Uniform Distribution Theory **6** (2011), no.2, 131–141

SOME METRICAL RESULTS ON THE APPROXIMATION BY CONTINUED FRACTIONS

Hendrik Jager

ABSTRACT. Let x be a real irrational number and p_n/q_n, $n = 1, 2, \ldots$ its sequence of continued fraction convergents. Define $d_n = q_{n+1}\left|q_n x - p_n\right|$. For almost all x the distribution functions of the sequences $\left|d_n - d_{n+1}\right|$ and $d_n + d_{n+1}$ are determined.

Communicated by Cor Kraaikamp

1. Introduction

Let x be a real irrational number with continued fraction expansion

$$x = [a_0; a_1, a_2, \ldots] \quad \text{and} \quad p_n/q_n,\ n = 1, 2, \ldots$$

its sequence of convergents.

The sequence $\theta_n,\ n = 1, 2, \ldots$ of approximation coefficients of x is defined by

$$\theta_n = q_n^2 \left| x - \frac{p_n}{q_n} \right|,\ n = 1, 2, \ldots \tag{1}$$

One of the most important aspects of the approximation by continued fractions is the fact that this is always a sequence in the unit interval.

At the basis of many results on the distribution of these coefficients θ_n lies the following fundamental metrical result.

Theorem 1. *Denote by Δ the unit triangle in the (α, β)-plane, that is the triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$.*

2010 Mathematics Subject Classification: 11K50, 11K31.
Keywords: Continued fraction, convergent, distribution function.

For every x the sequence (θ_n, θ_{n+1}) is a sequence in Δ and for almost all x it is distributed over Δ according to the density function μ, where

$$\mu(\alpha,\beta) = \frac{1}{\log 2}\,\frac{1}{\sqrt{1-4\alpha\beta}}.$$

For a proof see [2, section 5.3] or [3].

From Theorem 1 a proof of the Doeblin-Lenstra conjecture on the distribution of the sequences $\theta_n,\ n = 1, 2\dots$, see [1], follows easily. Also the distribution functions, for almost all x, of the sequences

$$\theta_n + \theta_{n+1},\ \left|\theta_n - \theta_{n+1}\right| \quad \text{and} \quad \theta_n\theta_{n+1},\ n = 1, 2, \dots$$

can be derived from it, see [3].

Less is known about another type of approximation coefficients, the sequences $d_n,\ n = 1, 2, \dots$, defined by

$$d_n = q_n q_{n+1}\left|x - \frac{p_n}{q_n}\right|,\ n = 1, 2, \dots \tag{2}$$

For every irrational x this is a sequence in the interval $[\frac{1}{2}, 1]$ with, for almost all x, a distribution function F where

$$F(z) = \frac{1}{\log 2}\big(z\log z + (1-z)\log(1-z) + \log 2\big), \tag{3}$$

see [1, Theorem 4].

The sequences from (1) and (2) are related by the fact that the two-dimensional sequence

$$(\theta_n, d_n),\ n = 1, 2, \dots,$$

is for all irrational x a sequence in the interior of the triangle in the (α, β)-plane with vertices $(0, 1)$, $(\frac{1}{2}, \frac{1}{2})$ and $(1, 1)$ and that for almost all x this sequence is distributed over this triangle according to the density function

$$\frac{1}{\log 2}\,\frac{1}{\alpha}, \tag{4}$$

see [3, Theorem 7].

The distribution (4) also yields an easy proof of the Doeblin-Lenstra conjecture. Other consequences of (4) are the distribution (3) of the d_n's and the uniform distribution in the unit interval of the sequence $d_n - \theta_n,\ n = 1, 2, \dots$ see [3, Theorem 9]. Hence the mean of the d_n's, $\mathcal{M}(d_n)$, differs $\frac{1}{2}$ from $\mathcal{M}(\theta_n)$. As $\mathcal{M}(\theta_n) = \frac{1}{4\log 2}$, see [1, Corollary 2], we thus have

$$\mathcal{M}(d_n) = \frac{1}{2} + \frac{1}{4\log 2} = 0.86067\dots, \quad \textit{a.e.}$$

By a simple transformation of the Doeblin-Lenstra conjecture one obtains the distribution function of the sequences $\log\theta_n, n = 1, 2, \ldots,$ over the interval $(-\infty, 0)$ and with this one easily shows that

$$\mathcal{M}(\log\theta_n) = -1 - \frac{1}{2}\log 2 = -1.34657\ldots, \quad a.e. \tag{5}$$

The step to

$$\mathcal{M}(\log d_n) = -1 - \frac{1}{2}\log 2 + \frac{\pi^2}{12\log 2} = -0.16000\ldots, \quad a.e.,$$

then follows from Lévy's celebrated result

$$\lim_{n\to\infty}\frac{1}{n}\log q_n = \frac{\pi^2}{12\log 2}, \quad a.e.,$$

because

$$\log d_n - \log\theta_n = \log q_{n+1} - \log q_n.$$

To obtain more information about the distribution properties of the sequences from (2) one would like a result similar to Theorem 1. This was given in [4] and reads as follows.

Theorem 2. *Let a be a positive integer and denote by R_a the quadrangle in the (α,β)-plane with vertices*

$$(1,1),\ \big(a/(a+1),1\big), \big((a+1)/(a+2),(a+1)/(a+2)\big) \quad and \quad \big(1,a/(a+1)\big).$$

Consider the set Ω as the union of the R_a, $a = 1, 2, \ldots,$ piled one upon another, where the edge from

$$\big((a+1)/(a+2),(a+1)/(a+2)\big) \text{ to } \big(1,a/(a+1)\big) \text{ of } R_a$$

is identified, by vertical projection, with the edge from

$$\big((a+1)/(a+2),1\big) \text{ to } (1,1) \text{ of } R_{a+1}.$$

Then, for every x, the sequence (d_n, d_{n+1}) $n = 1, 2, \ldots$ lies in Ω; more precisely

$$(d_n, d_{n+1}) \in R_a \text{ if and only if } a_{n+2} = a.$$

For almost all x the sequence is distributed over Ω according to the density function ν, where

$$\nu(\alpha,\beta) = \frac{1}{\log 2}\,\frac{1}{\alpha+\beta-1}.$$

A consequence of the first part of this theorem is that when a d_n is small, i.e., close to the left end point of the interval $\left[\frac{1}{2}, 1\right]$, its successor d_{n+1} is close to the other end point of the interval. In the case of the coefficients θ_n the situation is just the opposite. When a θ_n is close to 1, its successor, and also its predecessor, are close to 0. Therefore it is of interest to study the distribution of the sequences

$$\big|\theta_n - \theta_{n+1}\big| \quad \text{and} \quad \big|d_n - d_{n+1}\big|,\ n = 1, 2, \ldots \tag{6}$$

For the first sequence of (6) this was done in [3]. One has for instance for almost all x :

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \big|\theta_n - \theta_{n+1}\big| = \frac{4-\pi}{4\log 2} = 0.30960\ldots \tag{7}$$

The only application of Theorem 2 in [4] was an alternative proof of (3). The object of this paper is to use Theorem 2 to obtain the distribution function for the second sequence of (6) and a result similar to (7) for the d_n. Further we determine the distribution of the sequences $d_n + d_{n+1},\ n = 1, 2, \ldots$, for almost all x.

2. The distribution of the sequence $\big|d_n - d_{n+1}\big|$

Theorem 3. *Put*

$$m = m(z) = \left\lfloor \frac{1}{z} \right\rfloor, \qquad z > 0,$$

and define the function F on the interval $\left[0, \frac{1}{2}\right]$ by

$$\begin{aligned} F(z) = \frac{1}{\log 2}\Bigg(& z \log \frac{m(m+1)}{2} \\ & - (m-1)\big((1+z)\log(1+z) + (1-z)\log(1-z)\big) \\ & + \log\left(1 + \frac{1}{m}\right)\Bigg), \qquad 0 < z \le \frac{1}{2}, \\ F(0) = 0. & \end{aligned}$$

This function F is monotonically increasing from 0 in $z = 0$ to 1 in $z = \frac{1}{2}$ and has a continuous derivative.

It is for almost all x the distribution function of the sequence

$$\big|d_n - d_{n+1}\big|, \quad n = 1, 2, \ldots$$

The form of F looks perhaps unpleasant but when one plots the graph, one gets a nice, smooth curve, see Figure 1.

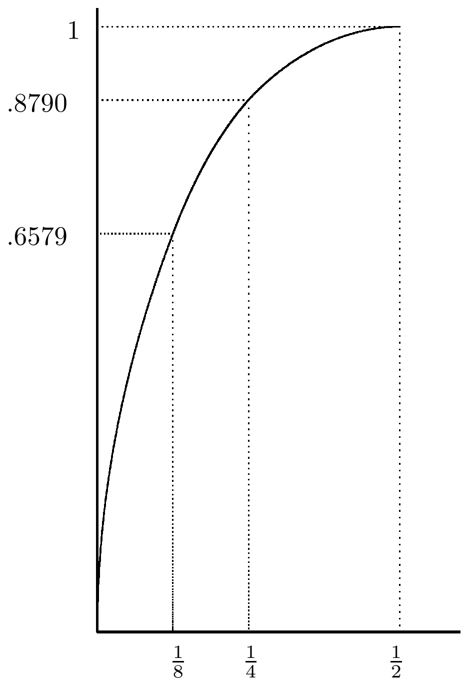

FIGURE 1. Distribution function of $|d_n - d_{n+1}|$.

EXAMPLES.

$$F\left(\frac{1}{4}\right) = \frac{41}{4} - \frac{1}{\log 2}\left(\frac{9}{4}\log 3 + \frac{5}{2}\log 5\right) = 0.87901\ldots,$$

$$F'\left(\frac{1}{4}\right) = 1 + \frac{1}{\log 2}\log\frac{27}{25} = 1.11103\ldots,$$

$$F\left(\frac{1}{8}\right) = \frac{157}{4} - \frac{1}{\log 2}\left(\frac{27}{2}\log 3 + \frac{49}{8}\log 7\right) = 0.65795\ldots$$

Proof. Let a be a positive integer. Denote by $R_a(z)$, with $0 < z < 1/(a+1)$, that part of R_a for which $|\alpha - \beta| < z$. Then, in view of Theorem 2, one has, for almost all x, that

$$\lim_{N\to\infty} \frac{1}{N}\#\{n; n \le N, |d_n - d_{n+1}| < z, a_{n+2} = a\} = \nu(R_a(z)).$$

Denote this expression by $F_a(z)$. After a tedious, but otherwise completely elementary calculation, one finds that for $0 < z < 1/(a+1)$

$$F_a(z) = \frac{1}{\log 2}\left(z \log \frac{a+2}{a} - (1+z)\log(1+z) - (1-z)\log(1-z)\right).$$

Define $F_a(z)$ on the interval $[1/(a+1], 1/2]$ by

$$F_a(z) = F_a\left(\frac{1}{a+1}\right) = \frac{1}{\log 2} \log \frac{(a+1)^2}{a(a+2)}.$$

One easily verifies that F_a has a continuous derivative on $[0, 1/2]$. The function F, defined by

$$F(z) = \sum_{a=1}^{\infty} F_a(z), \qquad 0 \le z \le \frac{1}{2},$$

is the distribution function of the sequence $\left|d_n - d_{n+1}\right|$. It can be written in the closed form given in the statement of the theorem. □

Theorem 4. *For almost all x one has*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left|d_n - d_{n+1}\right| = \frac{1-\gamma}{\log 2} - \frac{1}{2} = 0.10994\ldots,$$

where γ is Euler's constant.

Proof. Denote the first moment of F_a by M_a. Then

$$M_a = \int_0^{\frac{1}{a+1}} zF_a'(z)\mathrm{d}z = \frac{1}{\log 2}\left(\frac{1}{2}\log\frac{a+2}{a} - \frac{1}{a+1}\right).$$

Summing over a yields

$$\lim_{A\to\infty} \sum_{a=1}^{A} M_a = \frac{1}{\log 2} \lim_{A\to\infty} \left(\frac{1}{2}\log\frac{1}{2}(A+1)(A+2) - \sum_{a=2}^{A+1}\frac{1}{a}\right).$$

In this we substitute the classical

$$\sum_{a=2}^{A+1} \frac{1}{a} = \log(A+1) + \gamma - 1 + o(1), \quad A \to \infty,$$

and obtain, for almost all x, the expectation of $\left|d_n - d_{n+1}\right|$ as given by the theorem. □

3. The distribution of the sequence $d_n + d_{n+1}$

We introduce the following intervals:

$$\Delta_m^{(0)} = \left[\frac{2m+1}{m+1}, \frac{4m+4}{2m+3}\right), \quad m = 1, 2, \dots \tag{8}$$

and

$$\Delta_m^{(1)} = \left[\frac{4m+4}{2m+3}, \frac{2m+3}{m+2}\right), \quad m = 0, 1, 2, \dots \quad . \tag{9}$$

These intervals are pairwise disjoint. Their natural order is

$$\Delta_0^{(1)},\ \Delta_1^{(0)},\ \Delta_1^{(1)},\ \Delta_2^{(0)},\ \Delta_2^{(1)},\ \Delta_3^{(0)},\ \Delta_3^{(1)}, \dots$$

and together they fill up the interval $\left[\frac{4}{3}, 2\right)$.

Put

$$\Delta^{(0)} = \bigcup_{m=1}^{\infty} \Delta_m^{(0)} \quad \text{and} \quad \Delta^{(1)} = \bigcup_{m=0}^{\infty} \Delta_m^{(1)}.$$

The sets $\Delta^{(0)}$ and $\Delta^{(1)}$ have Lebesgue measure

$$\frac{5}{3} - 2\log 2 = 0.28037\dots \quad \text{and} \quad 2\log 2 - 1 = 0.38629\dots,$$

respectively.

Next we define for every $z \in \left[\frac{4}{3}, 2\right)$ two integers m and n by

$$m = m(z) = \left\lfloor \frac{z-1}{2-z} \right\rfloor \quad \text{and} \quad n = n(z) = \left\lfloor \frac{2z-2}{2-z} \right\rfloor.$$

Clearly, $n = 2m$ or $n = 2m+1$. Write $\lambda(z) = n - 2m$. The intervals from (8) and (9) are constructed in such a way that λ is the characteristic function of $\Delta^{(1)}$.

Finally, we use the abbreviation

$$h(z) = h(m,n) = \sum_{k=m+1}^{n} \frac{1}{k}.$$

Theorem 5. *Let F be the function defined on $\left[\frac{4}{3}, 2\right]$ by*

$$F(z) = \frac{1}{\log 2}\Big(2h(z)(z-1) + \lambda(z)\big(z - \log(z-1) - 2\big) + \log(m+1) - \log\big(n+1+\lambda(z)\big)\Big), \quad \frac{4}{3} \le z < 2,$$
$$F(2) = 1.$$

This function F is monotonically increasing from 0 in $z = \frac{4}{3}$ to 1 in $z = 2$ and has a continuous derivative.

It is for almost all x the distribution function of the sequence

$$d_n + d_{n+1}, \quad n = 1, 2, \ldots$$

Remark. On each interval $\Delta_m^{(0)}$ from (8), F is the simple linear function

$$\frac{1}{\log 2}\Big(2h(m,2m)(z-1) + \log\frac{m+1}{2m+1}\Big).$$

For example,

$$F(z) = \frac{1}{\log 2}\big(z - 1 + \log 2 - \log 3\big), \qquad z \in \Delta_1^{(0)} = \left[\frac{3}{2}, \frac{8}{5}\right),$$
$$F(z) = \frac{1}{\log 2}\left(\frac{7}{6}(z-1) + \log 3 - \log 5\right), \quad z \in \Delta_2^{(0)} = \left[\frac{5}{3}, \frac{12}{7}\right).$$

On an interval from (9) there are only minor changes from this. The coefficient $h(m,2m)$ is replaced by $h(m,2m+1)$ which means that a term $1/(2m+1)$ is added to it; the term $\log\big((m+1)/(2m+1)\big)$ changes into $\log\big((m+1)/(2m+3)\big)$ and finally the term

$$\frac{1}{\log 2}\big(z - \log(z-1) - 2\big),$$

which is asymptotically $\frac{1}{2\log 2}(z-2)^2$ for $z \to 2$, is added.

For example,

$$F(z) = \frac{1}{\log 2}\,\big(3z - \log(z-1) - 4 - \log 3\big), \qquad z \in \Delta_0^{(1)} = \left[\frac{4}{3}, \frac{3}{2}\right),$$
$$F(z) = \frac{1}{\log 2}\left(\frac{8}{3}z - \log(z-1) - \frac{11}{3} + \log 2 - \log 5\right), \quad z \in \Delta_1^{(1)} = \left[\frac{8}{5}, \frac{5}{3}\right).$$

On an interval $\Delta_m^{(0)}$ from (8) one has

$$F'(z) = \frac{1}{\log 2}\,2h(z)$$

and on an interval $\Delta_m^{(1)}$ from (9)

$$F'(z) = \frac{1}{\log 2}\left(2h(z) + 1 - \frac{1}{z-1}\right).$$

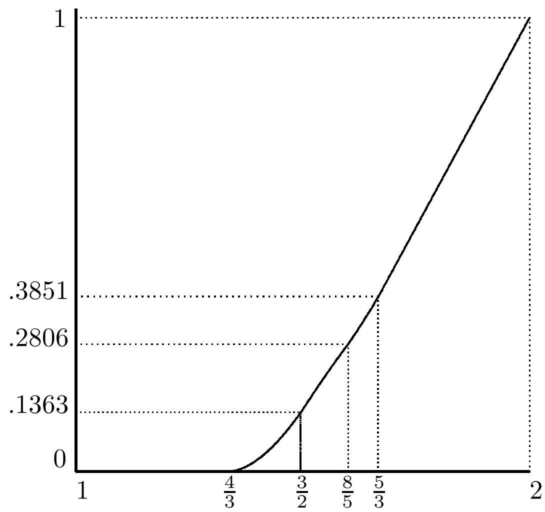

FIGURE 2. Distribution function of $d_n + d_{n+1}$.

From this and from

$$\lim_{z \uparrow 2} h(z) = \log 2$$

it follows that

$$\lim_{z \uparrow 2} F'(z) = 2.$$

Further,

$$\lim_{z \downarrow \frac{4}{3}} F'(z) = 0.$$

Between 5/3 and 2, the graph of F is on the scale of Figure 2 not distinguishable from a line segment.

P r o o f. Denote by $R_a(z)$ that part of R_a which lies under the line $\alpha + \beta = z$. Then

$$\lim_{N \to \infty} \frac{1}{N} \#\{n; n \le N, d_n + d_{n+1} < z, a_{n+2} = a\} = \nu\big(R_a(z)\big).$$

Call this expression $F_a(z)$ and consider it as a function on the interval $\left[\frac{4}{3}, 2\right]$. Divide this interval into three subintervals, $I_a^{(0)}$, $I_a^{(1)}$ and $I_a^{(2)}$ by the two points

$(2a+2)/(a+2)$ and $(2a+1)/(a+1)$. Some elementary integration leads to

$$F_a(z) = \begin{cases} 0, & z \in I_a^{(0)}, \\ \dfrac{1}{\log 2}\left(\dfrac{a+2}{a}(z-1) - \log(z-1) + \log\dfrac{a}{a+2} - 1\right), & z \in I_a^{(1)}, \\ \dfrac{1}{\log 2}\left(-z + \log(z-1) + 2 + \log\dfrac{(a+1)^2}{a(a+2)}\right), & z \in I_a^{(2)}. \end{cases}$$

To get the expression for F_a on $I_a^{(1)}$ one has to integrate over two separate regions, one to the left and one to the right of the line $\alpha = (a+1)/(a+2)$.

The easiest way to calculate the form of F_a on $I_a^{(2)}$ is to integrate over the complement of $R_a(z)$ and to subtract this from $\nu(R_a)$, which equals, as is well-known,

$$\frac{1}{\log 2}\log\frac{(a+1)^2}{a(a+2)}\cdot$$

To obtain the required distribution function we must sum over a. We take a z and observe that

$$z \in I_{m+1}^{(1)}, I_{m+2}^{(1)}, I_{m+3}^{(1)}, \ldots, I_n^{(1)}$$

and that z is not contained in any other interval $I_a^{(1)}$.

Hence

$$F(z) = \frac{1}{\log 2}\sum_{a=1}^{m}\left(-z + \log(z-1) + 2 + \log\frac{(a+1)^2}{a(a+2)}\right) + \frac{1}{\log 2}\sum_{a=m+1}^{n}\left(\frac{a+2}{a}(z-1) - \log(z-1) + \log\frac{a}{a+2} - 1\right),$$

and after some rearrangements one obtains the form as given in the statement of the theorem. □

REFERENCES

[1] BOSMA, W. – JAGER, H. – WIEDIJK, F.: *Some metrical observations on the approximation by continued fractions*, Nederl. Akad. Wetensch. Indag. Math. **45**, (1983), no. 3, 281–299.

[2] DAJANI, K. – KRAAIKAMP, C.: *Ergodic Theory of Numbers*, The Carus Mathematical Monographs **29**, The Mathematical Association of America, Washington, DC, 2002.

[3] JAGER, H.: *The distribution of certain sequences connected with the continued fraction*, Nederl. Akad. Wetensch. Indag. Math. **48** (1986), no. 1, 61–69.
[4] KRAAIKAMP, C.: *On symmetric and asymmetric diophantine approximation by continued fractions*, J. Number Theory **46** (1994), no. 2, 137–157.

Received September 10, 2010
Accepted October 19, 2010

Hendrik Jager
Oude Larenseweg 26
7214 PC Epse
the NETHERLANDS
E-mail: epserbos@xs4all.nl