.4) for the mean-value α in Theorem 1.1, and to positivity of the variance.

Proposition 5.2. *We have*

$$(5.1) \quad \log(2\rho(0)) = 0, \quad \alpha := (\log \rho)'(0) = -\frac{1}{2} \int_{(0,1]} \log d\mu, \quad (\log \rho)''(0) < 0.$$

More precisely, we have

$$(5.2) \quad \begin{aligned} (\log \rho)''(0) &= -\frac{1}{2} \int_0^1 (\log x + \alpha \lfloor \frac{1}{x} \rfloor)^2 d\mu(x) + \int_0^1 \int_0^1 (\log x + \alpha \lfloor \frac{1}{x} \rfloor) \log(1 + xy) d\mu(x) d\mu(y) \\ &= -\frac{1}{2} \int_0^1 \left(\log x + \alpha \lfloor \frac{1}{x} \rfloor + \int_0^1 \log \left(\frac{1 + y \lfloor \frac{1}{x} \rfloor}{1 + yx} \right) d\mu(y) \right)^2 d\mu(x). \end{aligned}$$

The fact that $(\log \rho)''(0) < 0$ is essential and it is equivalent to the non-degeneracy of the Gaussian law in Theorem 1.1. The proof of this fact is not straightforward. As explained in [BV05, Lemma 7] (see also the proof of (6b), page 343 there), it corresponds to a general result about convexity of pressure functions [Rue94, Section 4.6]. The existence of the explicit expression (5.2) contrasts with the context of statistics of continued fractions coefficients [BV05], where no explicit expression for the variance is known in general. The underlying construction is due to Benoist and Quint [BQ16a].

For the proof of Proposition 5.2, we mostly follow [Mor15, Proposition 3.3]. As in [BV05, Lemma 7], the argument is mainly adapted from [Bro96]; additionally, we will provide a simpler way to obtain non-positivity (see (5.9) below).

We will use repeatedly the following Lemma.

Lemma 5.3. *For $f, g \in L^1(\mu)$, we have $\mathbb{H}((f \circ \mathcal{G}) \times g) = f \times (\mathbb{H} g)$ almost-everywhere.*

Proof. For all $x \in [0, 1]$ and $n \geq 1$, we have $\mathcal{G}(h_n(x)) = x$, and so $((f \circ \mathcal{G}) \times g) \circ h_n = f \times (g \circ h_n)$. Summing over $n \geq 1$ against 2^{-n} yields the claimed equality. \square

Proof of Proposition 5.2. In Proposition 5.1 we showed $\rho(0) = \frac{1}{2}$, whereas the value $(\log \rho)'(0)$ is easily obtained by differentiating $\lambda(\tau, \rho(\tau))$ at $\tau = 0$ and using the values given in Proposition 4.2.(ii).

We now wish to study the second derivative $(\log \rho)''(0)$, with the goal of proving that it is negative. For this purpose, it is convenient to consider a univariate function in place of $\lambda(\tau, z)$. Let

$$\ell(w) := \lambda(\alpha^{-1}w, \frac{1}{2}e^w),$$

which is constructed in such a way that it is defined and analytic in a neighborhood of $w = 0$, and also $\ell'(0) = 0$. There is a simple relationship between $\ell''(0)$ and the besought quantity

$(\log \rho)''(0)$: differentiating twice the relation $\lambda(\tau, e^{\log \rho(\tau)}) = 1$ and evaluating at $\tau = 0$, we obtain

$$(5.3) \quad \begin{aligned} (\log \rho)''(0)\rho(0)\frac{\partial}{\partial z}\lambda(0, \tfrac{1}{2}) &= -(\alpha^2\rho(0)\frac{\partial}{\partial z}\lambda(0, \tfrac{1}{2}) + \frac{\partial^2}{\partial \tau^2}\lambda(0, \tfrac{1}{2})) \\ &\quad + \alpha\frac{\partial^2}{\partial \tau \partial z}\lambda(0, \tfrac{1}{2}) + \frac{\alpha^2}{4}\frac{\partial^2}{\partial z^2}\lambda(0, \tfrac{1}{2}) \\ &= -\alpha^2\ell''(0). \end{aligned}$$

Therefore, our task is to prove that $\ell''(0) > 0$. To obtain this, we write the eigenvalue equation relative to $\ell(w)$, with the aim of differentiating twice, while gathering information along the way.

For w in a neighborhood of 0 we let $\xi_w = f_{\frac{w}{\alpha}, \frac{1}{2}}e^w$ with the notation of Proposition 4.1. Recall that $\xi_0 = \mathbf{1}$. Then, the eigenvalue equation is

$$\ell(w)\xi_w = \mathbb{H}_{\frac{w}{\alpha}, \frac{1}{2}}e^w[\xi_w] = \mathbb{H}[e^{w\psi}\xi_w],$$

where

$$\psi(x) = \begin{cases} \lfloor 1/x \rfloor + \alpha^{-1} \log(x) & (0 < x \leq 1), \\ 0 & (x = 0). \end{cases}$$

For future reference, we notice that $\psi \in L^2(\mu)$.

By local uniform convergence, we may differentiate the above with respect to w , obtaining

$$(5.4) \quad \ell'(w)\xi_w + \ell(w)\frac{\partial}{\partial w}\xi_w = \mathbb{H}[e^{w\psi}(\psi\xi_w + \frac{\partial}{\partial w}\xi_w)].$$

Let $\chi := [\frac{\partial}{\partial w}\xi_w]_{w=0} \in H^\infty(\mathbb{D})$. Evaluating (5.4) at $w = 0$ and using $\ell(0) = 1$, $\ell'(0) = 0$ and $\xi_0 = \mathbf{1}$ yields

$$(5.5) \quad \chi = \mathbb{H}[\psi + \chi].$$

Before continuing the analysis, we focus on the function χ . In [BQ16a], formula (1.8), an explicit expression for solutions to cohomological equations related to (5.5) is obtained. Their construction is effectuated in our case as follows. Define

$$x \mapsto \chi_1(x) := -\alpha^{-1} \int_0^1 \log(1 + xy) \, d\mu(y)$$

It is obvious that χ_1 belongs to $H^\infty(\mathbb{D})$ and is real-valued. For all $x \in [0, 1]$, we have

$$\begin{aligned} \mathbb{H}[\chi_1](x) &= -\alpha^{-1} \int_0^1 \sum_{n \geq 1} 2^{-n} \log\left(1 + \frac{y}{x+n}\right) \, d\mu(y) \\ &= -\alpha^{-1} \int_0^1 \sum_{n \geq 1} 2^{-n} \log(x + y + n) \, d\mu(y) + \alpha^{-1} \sum_{n \geq 1} 2^{-n} \log(x + n) \\ &= -\alpha^{-1} \int_0^1 \log(x + 1/y) \, d\mu(y) - \mathbb{H}[\alpha^{-1} \log](x) \\ &= \chi_1(x) - 2 - \mathbb{H}[\alpha^{-1} \log](x). \end{aligned}$$

Since $\mathbb{H}[x \mapsto \lfloor 1/x \rfloor](x) = 2$, we obtain that χ_1 satisfies (5.5) on $[0, 1]$, hence on \mathbb{D} by analytic continuation. In particular, we have $(\chi - \chi_1) = \mathbb{H}(\chi - \chi_1)$, and so, by Proposition 4.1.(ii),

$$(5.6) \quad \chi - \chi_1 \in \mathbb{C}\mathbf{1}.$$

We may now continue our analysis. Differentiating again (5.4) with respect to w , we obtain

$$\ell''(w)\xi_w + 2\ell'(w)\frac{\partial}{\partial w}\xi_w + \ell(w)\frac{\partial^2}{\partial w^2}\xi_w = \mathbb{H}[e^{w\psi}(\psi^2\xi_w + 2\psi\frac{\partial}{\partial w}\xi_w + \frac{\partial^2}{\partial w^2}\xi_w)].$$

Evaluating at $w = 0$ yields

$$(5.7) \quad \ell''(0)\mathbf{1} + [\frac{\partial^2}{\partial w^2}\xi_w]_{w=0} = \mathbb{H}[\psi^2 + 2\psi\chi + (\frac{\partial^2}{\partial w^2}\xi_w)_{w=0}].$$

Here, note that the function $\psi^2 + 2\psi\chi + (\frac{\partial^2}{\partial w^2}\xi_w)_{w=0}$ is in $L^1([0, 1], \mu)$. We integrate both sides of (5.7) against $d\mu$, in order to eliminate the terms $\frac{\partial^2}{\partial w^2}\xi_w$. Since $\int_{[0,1]} \mathbb{H}[f] d\mu = \int_{[0,1]} f d\mu$ for any $f \in L^1(\mu)$, we arrive at

$$(5.8) \quad \begin{aligned} \ell''(0) &= \int_{[0,1]} (\psi^2 + 2\psi\chi) d\mu \\ &= \int_{[0,1]} (\psi^2 + 2\psi\chi_1) d\mu \end{aligned}$$

where we have used (5.6) together with the value $\int \psi d\mu = 0$. Define now $\widehat{\psi} := \tau + \chi_1 - \chi_1 \circ \mathcal{G}$; this definition is motivated by analogy with the proof of Proposition 3.3 of [Mor15]. Note that $\widehat{\psi}$ is real-valued. Using Lemma 5.3, we have

$$\begin{aligned} \int_{[0,1]} \chi_1^2 \circ \mathcal{G} d\mu &= \int_{[0,1]} \chi_1^2 \cdot \mathbb{H}[\mathbf{1}] d\mu = \int_{[0,1]} \chi_1^2 d\mu, \\ \int_{[0,1]} (\psi + \chi_1)(\chi_1 \circ \mathcal{G}) d\mu &= \int_{[0,1]} \mathbb{H}[\psi + \chi_1]\chi_1 d\mu = \int_{[0,1]} \chi_1^2 d\mu. \end{aligned}$$

In the last step, we used the fact that equation (5.5) holds with χ replaced by χ_1 . We deduce

$$(5.9) \quad \begin{aligned} \int_{[0,1]} (\widehat{\psi})^2 d\mu &= \int_{[0,1]} (\psi + \chi_1)^2 d\mu - 2 \int_{[0,1]} (\psi + \chi_1)(\chi_1 \circ \mathcal{G}) d\mu + \int_{[0,1]} (\chi_1 \circ \mathcal{G})^2 d\mu \\ &= \int_{[0,1]} (\psi + \chi_1)^2 d\mu - \int_{[0,1]} \chi_1^2 d\mu \\ &= \ell''(0). \end{aligned}$$

Now, we note that $\psi(x) = \lfloor 1/x \rfloor + \alpha^{-1} \log(x)$ tends to $+\infty$ as $x \rightarrow 0$, whereas $\chi_1 \in H^\infty(\mathbb{D})$ is a bounded function. We may therefore find $\beta > 0$ such that $|\widehat{\psi}(x)| \geq 1$ for all $x \in (0, \beta)$. We deduce that $\ell''(0) = \int_{[0,1]} (\widehat{\psi})^2 d\mu \geq \mu((0, \beta)) = \Psi(\beta) > 0$ as required.

Since, by (5.3), we have $(\log \rho)''(0) = -\frac{1}{2}\alpha^2 \ell''(0)$, the explicit formula (5.2) follows from (5.8) and (5.9). \square

6. PROOF OF THEOREM 1.2

By the definitions (2.16) and Cauchy's formula, for $|\tau| \leq \eta$ with η sufficiently small we have

$$2^N \mathbb{E}_N[(\mathbf{S}_N)^{-\tau}] = \frac{1}{2\pi i} \oint_{|z|=\frac{1}{4}} z^{-N} S_\tau(z) \frac{dz}{z},$$

where the circle is oriented counter-clockwise. By Proposition (5.1) and the residue theorem for some $0 < c < \frac{1}{2}$ we have

$$2^N \mathbb{E}_N((\mathbf{S}_N)^{-\tau}) = \frac{R(\tau)}{\rho(\tau)^N} + \frac{1}{2\pi i} \oint_{|z|=\frac{1}{2}+c} z^{-N} S_\tau(z) \frac{dz}{z},$$

where the circle is oriented counter-clockwise and

$$R(\tau) = \frac{\mathbb{P}_{s,\rho(\tau)}[\mathbf{1}](1)}{\rho(\tau)(1 - \rho(\tau)) \frac{\partial}{\partial z} \lambda(\tau, \rho(\tau))}.$$

Note that $R(\tau)$ is analytic for $|\tau| \leq \eta$ and $R(0) = 1$. We bound the integral trivially

$$\frac{1}{2\pi i} \oint_{|z|=\frac{1}{2}+c} z^{-N} S_\tau(z) \frac{dz}{z} \ll (\frac{1}{2} + c)^{-N} \ll 2^N (1 + 2c)^{-N},$$

uniformly for $|\tau| \leq \eta$. At the possible cost of reducing η , we may write $R(\tau) = \exp V(\tau)$, and similarly $2\rho(\tau) = \exp\{-U(\tau)\}$, for two functions U and V which are holomorphic for $|\tau| \leq \eta$.

By Proposition 5.2 we also have $U(0) = V(0) = 0$, $U'(0) = \frac{1}{2} \int_{(0,1]} \log d\mu$, $U''(0) > 0$. In the same range of τ , we conclude that

$$\mathbb{E}_N((\mathbf{S}_N)^{-\tau}) = \exp(NU(\tau) + V(\tau))\{1 + O((1 + 2c)^{-N})\}$$

for some $c > 0$. This is the statement of Theorem 1.2, up to changing τ to $-\tau$. Theorem 1.1 follows at once using Hwang's quasi-power theorem, for which we refer to [FS09, Lemma IX.1] and to the papers [Hwa98, Hwa96]. We note for future reference that the variance is related to ρ by

$$(6.1) \quad \sigma^2 = -(\log \rho)''(0).$$

Remark. – The foregoing computations may be seen as a materialization of the link between the spectra of transfer operators associated with the Farey and Gauss map. There have been many works on this topic: see *e.g.* [PS92], [Iso02], [Pre03], and also [FMT07] for works on the numerical aspect. Another useful reference is the introduction of [Dod17].

In the above, by using the Cauchy formula on the generating series (2.16), we have avoided completely discussing the Farey map and its associated transfer operator

$$\mathbb{L}[f](x) = \frac{1}{2}f\left(\frac{1}{x+1}\right) + \frac{1}{2}f\left(\frac{x}{x+1}\right),$$

whose functional analysis is made difficult by the fixed point with derivative 1 (neutral fixed point) of the Farey map at 0. Instead, we extracted the precise information we needed, which is the exponential convergence for iterates of the function $\mathbf{1}$.

7. THE MEAN-VALUE α

In this section, we detail and prove Theorem 2.1.

7.1. Variants of the integral expression.

Lemma 7.1. *For the quantity α defined in (5.1), we have $\alpha = \int_0^1 \log(1+x) d\mu(x)$.*

Proof. We remark that

$$\begin{aligned} \log(1+x) + \mathbb{H}[t \mapsto \log(1+t)](x) &= \log(1+x) + \sum_{n \geq 1} 2^{-n} \log\left(\frac{1+n+x}{n+x}\right) \\ &= \sum_{n \geq 1} 2^{-n} \log(n+x) \\ &= -\mathbb{H}[\log](x) \end{aligned}$$

by splitting the logarithm into a difference. Integrating against $d\mu(x)$ and using (4.1) yields the claimed equality in the form $2 \int_0^1 \log(1+x) d\mu(x) = -\int \log d\mu$. \square

We can now justify assertion (i). Using Lemma 7.1, we see that

$$\alpha = \int_0^1 \log(1+x) d\mu(x) = \int_0^1 \log(2-x) d\mu(x) = \log 2 + \int_0^1 \log\left(1 - \frac{x}{2}\right) d\mu(x)$$

where we have used the symmetry $d\mu(1-x) = -d\mu(x)$. Expanding the logarithm into a power series, we recover the definition (2.5) used in [Bac17].

Note that Conjectures 8.1 and 8.2 of [Bac17], which are concerned with similar identities, can be proven along the same lines.

7.2. Lipschitz points of Ψ . In the paper [Kin60], Kinney studies the set of Lipschitz points of the Minkowski function Ψ . There has been many consecutive works on this topic, see *e.g.* [MZ04, KS08, KS07].

Theorem 1 of [Kin60] states that there exists a set $V \subset [0, 1]$ with $\mu(V) = 1$ and Hausdorff dimension β on which the Lipschitz exponent of Ψ is β . The value of β is recognized, using Lemma 7.1, to be precisely

$$\beta = \frac{\log 2}{2\alpha}.$$

Theorem 1.1 may be used to recover Kinney's result. We first give the following lemma, which is easily proved by induction.

Lemma 7.2. *For all $N \geq 0$ and $m \in \{0, \dots, 2^N - 1\}$, we have*

$$(7.1) \quad \frac{s(m+1)}{s(2^N + m + 1)} - \frac{s(m)}{s(2^N + m)} = \frac{1}{s(2^N + m)s(2^N + m + 1)}.$$

By equations (7.1), (2.8), and Theorem 1.1 with $t = \pm \log N$ (say), we obtain

$$(7.2) \quad \log \left| \Phi\left(\frac{m+1}{2^N}\right) - \Phi\left(\frac{m}{2^N}\right) \right| = -2N\alpha + O(\sqrt{N} \log N)$$

for all m in a set $R_N \subset \{0, \dots, 2^N - 1\}$ of cardinality $|R_N| \geq 2^N - O(2^N/N^2)$. We then define

$$U = \bigcup_{N_0 \geq 0} \bigcap_{N \geq N_0} \bigcup_{\substack{1 \leq m \leq 2^N - 1 \\ (m-1, m, m+1) \in R_N^3}} \left[\frac{m}{2^N}, \frac{m+1}{2^N} \right].$$

Since $\sum_{N \geq 0} (1 - |R_N|/2^N) < \infty$, the set U has Lebesgue measure 1. Now let $x \in U$ and $\varepsilon > 0$ be fixed. Let $N_0 = N_0(\varepsilon) \geq 1/\varepsilon^3$ be an integer, such that $x \in \bigcap_{N \geq N_0} \bigcup_{(m-1, m, m+1) \in R_N^3} \left[\frac{m}{2^N}, \frac{m+1}{2^N} \right]$. Then for $0 < h \leq 2^{-N_0}$, whenever $2^{-1-N} \leq h \leq 2^{-N}$, we have

$$\left[\frac{m}{2^{N+1}}, \frac{m+1}{2^{N+1}} \right] \subset [x-h, x+h] \subset \left[\frac{m'-1}{2^N}, \frac{m'+2}{2^N} \right]$$

for some integers (m, m') satisfying $m' \in R_{N+1}$ and $(m-1, m, m+1) \in R_N^3$. From (7.2), we deduce

$$e^{-N(2\alpha + O(\varepsilon))} \leq |\Phi(x+h) - \Phi(x-h)| \leq 3e^{-N(2\alpha - O(\varepsilon))}.$$

Since both the left-hand side and the right-hand side are of the order $h^{2\alpha/\log 2 + O(\varepsilon)}$, we conclude that the Lipschitz exponent of Φ at any $x \in U$ is equal to $2\alpha/\log 2$. Additionally, Φ and Ψ are inverse bijections, therefore the Lipschitz exponent of Ψ at any $x \in V := \Phi(U)$ is $\log 2/(2\alpha)$. Because $\mu(V)$ is equal to the Lebesgue measure of U , which is 1, we obtain that on V the Lipschitz exponent of Ψ is β .

Finally, one has

$$V = \bigcup_{N_0 \geq 0} \bigcap_{N \geq N_0} \bigcup_{\substack{1 \leq m \leq 2^N - 1 \\ (m-1, m, m+1) \in R_N^3}} \left[\frac{s(m+1)}{s(2^N + m + 1)}, \frac{s(m)}{s(2^N + m)} \right]$$

and, for any fixed $u \geq 0$

$$\begin{aligned} \sum_{\substack{1 \leq m \leq 2^N - 1 \\ (m-1, m, m+1) \in R_N^3}} \left| \frac{s(m)}{s(2^N + m)} - \frac{s(m+1)}{s(2^N + m + 1)} \right|^u &= 2^N (1 + O(N^{-2})) e^{-2uN\alpha + o(N)} \\ &= e^{(\log 2 - 2u\alpha)N + o(N)}. \end{aligned}$$

by (7.2). The fact that V has Hausdorff dimension $\leq \beta = \frac{\log 2}{2\alpha}$ then follows immediately. To prove the opposite inequality, we proceed as in [Kin60] and observe that since Ψ has Lipschitz exponent β , then for all $\varepsilon > 0$ there exist $h > 0$ and a subset $V' \subseteq V$ with $\mu(V') \geq \frac{1}{2}$, say, such that

$$(7.3) \quad \Psi(x+h', x-h') = \mu([x-h', x+h']) < (2h')^{\beta-\varepsilon}$$

for all $x \in V$ and all $0 < h' \leq h$. Now let $\{I_i \mid i \in I\}$ be a cover of V' with intervals I_i satisfying $|I_i| \leq h$. Then, taking $x_i \in V' \cap I_i$ we have $I_i \subseteq I'_i := [x_i - |I_i|, x_i + |I_i|]$ and $|I'_i| \leq 2|I_i|$. In particular, by (7.3),

$$\frac{1}{2} \leq \mu(V') \leq \sum_{i \in I} \mu(I'_i) < \sum_{i \in I} (2|I'_i|)^{\beta-\varepsilon} \leq \sum_{i \in I} (4|I_i|)^{\beta-\varepsilon}.$$

This implies that the Hausdorff dimension of V' (and henceforth that of V) is $\geq \beta - \varepsilon$. Since ε is arbitrary, the result follows.

7.3. The constant 2α as a metric entropy. We now explain how one may interpret formula (1.3) as the computation of a certain partition entropy. This may be used to derive another expression for α .

Recall that the Farey map \mathcal{F} was defined at (2.13) (see Figure 5a). Also, in this section only, we denote by ν the Lebesgue measure on $[0, 1]$. We start with the partition

$$[0, 1) = \bigvee_{k=0}^N F^{-k}[0, 1) = \bigcup_{m=0}^{2^N-1} J_{m,N},$$

where $J_{m,N}$ are consecutive segments; explicitly, we have $J_{m,N} = [\Phi(\frac{m}{2^N}), \Phi(\frac{m+1}{2^N})]$. By (7.1) we have

$$\nu(J_{m,N}) = \frac{1}{s(2^N + m)s(2^N + m + 1)}, \quad \mu(J_{m,N}) = 2^{-N}.$$

Since $s(2^{N+1}) = s(2^N)$, we deduce

$$(7.4) \quad -\frac{1}{2} \sum_{m=0}^{2^N-1} \mu(J_{m,N}) \log \nu(J_{m,N}) = \frac{2}{2^N} \sum_{m=0}^{2^N-1} \log s(2^N + m).$$

The left-hand side can be seen as a relative entropy with respect to the partition $(J_{m,N})_m$. As N tends to infinity, we can adapt the proof of Rokhlin's formula [PY98, Theorem 12.10] to show that

$$(7.5) \quad -\frac{1}{N} \sum_{m=0}^{2^N-1} \mu(J_{m,N}) \log \nu(J_{m,N}) \longrightarrow \int_0^1 \log |\mathcal{F}'(x)| d\mu(x).$$

To see this, note first that the Rényi condition [PY98, formula (12.2)] holds in our case in the weaker form

$$\sup_{(x,y) \in [0,1]^2} \left| \frac{\psi'(x)}{\psi'(y)} \right| \leq N^{O(1)}$$

for any inverse branch ψ of \mathcal{F}^N ; the steps (i) and (ii) on page 133 of [PY98] are easily carried out and yield (7.5).

For the equations (1.3), (7.4) and (7.5) to agree, the equality (2.6) must hold. Indeed, since $\log |F'(x)| = -2 \log(\max\{x, 1-x\})$ and $d\mu(1-x) = -d\mu(x)$, we deduce

$$\int_0^1 \log |\mathcal{F}'(x)| d\mu(x) = -4 \int_{1/2}^1 (\log x) d\mu(x) = 2 \int_0^1 \log(1+x) d\mu(x)$$

where we have used the equality $\mathbb{H}[t \mapsto \mathbf{1}_{t>1/2} \log t](x) = -\frac{1}{2} \log(1+x)$. By Lemma 7.1, the claimed equality (2.6) follows.

7.4. The constant α as a Lyapunov exponent. We have already noted in (2.1) that the quantity $\log \mathbf{S}_N$ may be expressed as a random product in the following way: define, as earlier, the measure η on $GL_2(\mathbb{R})$ by $\eta = \frac{1}{2} \delta_{A_0} + \frac{1}{2} \delta_{A_1}$, where $A_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $A_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Let also $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and $f(x) = x \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then

$$\log \mathbf{S}_N = \log |f(g_N \cdots g_1 v)|$$

where $g_j \in GL_2(\mathbb{R})$ are taken independently at random according to η (which simply means $g_j = A_0$ or A_1 with equal probability). Then, by definition, the mean-value α of $\log \mathbf{S}_N$ should

coincide with the first Lyapunov exponent λ_1 of η ; by a formula due to Furstenberg [Fur63, Theorem 8.5], we have

$$(7.6) \quad \lambda_1 = \int_{GL_2(\mathbb{R}) \times \mathbb{P}(\mathbb{R}^2)} \log \frac{|f(gv)|}{|f(v)|} d\eta(g) d\xi(v)$$

for any measure ξ on $\mathbb{P}(\mathbb{R}^2)$ with $\eta * \xi = \xi$. Let us now explain how this expression may be used to recover the formula (2.4), which amounts to finding an admissible measure ξ . Parametrizing $\mathbb{P}(\mathbb{R}^2) = \mathbb{R} \cup \{\infty\}$, consider the Minkowski measure μ as being defined on all $\mathbb{R} \cup \{\infty\}$ but supported on $[0, 1]$. Let $T = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$; note that $TA_1 = A_0T$. Moreover, the Minkowski measure satisfies

$$\mu = \frac{1}{2}(A_1^* \mu + (A_1 T)^* \mu).$$

Indeed, we have $\mathbb{H}^* \mu = \mu$, but from the definition of \mathbb{H} ,

$$\mathbb{H}^* \mu = \sum_{n \geq 1} 2^{-n} (TA_0^n)^* \mu = \frac{1}{2}(A_1 T)^* \mu + \frac{1}{2} A_1^* \mathbb{H}^* \mu$$

since $TA_0^n = A_1 T A_0^{n-1}$. Define ξ as

$$\xi = \frac{1}{2} \mu + \frac{1}{2} (T^* \mu).$$

A small computation, using $A_0 = TA_1 T$ and $A_0 T = TA_1$, yields

$$\begin{aligned} \eta * \xi &= \frac{1}{4}(A_1^* \mu + (A_1 T)^* \mu + (TA_1 T)^* \mu + (TA_1)^* \mu) \\ &= \frac{1}{2} \mu + \frac{1}{2} T^* \mu = \xi \end{aligned}$$

so the conditions for the integral (7.6) to hold are met. The integral on the right-hand side of (7.6) then reads

$$\begin{aligned} \lambda_1 &= \int_{\mathbb{R}_+} \frac{1}{2} \log |f(A_0(\frac{x}{1})) f(A_1(\frac{x}{1}))| d\xi(x) \\ &= \frac{1}{2} \int_0^\infty \log(1+x) d\xi(x) \\ &= \frac{1}{4} \int_0^1 (\log(1+x) + \log(1+x^{-1})) d\mu(x) \end{aligned}$$

which evaluates to α by (2.4) and Lemma 7.1.

7.5. Explicit expression for the variance. We recall that the variance in Theorem 1.1 is related to $\tau \mapsto \rho(\tau)$ by formula (6.1). The expression stated in Theorem 2.1.(v) immediately follows using formula (5.2). Note that the alternative expression

$$\sigma^2 = \frac{1}{2} \int_0^1 (\log x + \alpha \lfloor \frac{1}{x} \rfloor)^2 d\mu(x) - \int_0^1 \int_0^1 (\log x + \alpha \lfloor \frac{1}{x} \rfloor) \log(1+xy) d\mu(x) d\mu(y)$$

is particularly well-suited for numerical computation on the lines of [Bac17].

Acknowledgements. The authors are grateful to B. Vallée and to the anonymous referee for numerous insightful remarks on earlier versions of this manuscript. SD thanks É. de Panafieu, J. Emme, P. Hubert, S. Troubetzkoy and G. Merlet for helpful discussions. Part of this work was done during a visit of SD at University of Genova, and a visit of LS at Aix-Marseille University. Both institutions are thanked for their support. The work of the SB is partially supported by PRIN “Number Theory and Arithmetic Geometry”.

REFERENCES

- [AS03] J.-P. Allouche and J. Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003, Theory, applications, generalizations.
- [Bac17] R. Bacher, *The Stern sequence and moments of Minkowski's question mark function*, arXiv:1703.07268, 2017.
- [Bal00] V. Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific Publishing Co., Inc., River Edge, NJ, 2000.
- [Bel54] R. Bellman, *Limit theorems for non-commutative operations. I*, Duke Math. J. **21** (1954), 491–500.
- [BQ16a] Y. Benoist and J.-F. Quint, *Central limit theorem for linear groups*, Ann. Probab. **44** (2016), no. 2, 1308–1340.
- [BQ16b] ———, *Random walks on reductive groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 62, Springer, Cham, 2016.
- [Bro96] A. Broise, *Transformations dilatantes de l'intervalle et théorèmes limites*, Astérisque (1996), no. 238, 1–109, Études spectrales d'opérateurs de transfert et applications.
- [BV05] V. Baladi and B. Vallée, *Euclidean algorithms are Gaussian*, J. Number Theory **110** (2005), no. 2, 331–386.
- [CS17] M. Coons and L. Spiegelhofer, *The maximal order of hyper-(b-ary)-expansions*, Electron. J. Combin. **24** (2017), no. 1, P1.15.
- [CT14] M. Coons and J. Tyler, *The maximal order of Stern's diatomic sequence*, Mosc. J. Comb. Number Theory **4** (2014), no. 3, 3–14.
- [CW00] N. Calkin and H. S. Wilf, *Recounting the rationals*, Amer. Math. Monthly **107** (2000), no. 4, 360–363.
- [Def16] C. Defant, *Upper bounds for Stern's diatomic sequence and related sequences*, Electron. J. Combin. **23** (2016), no. 4, Paper 4.8, 47.
- [Den38] A. Denjoy, *Sur une fonction réelle de Minkowski*, J. Math. Pures Appl. IX. Sér. **17** (1938), 105–151.
- [Dod17] P. S. Dodds, *On the thermodynamic formalism for the Farey map*, arXiv:1701.04486, 2017.
- [EH18] J. Emme and P. Hubert, *Central limit theorem for probability measures defined by sum-of-digits function in base 2*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (2018), to appear.
- [FMT07] G. Froyland, R. Murray, and D. Terhesiu, *Efficient computation of topological entropy, pressure, conformal measures, and equilibrium states in one dimension*, Phys. Rev. E (3) **76** (2007), no. 3, 036702, 5.
- [FS09] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009.
- [Fur63] H. Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428.
- [GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics*, second ed., Addison-Wesley Publishing Company, Reading, MA, 1994, A foundation for computer science.
- [Hen94] D. Hensley, *The number of steps in the Euclidean algorithm*, J. Number Theory **49** (1994), no. 2, 142–182.
- [Hwa96] H.-K. Hwang, *Large deviations for combinatorial distributions. I. Central limit theorems*, Ann. Appl. Probab. **6** (1996), no. 1, 297–319.
- [Hwa98] ———, *On convergence rates in the central limit theorems for combinatorial structures*, European J. Combin. **19** (1998), no. 3, 329–343.
- [Iso02] S. Isola, *On the spectrum of Farey and Gauss maps*, Nonlinearity **15** (2002), no. 5, 1521–1539.
- [Kat95] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition.
- [Kin60] J. R. Kinney, *Note on a singular function of Minkowski*, Proc. Amer. Math. Soc. **11** (1960), 788–794.
- [KMS16] M. H. Kesseböhmer, S. Munday, and B. O. Stratmann, *Infinite ergodic theory of numbers*, De Gruyter Graduate, De Gruyter, Berlin, 2016.
- [KS07] M. Kesseböhmer and B. O. Stratmann, *A multifractal analysis for Stern-Brocot intervals, continued fractions and Diophantine growth rates*, J. Reine Angew. Math. **605** (2007), 133–163.
- [KS08] ———, *Fractal analysis for sets of non-differentiability of Minkowski's question mark function*, J. Number Theory **128** (2008), no. 9, 2663–2686.
- [Lag92] J. C. Lagarias, *Number theory and dynamical systems*, The unreasonable effectiveness of number theory (Orono, ME, 1991), Proc. Sympos. Appl. Math., vol. 46, Amer. Math. Soc., Providence, RI, 1992, pp. 35–72.
- [Lan14] J. Lansing, *Largest values for the Stern sequence*, J. Integer Seq. **17** (2014), no. 7, Article 14.7.5, 18.
- [Leh29] D. H. Lehmer, *On Stern's Diatomic Series*, Amer. Math. Monthly **36** (1929), no. 2, 59–67.
- [May91] D. H. Mayer, *Continued fractions and related transformations*, Ergodic theory, symbolic dynamics, and hyperbolic spaces (Trieste, 1989), Oxford Sci. Publ., Oxford Univ. Press, New York, 1991, pp. 175–222.
- [Min05] H. Minkowski, *Zur Geometrie der Zahlen.*, Verh. d. 3. intern. Math.-Kongr. Heidelb., 164-173 (1905)., 1905.

- [Mor15] I. D. Morris, *A short proof that the number of division steps in the Euclidean algorithm is normally distributed*, arXiv:1502.07616, 2015.
- [MR17] M. Misiurewicz and A. Rodrigues, *Counting preimages*, Ergodic Theory Dynam. Systems (2017), to appear.
- [MZ04] N. Moshchevitin and A. Zhigljavsky, *Entropies of the partitions of the unit interval generated by the Farey tree*, Acta Arith. **115** (2004), no. 1, 47–58.
- [Pan08] G. Panti, *Multidimensional continued fractions and a Minkowski function*, Monatsh. Math. **154** (2008), no. 3, 247–264.
- [Pau17] R. Paulin, *Largest values of the Stern sequence, alternating binary expansions and continuants*, J. Integer Seq. **20** (2017), no. 2, Art. 17.2.8, 13.
- [Pre03] T. Prellberg, *Towards a complete determination of the spectrum of a transfer operator associated with intermittency*, J. Phys. A **36** (2003), no. 10, 2455–2461.
- [PS92] T. Prellberg and J. Slawny, *Maps of intervals with indifferent fixed points: thermodynamic formalism and phase transitions*, J. Statist. Phys. **66** (1992), no. 1-2, 503–514.
- [PY98] M. Pollicott and M. Yuri, *Dynamical systems and ergodic theory*, London Mathematical Society Student Texts, vol. 40, Cambridge University Press, Cambridge, 1998.
- [Rue94] D. Ruelle, *Dynamical zeta functions for piecewise monotone maps of the interval*, CRM Monograph Series, vol. 4, American Mathematical Society, Providence, RI, 1994.
- [Rue04] ———, *Thermodynamic formalism*, second ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2004, The mathematical structures of equilibrium statistical mechanics.
- [Sal43] R. Salem, *On some singular monotonic functions which are strictly increasing*, Trans. Amer. Math. Soc. **53** (1943), 427–439.
- [Ste58] M. Stern, *Über eine zahlentheoretische Funktion*, J. Reine Angew. Math. **55** (1858), 193–220.
- [Val03] B. Vallée, *Dynamical analysis of a class of Euclidean algorithms*, Theoret. Comput. Sci. **297** (2003), no. 1-3, 447–486, Latin American theoretical informatics (Punta del Este, 2000).

SB: DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI GENOVA, VIA DODECANESO 35, 16146 GENOVA, ITALY
E-mail address: `bettin@dim.unige.it`

SD: AIX MARSEILLE UNIVERSITÉ, CNRS, CENTRALE MARSEILLE, I2M UMR 7373, 13453 MARSEILLE, FRANCE
E-mail address: `sary-aurelien.drappeau@univ-amu.fr`

LS: INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TU WIEN, WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA
E-mail address: `lukas.spiegelhofer@tuwien.ac.at`