ACTA ARITHMETICA
Online First version

Another note on "Euclidean algorithms are Gaussian" by V. Baladi and B. Vallée

by

JUNGWON LEE and HAE-SANG SUN (Ulsan)

1. Introduction. Every rational number $0 < r < 1$ has a unique, simple finite continued fraction expansion $r = [0; m_1, \ldots, m_\ell]$ with integers $m_1, \ldots, m_{\ell-1} \geq 1$ and $m_\ell \geq 2$. One can regard the length $\ell := \ell(r)$ of the continued fraction as a random variable on the set of rational numbers with a fixed denominator,

$$\Sigma_N = \{m/N \mid 1 \leq m < N,\ (m, N) = 1\},$$

equipped with the uniform probability. It has been expected that ℓ follows the asymptotic Gaussian distribution as N goes to infinity.

A prominent result goes back to Hensley [5] who showed an average version of this conjecture, i.e., the asymptotic Gaussian distribution of ℓ on the larger probability space

$$\Omega_N = \{u/v \mid 1 \leq u < v \leq N,\ (u, v) = 1\}$$

with the uniform probability $\mathbb{P}_N$. The result was further generalised by Baladi and Vallée [1] in a remarkable way that is based on the dynamical analysis of the Euclidean algorithm.

For $r \in \Omega_N$, Baladi and Vallée considered a non-negative real value $c(m)$ associated to each possible digit $m \geq 1$ of the continued fraction expansion of r, with one mild assumption $c(m) = O(\log m)$, and they defined the total cost C of r by $C(r) := \sum_{i=1}^{\ell(r)} c(m_i)$. Then C can be regarded as a random variable on Ω_N. They proved:

THEOREM A (Central Limit Theorem, [1]). *The distribution of the total cost C on Ω_N is asymptotically Gaussian, with speed of convergence*

2010 *Mathematics Subject Classification*: Primary 37A30; Secondary 11Y16.
Key words and phrases: dynamical systems, transfer operator, Dirichlet series, Perron's formula.
Received 18 April 2017; revised 16 February 2018 and 18 September 2018.
Published online 7 March 2019.

DOI: 10.4064/aa170418-6-3

$O(1/\sqrt{\log N})$, *i.e., for suitable positive constants* μ *and* δ,

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

We remark that the work of Baladi–Vallée not only generalised Hensley's result that corresponds to $c\equiv 1$, but also improved it with an optimal error.

1.1. Main results. We shall outline the main steps of the proof of Theorem A as follows. Let us first state a crucial criterion for the asymptotic Gaussian distribution.

THEOREM 1.1 (Hwang's Quasi-Power Theorem, [1]). *Assume that the moment generating functions for a sequence of random variables* X_N *on probability spaces* $(\Xi_N,\mathbb{P}_N)$ *are analytic in a neighborhood* W *of zero, and*

$$\mathbb{E}_N[\exp(wX_N)]=\exp(\beta_N U(w)+V(w))(1+O(\kappa_N^{-1}))$$

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

$$\mathbb{E}(X_N)=\beta_N U'(0)+V'(0)+O(\kappa_N^{-1}),$$
$$\mathbb{V}(X_N)=\beta_N U''(0)+V''(0)+O(\kappa_N^{-1}).$$

Further, the distribution of X_N *on* Ξ_N *is asymptotically Gaussian, with speed of convergence* $O(\kappa_N^{-1}+\beta_N^{-1/2})$.

To obtain the quasi-power expression for the moment generating function of the total cost C, Baladi and Vallée studied a certain Dirichlet series whose coefficients are related to the generating function. A crucial point is the observation that the Dirichlet series admits an alternative expression in terms of the so-called transfer operator, and a Tauberian argument to estimate the coefficients of the Dirichlet series is deduced from the spectral analysis of the transfer operator.

However, Perron's formula of order two used in [1, (2.19)] only provides an estimate for iterated average sums of coefficients, which does not directly yield the necessary quasi-power expression for the total cost C on Ω_N. So Baladi and Vallée introduced the smoothed probabilistic model $\Omega_N(\varepsilon)$ containing Ω_N for sufficiently small $\varepsilon=\varepsilon(N)$ and showed that the distribution of C on $\Omega_N(\varepsilon)$ is asymptotically Gaussian. By showing that the difference between the two probabilities, on Ω_N and $\Omega_N(\varepsilon)$, is $O(\varepsilon)$, they obtained the result.

The purpose of this short article is to make a few remarks on the work of Baladi–Vallée [1]. First, we obtain Theorem A without the smoothing process. More precisely, it is possible to get directly a quasi-power estimate for the moment generating function of C on Ω_N by applying a version of Perron's formula with error terms. Furthermore, the smoothing process is also doable for Σ_N, that is, there is an auxiliary space $\Sigma_N(\varepsilon)$ containing Σ_N (see Section 4 for a precise definition) on which the cost C is asymptotically Gaussian:

THEOREM B. *The distribution of the total cost C on $\Sigma_N(\varepsilon)$ is asymptotically Gaussian, with speed of convergence $O(1/\sqrt{\log N})$.*

Finally, we present Question C, which is a Σ_N-version of the last step of the smoothing process, i.e., the statement that the difference between two probabilistic models Σ_N and $\Sigma_N(\varepsilon)$ is $O(\varepsilon)$. The conjecture on the asymptotic Gaussian behavior of the length ℓ of continued fractions on Σ_N immediately follows from an affirmative answer to this question.

2. Work of Baladi–Vallée. The distribution of the cost C on Ω_N is determined by the Lévy moment generating function

$$\mathbb{E}_N[\exp(wC)] = \frac{1}{|\Omega_N|}\sum_{r\in\Omega_N}\exp[wC(r)]$$

for complex w close enough to 0. Baladi and Vallée studied the Dirichlet series

$$L(s,w) = \sum_{n\geq 1}\frac{c_n(w)}{n^s}, \qquad c_n(w) = \sum_{r\in\Sigma_n}\exp[wC(r)],$$

for $\Re s > 1$ and $|w|$ sufficiently small. Note that

$$\sum_{n\leq N} c_n(w) = \sum_{r\in\Omega_N}\exp[wC(r)]$$

is essentially the moment generating function of C on Ω_N. Therefore, the statistics of C follows from Tauberian type arguments for $L(s,w)$.

A crucial point is that analytic properties of $L(s,w)$ can be investigated by the following thermodynamical formalism, the so-called transfer operator method. For $r = u/v \in \Omega_N$, executing the Euclidean algorithm on the inputs u and v yields a unique continued fraction expansion $r = [0; m_1,\dots,m_\ell]$ with $m_1,\dots,m_{\ell-1}\geq 1$ and $m_\ell \geq 2$. Observe that each digit can be written as $m_i = \lfloor T^{i+1}(r)\rfloor$, where T is the Gauss map

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

Here, $\lfloor x\rfloor$ denotes the integer part of x. The continued fraction expansions can be viewed as rational trajectories of a one-dimensional dynamical system $([0,1],T)$ that reaches 0 in a finite number of steps.

Let $\mathcal{H}$ be the set of inverse branches of T that are of the form $h_{[m]}(x) = 1/(m+x)$ for some $m\geq 1$. The digit cost c can be regarded as a function on $\mathcal{H}$ via $c(h_{[m]}) := c(m)$. The weighted transfer operator $H_{s,w}$ on $C^1([0,1])$, which depends on two complex parameters s and w, is defined by

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

We write $F_{s,w}$ for the same operator with $\mathcal{H}$ replaced by the final set $\mathcal{F} := \left\{\frac{1}{m+x} \mid m \geq 2\right\} \subseteq \mathcal{H}$.

Let $\Omega := \bigcup_{n\geq 1} \Omega_n$ and $\mathcal{H}^* := \bigcup_{n\geq 1} \mathcal{H}^n$. The above arguments show that each $r = u/v \in \Omega$ can be written as

$$u/v = h_{[m_1]} \circ \cdots \circ h_{[m_{\ell(r)}]}(0) =: h(0)$$

with $h_{[m_i]} \in \mathcal{H}$, $1 \leq i \leq \ell(r) - 1$ and $h_{[m_{\ell(r)}]} \in \mathcal{F}$. Finally, one can observe the following key relation between $L(2s, w)$ and the transfer operator $H_{s,w}$:

$$\begin{aligned} L(2s, w) &= \sum_{r\in\Omega} \frac{1}{v^{2s}} \exp[wC(r)] = \sum_{h\in\mathcal{H}^*} |h'(0)|^s \exp[wc(h)] \\ &= F_{s,w} \circ (I - H_{s,w})^{-1}[1](0). \end{aligned}$$

This relation and the estimate of the operator norm of $(I - H_{s,w})^{-1}$ due to Dolgopyat [3] enable Baladi and Vallée to show that $L(2s, w)$ can be meromorphically continued to $\mathbb{C}$ and has a simple pole at $s = \sigma(w)$, which is analytic in w near 0 and $\sigma(0) = 1$. Furthermore, a crucial bound of $L(2s, w)$ is obtained on a vertical strip containing $s = 1$, as follows.

LEMMA 2.1. *For all ξ with $0 < \xi < 1/5$, we can find α_0 such that for any $\widehat{\alpha}_0$ with $0 < \widehat{\alpha}_0 < \alpha_0 \leq 1/2$, there exist a neighborhood W' of* 0 *and constants $M, M' > 0$ such that for all $w \in W'$ we have:*

(1) $\Re\sigma(w) > 1 - (\alpha_0 - \widehat{\alpha}_0)$.
(2) $L(2s, w)$ *has only a simple pole at $s = \sigma(w)$ in the strip $|\Re s - 1| \leq \alpha_0$.*
(3) $|L(2s, w)| \leq M \max(1, |t|^\xi)$ *on the vertical line $\Re s = 1 \pm \alpha_0$, $t = \Im s$.*
(4) $|L(2s, w)| \leq M'|t|^\xi$ *in the strip for sufficiently large $|t|$.*

Proof. The first three statements are just [1, Lemma 8]. The statement (4) comes from [1, Theorem 2]. ■

Using the bound in Lemma 2.1 and the Perron formula of order two, Baladi and Vallée showed that for $0 < \widehat{\gamma} < \alpha_0$,

$$\sum_{Q\leq N} \sum_{n\leq Q} c_n(w) = A(w)N^{2\sigma(w)+1}(1 + O(N^{-\widehat{\gamma}})) \tag{2.1}$$

where $A(w)$ is non-vanishing.

In order to get a result on C for Ω_N, they first considered a smoothed version of (2.1). In other words, for $\varepsilon(N) = N^{-\widehat{\gamma}/2}$, it can be deduced from the formula (2.1) that

$$\sum_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} \sum_{n\le Q} c_n(w) = \lfloor N\varepsilon(N)\rfloor A(w)(2\sigma(w)+1)N^{2\sigma(w)}(1+O(N^{-\widehat{\gamma}/2})). \tag{2.2}$$

This implies that the moment generating function $\mathbb{E}_N[\exp(wC)|\Omega_N(\varepsilon)]$ of C on the auxiliary probability space

$$\Omega_N(\varepsilon) := \bigcup_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} \Omega_Q \times \{Q\}$$

satisfies a quasi-power estimate

$$\mathbb{E}_N[\exp(wC)|\Omega_N(\varepsilon)] = \frac{A(w)(2\sigma(w)+1)}{3A(0)} N^{2(\sigma(w)-\sigma(0))}(1+O(N^{-\widehat{\gamma}/2})). \tag{2.3}$$

The definition of $\Omega_N(\varepsilon)$ in [1] is invalid and the correct one can be found in Cesaratto [2]. The formula (2.3) corresponds to the fact that C asymptotically follows the Gaussian distribution on the smoothed probability space $\Omega_N(\varepsilon)$ due to Theorem 1.1.

For the last step, with the choice of $\varepsilon(N)$, Baladi and Vallée proved that the difference between the probabilities $\mathbb{P}_N$ on Ω_N and $\mathbb{P}_{N,\varepsilon}$ on $\Omega_N(\varepsilon)$ is $O(\varepsilon(N))$. Note that $\mathbb{P}_N$ and $\mathbb{P}_{N,\varepsilon}$ are not defined on the same probability space. However, they can be compared by dealing with sets $A_{N,\varepsilon} \subseteq \Omega_N(\varepsilon)$ coming from subsets $A \subseteq \Omega_N$, which are essentially of the form

$$A_{N,\varepsilon} := \bigcup_{Q=N-\lfloor N\varepsilon\rfloor}^{N} (A\cap\Omega_Q)\times\{Q\}.$$

In other words, for any $A \subseteq \Omega_N$,

$$|\mathbb{P}_N(A) - \mathbb{P}_{N,\varepsilon}(A_{N,\varepsilon})| = O(\varepsilon(N)),$$

from which Baladi and Vallée deduced Theorem A.

Remark 2.2. They also established the Local Limit Theorem [1, Theorem 4] for ℓ on Ω_N. The result basically follows from the quasi-power estimate for the moment generating function with the saddle-point method. With the same μ and δ,

$$\mathbb{P}_N\left[x - \frac{1}{2\delta\sqrt{\log N}} < \frac{\ell(r)-\mu\log N}{\delta\sqrt{\log N}} \le x + \frac{1}{2\delta\sqrt{\log N}}\right] = \frac{e^{-x^2/2}}{\delta\sqrt{2\pi\log N}} + O\left(\frac{1}{\log N}\right).$$

3. Distributional analysis of C on Ω_N. In this section, we give a direct proof of Theorem A without using the smoothing process. Instead of Perron's formula of order two used in [1], we use a version of Perron's formula with error terms that is Lemma 3.19 in Titchmarsh [6]. Recall that Baladi–Vallée used one without error estimates and that the smoothing process is required to obtain the desired estimate for the moment generating function of C on Ω_N. However, the following version of the formula enables us to obtain a direct quasi-power estimate for the original cost by taking the optimal error term. Thus, we deduce the asymptotic Gaussian distribution of C on Ω_N from Theorem 1.1.

Theorem 3.1 (Truncated Perron's Formula). *Let $F(s) = \sum_{n\geq 1} a_n/n^s$ for $\Re(s) := \sigma > \sigma_a$, the abscissa of absolute convergence of $F(s)$. Then for $D > \sigma_a$, one has*

$$\sum_{n\leq x} a_n = \frac{1}{2\pi i}\int_{D-iT}^{D+iT} F(s)\frac{x^s}{s}\,ds + O\biggl(\frac{x^D|F|(D)}{T}\biggr) + O\biggl(\frac{A(2x)x\log x}{T}\biggr) + O\biggl(A(N)\min\biggl\{\frac{x}{T|x-N|},1\biggr\}\biggr),$$

where

$$|F|(\sigma) = \sum_{n\geq 1}\frac{|a_n|}{n^\sigma}$$

for $\sigma > \sigma_a$, N is the nearest integer to x, and $a_n = O(A(n))$, with $A(n)$ non-decreasing.

We first introduce the set up. For $r \in \Omega_N$, we write $r = [0; m_1, \ldots, m_\ell]$. Note that ℓ corresponds to the case $c \equiv 1$ and satisfies $\ell(r) = O(\log N)$. Now, for a general c satisfying $c(m) = O(\log m)$, the Dirichlet series

$$\sum_{m\geq 1}\exp[wc(m)]\cdot\frac{1}{(m+x)^{2s}}$$

converges when $(\Re s, \Re w)$ belongs to a real neighborhood of $(1, 0)$. That is, for $\Re s$ near 1, $\exp[\Re w\cdot c(m)] < m^{\Re s}$ ($\sim m^1$). Thus for each r, the total cost satisfies

$$C(r) = c(m_1) + \cdots + c(m_\ell) < \eta\cdot\log(m_1\cdots m_\ell)$$

with η being sufficiently small. This implies that $C(r) = O(\log N)$ once we show $m_1\cdots m_\ell \leq N$.

Lemma 3.2. *For $r \in \Omega_N$, let $r = [0; m_1, \ldots, m_\ell]$. Then*

$$m_1\cdots m_\ell \leq N.$$

Proof. Recall that we have the expression $r = h_{[m_1]}\circ\cdots\circ h_{[m_\ell]}(0)$ with inverse branches $h_{[m_i]}(x) = \frac{1}{m_i+x}$ which correspond to $\left[\begin{smallmatrix}0 & 1\\ 1 & m_i\end{smallmatrix}\right] \in \mathrm{GL}_2(\mathbb{Z})$.

With the canonical $\mathrm{GL}_2(\mathbb{Z})$-action on the rational numbers, r can be written as

$$r = h_{[m_1]} \circ \cdots \circ h_{[m_\ell]}(0) = \begin{bmatrix} 0 & 1 \\ 1 & m_1 \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & m_\ell \end{bmatrix} \cdot \frac{0}{1} = \frac{b}{d}$$

where $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ with $ad - bc = \pm 1$ is the product of the matrices on the L.H.S. Therefore, we get $m_1 \cdots m_\ell < d \le N$ by looking at the $(2,2)$-component. ∎

In this case,

$$c_n(w) = \sum_{r \in \Sigma_n} \exp[wC(r)] = O(n^{1+k\Re w})$$

for some $k > 0$. Therefore, we may take $A(n) = n^{1+k\Re w}$ in Perron's formula. Together with a choice of optimal T, we have the following.

Theorem 3.3. *For a non-vanishing $B(w)$ and $\gamma > 0$, we have*

$$\sum_{n \le N} c_n(w) = B(w) N^{2\sigma(w)} (1 + O(N^{-\gamma})).$$

Proof. The analytic properties of $L(2s, w)$, summarized in Lemma 2.1, allow us to do contour integration using Cauchy's residue theorem:

$$\frac{1}{2\pi i} \int_{\mathcal{U}_T(w)} L(2s, w) \frac{N^{2s}}{2s}\, d(2s) = \frac{E(w)}{\sigma(w)} N^{2\sigma(w)}.$$

Here $E(w)$ is the residue of $L(2s, w)$ at the simple pole $s = \sigma(w)$, and $\mathcal{U}_T(w)$ is the positively oriented rectangle with vertices $1 + \alpha_0 + iT$, $1 - \alpha_0 + iT$, $1 - \alpha_0 - iT$, and $1 + \alpha_0 - iT$. Together with Perron's formula in Theorem 3.1, we have

$$\begin{aligned} \sum_{n \le N} c_n(w) = \frac{E(w)}{\sigma(w)} N^{2\sigma(w)} &+ O\left(\frac{N^{2(1+\alpha_0)}}{T}\right) \\ &+ O\left(\frac{A(2N) N \log N}{T}\right) + O(A(N)) \\ &+ O\left(\int_{1-\alpha_0-iT}^{1-\alpha_0+iT} |L(2s, w)| \frac{N^{2(1-\alpha_0)}}{|s|}\, ds\right) \\ &+ O\left(\int_{1-\alpha_0 \pm iT}^{1+\alpha_0 \pm iT} |L(2s, w)| \frac{N^{2\Re s}}{T}\, ds\right). \end{aligned}$$

Note that the last two error terms are from the contour integral, corresponding to the left vertical side and the horizontal sides of the rectangle $\mathcal{U}_T(w)$.

Let us write the last formula as

$$\sum_{n\le N} c_n(w) = \frac{E(w)}{\sigma(w)} N^{2\sigma(w)}(1 + \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}).$$

We choose $\widehat{\alpha}_0$ with

$$\tfrac{32}{79}\alpha_0 < \widehat{\alpha}_0 < \alpha_0$$

and set

$$T = N^{2\alpha_0 + 4\widehat{\alpha}_0}.$$

Notice that $\frac{E(w)}{\sigma(w)}$ is bounded in the neighborhood W' since $\sigma(0) = 1$. Then, the error terms are bounded as follows:

- The error term I is $O(N^{2(1-2\widehat{\alpha}_0-\Re\sigma(w))})$. By Lemma 2.1, the exponent satisfies
$$2(1 - 2\widehat{\alpha}_0 - \Re\sigma(w)) < 2(\alpha_0 - 3\widehat{\alpha}_0) < 0.$$
- For any ε_1 with $0 < \varepsilon_1 < \widehat{\alpha}_0/2$, we can choose W' again from Lemma 2.1 small enough to have $k\Re w < \varepsilon_1/2$ so that $A(N) = O(N^{1+\varepsilon/2})$. Then the exponent of N in II is equal to
$$1 - 2(\Re\sigma(w) + k\Re w) - (2\alpha_0 + 4\widehat{\alpha}_0) \le -2\alpha_0 + \tfrac{5}{2}\widehat{\alpha}_0 < 0.$$
Here, recall that $0 < \alpha_0 \le 1/2$.
- Similarly, the error term III is $O(N^{1+k\Re w-2\Re\sigma(w)})$. The exponent satisfies
$$\begin{aligned}1 + k\Re w - 2\Re\sigma(w) &< -1 + 2(\alpha_0 + \widehat{\alpha}_0) + \varepsilon_1/2\\ &< -1 + 2\alpha_0 - \tfrac{7}{4}\widehat{\alpha}_0 \le -\tfrac{7}{4}\widehat{\alpha}_0 < 0.\end{aligned}$$
- For any $0 < \xi < 1/5$, we also have $|L(2s, w)| \le M|t|^{\xi}$ by Lemma 2.1 where $\Im s = t$. The error term IV is $O(N^{2(1-\alpha_0-\Re\sigma(w))}T^{\xi})$ and the exponent of N is
$$\begin{aligned}&2(1 - \alpha_0 - \Re\sigma(w)) + (2\alpha_0 + 4\widehat{\alpha}_0)\xi\\ &\quad< 2(1 - \alpha_0 - (1 - \alpha_0 + \widehat{\alpha}_0)) + \tfrac{1}{5}(2\alpha_0 + 4\widehat{\alpha}_0) = \tfrac{2}{5}(\alpha_0 - 3\widehat{\alpha}_0) < 0.\end{aligned}$$
- The last term V is $O(T^{\xi-1} \cdot N^{2(1+\alpha_0-\Re\sigma(w))}\log N)$. Hence, the exponent satisfies
$$\begin{aligned}&(2\alpha_0 + 4\widehat{\alpha}_0)(\xi - 1) + 2(1 + \alpha_0 - \Re\sigma(w)) + \varepsilon_1/2\\ &\quad< -\tfrac{4}{5}(2\alpha_0 + 4\widehat{\alpha}_0) + \tfrac{1}{4}\widehat{\alpha}_0 + 2(2\alpha_0 - \widehat{\alpha}_0) < \tfrac{12}{5}\big(\alpha_0 - \tfrac{99}{48}\widehat{\alpha}_0\big) < 0.\end{aligned}$$

By taking

$$\gamma = \min\big(\tfrac{7}{4}\widehat{\alpha}_0, \tfrac{2}{5}(3\widehat{\alpha}_0 - \alpha_0), \tfrac{12}{5}\big(\tfrac{99}{48}\widehat{\alpha}_0 - \alpha_0\big)\big),$$

we obtain the theorem. ∎

Finally, with $0 < \gamma < \alpha_0$ from Theorem 3.3, the moment generating function of the total cost C on Ω_N admits a quasi-power expression

$$\mathbb{E}_N[\exp(wC)] = \frac{B(w)}{B(0)} N^{2(\sigma(w)-\sigma(0))}(1 + O(N^{-\gamma})).$$

REMARK 3.4. Of course, this theorem enables us to prove Theorem A directly from Theorem 1.1 without the smoothing process. In the following section, we observe that the smoothing process is also doable for Σ_N. This yields the asymptotic Gaussian distribution of C on $\Sigma_N(\varepsilon)$.

4. Distributional analysis of C on Σ_N. As before, we also define a smoothed probability space $\Sigma_N(\varepsilon)$ as follows. For $\varepsilon(N) = N^{-\gamma/2}$ and $\gamma > 0$ from Theorem 3.3, we consider the probability space

$$\Sigma_N(\varepsilon) := \bigcup_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} \Sigma_Q$$

with the uniform probability $\mathbb{P}_{N,\varepsilon}$. Then the moment generating function of C on $\Sigma(\varepsilon)$ is

$$\mathbb{E}_N[\exp(wC)|\Sigma_N(\varepsilon)] = \frac{1}{|\Sigma_N(\varepsilon)|} \sum_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} c_Q(w).$$

Let us write $\Psi_w(N) = \sum_{n\le N} c_n(w)$. Clearly,

$$\sum_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} c_Q(w) = \Psi_w(N) - \Psi_w(N - \lfloor N\varepsilon(N)\rfloor),$$

and $|\Sigma_N(\varepsilon)| = \sum_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} c_Q(0)$. The following smoothing process is similar to the one in Baladi–Vallée [1] and gives information on C for $\Sigma_N(\varepsilon)$.

PROPOSITION 4.1. *With the same setting as in Theorem* 3.3, *we have*

$$\sum_{Q=N-\lfloor N\varepsilon(N)\rfloor}^{N} c_Q(w) = 2\lfloor N\varepsilon(N)\rfloor B(w)\sigma(w)N^{2\sigma(w)-1}(1 + O(N^{-\gamma/2})).$$

Proof. For simplicity, we may write $F_w(N) = B(w)N^{2\sigma(w)}$. By Theorem 3.3, we have

$$\begin{aligned}
&\Psi_w(N) - \Psi_w(N - \lfloor N\varepsilon(N)\rfloor)\\
&\qquad = [F_w(N) - F_w(N - \lfloor N\varepsilon(N)\rfloor)] + O(F_w(N)N^{-\gamma})\\
&\qquad = \lfloor N\varepsilon(N)\rfloor F_w'(N) + O(F_w(N)N^{-\gamma})\\
&\qquad = \lfloor N\varepsilon(N)\rfloor F_w'(N)\left[1 + O\left(\frac{1}{\lfloor N\varepsilon(N)\rfloor}\cdot\frac{F_w(N)N^{-\gamma}}{F_w'(N)}\right)\right].
\end{aligned}$$

Note that $\frac{F_w(N)}{F'_w(N)} = \frac{N}{2\sigma(w)}$ and $\sigma(w)$ is bounded, holomorphic on a neighborhood W'. Since $\varepsilon(N) = N^{-\gamma/2}$, the last error term is equal to $O(N^{-\gamma/2})$, and this finishes the proof. ∎

Therefore the moment generating function of C on $\Sigma_N(\varepsilon)$ satisfies

$$\mathbb{E}_N[\exp(wC)|\Sigma_N(\varepsilon)] = \frac{B(w)\sigma(w)}{B(0)} N^{2(\sigma(w)-\sigma(0))}(1 + O(N^{-\gamma/2})),$$

and from Theorem 1.1 we are able to conclude the following.

THEOREM B. *The distribution of the total cost C on $\Sigma_N(\varepsilon)$ is asymptotically Gaussian, with speed of convergence $O(1/\sqrt{\log N})$.*

REMARK 4.2. The smoothing process for Ω_N and $\Omega_N(\varepsilon)$ in [1] is possible, since $\Omega_{N'}$ is a subset of Ω_N for $N' < N$. Thus, one can compare the different probabilities $\mathbb{P}_{N,\varepsilon}$ and $\mathbb{P}_N$ by only dealing with subsets $A_{N,\varepsilon}$ in $\Omega_N(\varepsilon)$ which naturally come from Ω_N. However, there is no such inclusion among Σ_N's. Hence, in order to have the statistical indistinguishability of the probabilistic models Σ_N and $\Sigma_{N,\varepsilon}$, we should specify which sets should be compared. It is sufficient to consider $C^{-1}(S)$ in Σ_N and $\Sigma_N(\varepsilon)$ for any $S \subset \mathbb{R}_{\geq 0}$.

Let us denote by $\mathbb{P}_{N,\varepsilon}$ the probability on $\Sigma_N(\varepsilon)$ with uniform density. It is now natural to ask:

QUESTION C. *For any $S \subset \mathbb{R}_{\geq 0}$, do we have*

$$|\mathbb{P}_{N,\varepsilon}(C \in S) - \mathbb{P}_N(C \in S)| = O(\varepsilon(N))?$$

The conjecture on the asymptotic Gaussian distribution of the length ℓ of continued fractions on Σ_N from the introduction immediately follows from an affirmative answer to this question since we can deduce (2.4) for $\Sigma_N(\varepsilon)$ from Theorem B, which enables us to estimate $\mathbb{P}_{N,\varepsilon}(\ell = k)$ for any integer k. However, at this time there is no explicit idea how to study $\mathbb{P}_N$ on Σ_N. Some computations in [4] suggest that to answer this question it may be necessary to obtain a fundamental result on properties of ℓ on Σ_N.

Acknowledgements. We are grateful to Valérie Berthé for careful comments and conversations about the topic. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT)(NRF-2017R1A2B4012408).

References

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

[2] E. Cesaratto, *A note on "Euclidean algorithms are Gaussian" by V. Baladi and B. Vallée*, J. Number Theory 129 (2009), 2267–2273.

[3] D. Dolgopyat, *On decay of correlations in Anosov flows*, Ann. of Math. (2) 147 (1998), 357–390.
[4] S. R. Finch, *Mathematical Constants*, Cambridge Univ. Press, 2003.
[5] D. Hensley, *The number of steps in the Euclidean algorithm*, J. Number Theory 49 (1994), 142–182.
[6] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., Clarendon Press, Oxford, 1986.

Jungwon Lee, Hae-Sang Sun
Department of Mathematical Sciences
Ulsan National Institute of Science and Technology
Ulsan, Korea
E-mail: jungwon@unist.ac.kr
haesang@unist.ac.kr