A NOTE
ON THE CONTINUED FRACTION OF MINKOWSKI

HENDRIK JAGER

ABSTRACT. Denote by $\Theta_1, \Theta_2, \cdots$ the sequence of approximation coefficients of Minkowski's diagonal continued fraction expansion of a real irrational number $x$. For almost all $x$ this is a uniformly distributed sequence in the interval $[0, \frac{1}{2}]$. The average distance between two consecutive terms of this sequence and their correlation coefficient are explicitly calculated and it is shown why these two values are close to $1/6$ and $0$, respectively, the corresponding values for a random sequence in $[0, \frac{1}{2}]$.

Communicated by Cor Kraaikamp

Let $x$ be a real irrational number and $p_n/q_n$, $n = 1, 2, \cdots$ its sequence of regular continued fraction convergents.

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

$$\theta_n = q_n^2 \left| x - \frac{p_n}{q_n} \right|, \ n = 1, 2, \cdots \quad (1)$$

is always a sequence in the unit interval. The subsequence of (1), consisting of all elements $\theta_n$ with $\theta_n < \frac{1}{2}$, is denoted here by

$$\Theta_n, \ n = 1, 2, \cdots \quad (2)$$

This is the sequence of approximation coefficients of Minkowski's diagonal continued fraction, see [5] and [6]. It is for almost all $x$ uniformly distributed in the interval $[0, \frac{1}{2}]$, which follows from the Doeblin-Lenstra conjecture proved in [11] or, without ergodic theoretical methods, from a result of Erdős, [2], as pointed out by Ito and Nakaeda in [3].

2010 Mathematics Subject Classification: 11K50.
Keywords: Continued fractions, approximation coefficients, metrical theory.
Reading the paper \cite{PillichshammerSteinerberger} by Pillichshammer and Steinerberger on the average distance between consecutive terms of uniformly distributed sequences, it occurred to the author that the sequences \cite{Minkowski} are also ‘typical examples of uniformly distributed sequences’ and he wondered what the average distance or expectation

\[ E(|\Theta_{n+1} - \Theta_n|) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |\Theta_{n+1} - \Theta_n| \]

might be for them.

The answer turned out to be an explicitly calculable constant with a numerical value \(0.164017 \cdots\), for almost all \(x\). This is close to \(\frac{1}{6}\), the value for a random sequence in \([0, \frac{1}{2}]\), as remarked by Professor Steinerberger to whom the author mentioned the result. This remark motivated the calculation of the correlation between two consecutive elements of \cite{Minkowski}. It yielded, for almost all \(x\), the value \(0.019290 \cdots\) which is again close to the corresponding value for a random sequence, namely 0. More precisely, we shall prove the following theorem.

**Theorem.** Denote by \(\Theta_1, \Theta_2, \cdots\)

the sequence of approximation coefficients of Minkowski’s diagonal continued fraction expansion of a real irrational number \(x\).

By definition this is always a sequence in \([0, \frac{1}{2}]\). For almost all \(x\) these sequences show the following three similarities with a random sequence in \([0, \frac{1}{2}]\):

1. They are uniformly distributed in the interval \([0, \frac{1}{2}]\);
2. The average distance between two consecutive terms is close to \(\frac{1}{6}\), namely,

\[ \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} |\Theta_{n+1} - \Theta_n| = \frac{25}{24} + \log 2 - \frac{\pi}{2} = 0.164017 \cdots, \quad (3) \]

3. The correlation coefficient \(R(\Theta_n, \Theta_{n+1})\) between two consecutive terms is almost 0, namely,

\[ R(\Theta_n, \Theta_{n+1}) = \frac{111}{5} - 32 \log 2 = 0.019290 \cdots \]

**Proof of the theorem.**

1. Where to find a proof of part 1 of the theorem was already mentioned.

2. For a proof of part 2 of the theorem we turn to the fundamental paper by Kraaikamp \cite{Kraaikamp}, on Minkowski’s expansion. The casual remark on page 207: “In a similar way one could determine the distribution of the sequence \((\Theta_{k-1} - \Theta_k)_{k \geq 0}\) on the interval \([-\frac{1}{2}, \frac{1}{2}]\)” tells one how this mean value can be calculated.
A NOTE ON THE CONTINUED FRACTION OF MINKOWSKI

Let \( 0 < z < \frac{1}{2} \). Denote by \( \Omega_1(z) \) that part of the first quadrant of the \((x, y)\)-plane for which
\[
0 < x < \frac{1}{2}, \quad y < x, \quad y > x - z,
\]
and by \( \Omega_2(z) \) that part of \( \Omega_1(z) \) which lies below the parabola
\[
(x - y)^2 + x + y = \frac{3}{4}.
\]
We now apply Kraaikamp’s Theorem 4.2: Define \( F_1 \) by
\[
F_1(z) = 4 \int \int_{\Omega_1(z)} \frac{dx \, dy}{\sqrt{1 - 4xy}}
\]
and \( F_2 \) by
\[
F_2(z) = 4 \int \int_{\Omega_2(z)} \frac{dx \, dy}{\sqrt{1 + 4xy}}.
\]
Then
\[
F = F_1 + F_2
\]
is the distribution function, for almost all \( x, \) of the sequence
\[
|\Theta_{n+1} - \Theta_n|, \ n = 1, 2, \ldots
\]
We attack both integrals with the substitution
\[
\xi = x + y, \quad \eta = x - y
\]
which has a determinant of Jacobi equal to \(-2\). We then find that
\[
F_1(z) = 2 \int \int_{\Omega_1^*(z)} \frac{d\xi \, d\eta}{\sqrt{1 - \xi^2 + \eta^2}}
\]
and
\[
F_2(z) = 2 \int \int_{\Omega_2^*(z)} \frac{d\xi \, d\eta}{\sqrt{1 + \xi^2 - \eta^2}},
\]
where \( \Omega_1^*(z) \) and \( \Omega_2^*(z) \) are the images of \( \Omega_1(z) \) and \( \Omega_2(z) \). Hence
\[
F_1(z) = 2 \int_0^z \left( \int_{\xi=\eta}^{\xi=1-\eta} \frac{d\xi}{\sqrt{1 - \xi^2 + \eta^2}} \right) \, d\eta = 2 \int_0^z (\arctan \frac{1 - \eta}{\sqrt{2\eta}} - \arctan \eta) \, d\eta
\]
from which we see that
\[
F_1'(z) = 2 \left( \arctan \frac{1 - z}{\sqrt{2z}} - \arctan z \right)
\]
and from which it follows that
\[ F_1(z) = 2\left(\sqrt{2z} + \log(1 + z - \sqrt{2z}) - z \arctan z + z \arctan \frac{1 - z}{\sqrt{2z}}\right). \]

In a similar way one finds
\[ F'_2(z) = 2\left(\log 2 + \log(1 - z)\right) \] (5)
and
\[ F_2(z) = 2\left(z \log 2 - z - (1 - z) \log(1 - z)\right). \] (6)
Thus we have obtained an explicit expression for the distribution function \( F \) from (4) on the interval \([0, \frac{1}{2}]\).

Now the rather complicated function \( F \) is on this interval surprisingly well approximated by the parabola \( G \)
\[ G(z) = -4z^2 + 4z. \] (7)
If one plots both graphs on an 27-inch computer screen, one hardly sees the difference. On \([0, \frac{1}{2}]\) one has
\[ 0 < F(z) - G(z) < 0.012 \]
and on \([\frac{1}{4}, \frac{1}{2}]\) the difference is even smaller than 0.006.

Let \( \xi_1, \xi_2, \xi_3, \cdots \) be a hypothetical sequence, uniformly distributed in the interval \([0, \frac{1}{2}]\), for which the sequence
\[ |\xi_2 - \xi_1|, |\xi_3 - \xi_2|, |\xi_4 - \xi_3|, \cdots, \]
has the \( G \) from (7) as distribution function. Then
\[ \mathcal{E}(|\xi_{n+1} - \xi_n|) = \frac{1}{6}. \]
Further
\[ \mathcal{E}((\xi_{n+1} - \xi_n)^2) = \frac{1}{24}, \]
and from this and from
\[ \mathcal{E}(\xi_n^2) = \frac{1}{12}, \]
it follows that
\[ \mathcal{E}(\xi_n \xi_{n+1}) = \frac{1}{16}. \]
So
\[ \mathcal{E}(\xi_n \xi_{n+1}) = \mathcal{E}(\xi_n)\mathcal{E}(\xi_{n+1}), \]
consequently the correlation coefficient between \( \xi_n \) and \( \xi_{n+1} \), denoted as
\[ \mathcal{R}(\xi_n, \xi_{n+1}), \] is 0.
A NOTE ON THE CONTINUED FRACTION OF MINKOWSKI

Therefore one may expect that the limit (3) is close to $\frac{1}{6}$ and that the correlation $\mathcal{R}(\Theta_n, \Theta_{n+1})$ between two consecutive approximation coefficients of Minkowski’s continued fraction is almost 0.

After the observation of the similarity between the obtained distribution function $F$ and a parabola and of some properties of a hypothetical sequence with this parabola as distribution function for the distances of consecutive terms, we finish the proof of part 2 of the theorem.

The first moment of the distribution $F$ equals

$$2 \int_0^{\frac{1}{2}} z \left( \arctan \frac{1-z}{\sqrt{2}z} - \arctan z \right) dz + 2 \int_0^{\frac{1}{2}} z \left( \log 2 + \log(1-z) \right) dz.$$

These two integrals are

$$\frac{5}{3} - \frac{\pi}{2} \quad \text{and} \quad \log 2 - \frac{5}{8},$$

respectively, which proves part 2 of the theorem.

3. For a proof of part 3 of the theorem we need to calculate

$$\mathcal{E}(\Theta_{n+1} - \Theta_n)^2 = \int_0^{\frac{1}{2}} z \frac{d}{dz} F(\sqrt{z}) dz. \quad (9)$$

Elementary but tedious calculations show that this integral equals

$$-\frac{53}{60} + \frac{4}{3} \log 2.$$

The reader may verify this with Mathematica (details are also available, on request via e-mail, from the author).

It follows that

$$\mathcal{E}(\Theta_n \Theta_{n+1}) = \frac{21}{40} - \frac{2}{3} \log 2,$$

which then yields

$$\mathcal{R}(\Theta_n, \Theta_{n+1}) = \frac{111}{5} - 32 \log 2 = 0.019290 \cdots,$$

positive but, as expected, close to 0. \hfill \Box

In [4] approximation coefficients connected with the nearest medians are added to the sequence [1] which yields in a natural way a uniformly distributed sequence in $[0, 1]$. It would be very interesting to find corresponding results for this sequence, in particular to see whether it is also very close, perhaps even closer, to the hypothetical series [8].
REFERENCES


Received October 10, 2016
Accepted February 15, 2017

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