COLLOQUIUM MATHEMATICUM

VOL. 79 1999 NO. 1

ON THE METRIC THEORY OF CONTINUED FRACTIONS

BY

JOËL RIVAT (LYON)

Introduction. For any positive integer n we denote by $P(n)$ the Lebesgue measure of the set of irrational numbers $x \in (0,1)$ whose closest rational approximation with denominator $\le n$ is a convergent of the continued fraction expansion of x.

The question of the behaviour of $P(n)$ was asked by M. Deléglise to A. Schinzel. Recently I. Aliev, S. Kanemitsu and A. Schinzel [1] proved that

$$P(n) = \frac{1}{2} + \frac{6}{\pi^2}(\log 2)^2 + O\left(\frac{1}{n}\right).$$

In this article we shall improve this result to the following

THEOREM. *There exists $c > 0$ such that*

$$P(n) = \frac{1}{2} + \frac{6}{\pi^2}(\log 2)^2 + O\left(\frac{1}{n}\exp\left(-c\,\frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\right)\right). \tag{1}$$

Under the Riemann hypothesis we have

$$P(n) = \frac{1}{2} + \frac{6}{\pi^2}(\log 2)^2 + O(n^{-4/3+\varepsilon}). \tag{2}$$

REMARK. I. Aliev, S. Kanemitsu and A. Schinzel [1] also note that the main term, but not the error term, can be derived from Theorem 1.3 of P. Kargaev and A. Zhigljavsky [2].

Classical results. We denote by $\lfloor x\rfloor$ the greatest integer not exceeding x and write $\psi(x) = x - \lfloor x\rfloor - 1/2$.

LEMMA 1. *Let f be a function with a continuous derivative in the interval $[a,b]$. Then*

$$\sum_{a<n\le b} f(n) = \int_a^b f(x)\,dx + \psi(a)f(a) - \psi(b)f(b) + \int_a^b \psi(x)f'(x)\,dx.$$

Proof. See for example Titchmarsh [4], formula 2.1.2, page 13. ∎

1991 *Mathematics Subject Classification*: Primary 11K50.

This work was started while the author was visiting the Polish Academy of Sciences in Warsaw.

Applying this lemma and writing $\phi(x)=\psi(x)/x$, $\log^+ x=\max(\log x,1)$ we obtain

LEMMA 2. *For arbitrary positive numbers $a<b$ we have*

$$\sum_{a<n\le b}\frac{1}{n}=\log b-\log a+\phi(a)-\phi(b)+O\Big(\frac{1}{a^2}\Big),$$

$$\sum_{a<n\le b}\frac{\log n}{n}=\frac{(\log b)^2-(\log a)^2}{2}+\phi(a)\log a-\phi(b)\log b+O\Big(\frac{\log^+ a}{a^2}\Big),$$

$$\sum_{a<n\le b}\frac{1}{n^2}=\frac{1}{a}-\frac{1}{b}+O\Big(\frac{1}{a^2}\Big).$$

COROLLARY 1. *For any positive number x, we have*

$$\sum_{x/2<k\le x}\frac{1}{k}=\log 2+\phi\Big(\frac{x}{2}\Big)-\phi(x)+O\Big(\frac{1}{x^2}\Big),$$

$$\sum_{x/2<k\le x}\frac{\log k}{k}=\log\Big(\frac{x}{2}\Big)\log 2+\frac{(\log 2)^2}{2}+\phi\Big(\frac{x}{2}\Big)\log\Big(\frac{x}{2}\Big)-\phi(x)\log x+O\Big(\frac{\log^+ x}{x^2}\Big),$$

$$\sum_{x/2<k\le x}\frac{1}{k^2}=\frac{1}{x}+O\Big(\frac{1}{x^2}\Big).$$

LEMMA 3. *There exists $c>0$ such that for any $x\ge 1$ we have*

$$\sum_{1\le d\le x}\frac{\mu(d)}{d}=O\Big(\exp\Big(-c\,\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\Big)\Big),$$

$$\sum_{1\le d\le x}\frac{\mu(d)}{d^2}=\frac{6}{\pi^2}+O\Big(\frac{1}{x}\exp\Big(-c\,\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\Big)\Big).$$

Under the Riemann hypothesis, for any $x\ge 1$ we have

$$\sum_{1\le d\le x}\frac{\mu(d)}{d}=O(x^{-1/2+\varepsilon}),\qquad \sum_{1\le d\le x}\frac{\mu(d)}{d^2}=\frac{6}{\pi^2}+O(x^{-3/2+\varepsilon}).$$

Proof. By partial summation, for any $1\le x\le y$ we have

$$\sum_{x<d\le y}\frac{\mu(d)}{d}=\frac{1}{y}\sum_{x<d\le y}\mu(d)+\int_x^y\Big(\sum_{x<d\le t}\mu(d)\Big)\frac{dt}{t^2},$$

$$\sum_{x<d\le y}\frac{\mu(d)}{d^2}=\frac{1}{y^2}\sum_{x<d\le y}\mu(d)+2\int_x^y\Big(\sum_{x<d\le t}\mu(d)\Big)\frac{dt}{t^3}.$$

By Satz 3 of A. Walfisz [5], page 191, there exists $c' > 0$ such that

$$\sum_{1\le d\le x} \mu(d) = O\biggl(x \exp\biggl(-c' \frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\biggr)\biggr).$$

Writing

$$\delta(t) = \exp\biggl(-c' \frac{(\log t)^{3/5}}{(\log\log t)^{1/5}}\biggr)$$

we have $\sum_{x<d\le t} \mu(d) \ll t\delta(t)$, hence

$$\begin{aligned}\sum_{x<d\le y} \frac{\mu(d)}{d} &\ll \delta(y) + \int_x^y \delta(t)\,\frac{dt}{t}\\ &\ll \delta(y) + \delta(x)(\log x)^2 \int_x^y \frac{dt}{t(\log t)^2} \ll \delta(x)\log x,\end{aligned}$$

$$\sum_{x<d\le y} \frac{\mu(d)}{d^2} \ll \frac{\delta(y)}{y} + \int_x^y \delta(t)\,\frac{dt}{t^2} \ll \frac{\delta(y)}{y} + \delta(x)\int_x^y \frac{dt}{t^2} \ll \frac{\delta(x)}{x},$$

In 1909, Landau [3] proved that $\sum_{d=1}^{\infty} \mu(d)/d = 0$. Hence for any c with $0 < c < c'$ we have

$$\sum_{1\le d\le x} \frac{\mu(d)}{d} = -\lim_{y\to\infty} \sum_{x<d\le y} \frac{\mu(d)}{d} \ll \delta(x)\log x \ll \exp\biggl(-c \frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\biggr),$$

which proves the first estimate of Lemma 3.

Furthermore,

$$\sum_{1\le d\le x} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} - \sum_{d>x} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} + O\biggl(\frac{\delta(x)}{x}\biggr),$$

which proves the second estimate of Lemma 3.

The proof of the estimates under the Riemann hypothesis is similar.

Proof of the theorem. In [1] I. Aliev, S. Kanemitsu and A. Schinzel reduce the problem to the evaluation of an elementary sum by proving

LEMMA 4 (Aliev, Kanemitsu, Schinzel). *For $n > 1$ we have*

$$P(n) = \frac{1}{2} + 2\sum_{b,c} \frac{1}{bc}$$

where the sum is taken over all integers b, c such that $1 \le b \le n < c < 2b$ and $(b, c) = 1$.

Applying this lemma we have

$$P(n) = \frac{1}{2} + 2\sum_{b=1}^{n} \sum_{n<c<2b} \frac{1}{bc} \sum_{d\,|\,(b,c)} \mu(d)$$
$$= \frac{1}{2} + 2\sum_{d=1}^{n} \frac{\mu(d)}{d^2} \sum_{1\le k\le n/d} \frac{1}{k} \sum_{n/d<m<2k} \frac{1}{m} = \frac{1}{2} + 2\sum_{d=1}^{n} \frac{\mu(d)}{d^2} S\left(\frac{n}{d}\right)$$

where

$$S(x) = \sum_{x/2<k\le x} \frac{1}{k} \sum_{x<m<2k} \frac{1}{m}.$$

LEMMA 5. *For any positive number x we have*

$$S(x) = \frac{(\log 2)^2}{2} - \frac{1}{4x} + O\left(\frac{\log^+ x}{x^2}\right).$$

Proof. The argument is similar to those used in [1]. We use Corollary 1, which is a simple application of the Euler–Maclaurin summation:

$$S(x) = \sum_{x/2<k\le x} \frac{1}{k} \sum_{x<m\le 2k} \frac{1}{m} - \sum_{x/2<k\le x} \frac{1}{2k^2}$$
$$= \sum_{x/2<k\le x} \frac{1}{k}(\log 2k - \log x + \phi(x) - \phi(2k) + O(x^{-2})) - \sum_{x/2<k\le x} \frac{1}{2k^2}$$
$$= \sum_{x/2<k\le x} \frac{\log k}{k} - \sum_{x/2<k\le x} \frac{1}{4k^2} - (\log(x/2) - \phi(x) + O(x^{-2})) \sum_{x/2<k\le x} \frac{1}{k}$$
$$= \frac{(\log 2)^2}{2} - \frac{1}{4x} + O\left(\frac{\log^+ x}{x^2}\right). \blacksquare$$

If we replace $S(n/d)$ by the asymptotic formula above and do a straightforward summation over d we obtain the result of I. Aliev, S. Kanemitsu and A. Schinzel in [1].

However, we observe that if d is large, then n/d is small and therefore the error term in the asymptotic formula above is bad. Hence we need a different argument when d is large.

Let R be an integer such that $R \asymp n^{1/3}$ (this choice will be explained later). We then have

$$P(n) = \frac{1}{2} + 2\sum_{1\le d\le n/R} \frac{\mu(d)}{d^2} S\left(\frac{n}{d}\right) + 2\sum_{1\le r<R} \sum_{n/(r+1)<d\le n/r} \frac{\mu(d)}{d^2} S\left(\frac{n}{d}\right).$$

We observe that for any real number $x > 0$ we have $S(x) = S(\lfloor x \rfloor)$. Indeed, if k and m are integers we have

$$x/2 < k \le x \Leftrightarrow \lfloor x \rfloor/2 < k \le \lfloor x \rfloor, \qquad x < m \le 2k \Leftrightarrow \lfloor x \rfloor < m \le 2k.$$

Now for $n/(r+1) < d \le n/r$ we have $\lfloor n/d \rfloor = r$. Hence

$$(3) \qquad P(n) = \frac{1}{2} + 2 \sum_{1 \le d \le n/R} \frac{\mu(d)}{d^2} S\left(\frac{n}{d}\right) + 2 \sum_{1 \le r < R} S(r) \sum_{n/(r+1) < d \le n/r} \frac{\mu(d)}{d^2}.$$

We will use Lemma 5 to replace $S(n/d)$ and $S(r)$ by the corresponding asymptotic formula. We recall that $R \asymp n^{1/3}$, which implies that $\log(n/R) \asymp \log n$. We deduce from Lemma 3 that there exists $c > 0$ such that for $1 \le r \le R$,

$$\sum_{d > n/r} \frac{\mu(d)}{d^2} \ll \frac{r}{n} \exp\left(-c \frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\right).$$

The term $(\log 2)^2/2$ from Lemma 5 for $S(n/d)$ and $S(r)$ contributes to $P(n)$ (in (3)) the amount

$$(\log 2)^2 \left(\sum_{1 \le d \le n/R} \frac{\mu(d)}{d^2} + \sum_{1 \le r < R} \sum_{n/(r+1) < d \le n/r} \frac{\mu(d)}{d^2} \right)$$
$$= (\log 2)^2 \sum_{1 \le d \le n} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} (\log 2)^2 - (\log 2)^2 \sum_{d > n} \frac{\mu(d)}{d^2},$$

which gives the constant term $\frac{6}{\pi^2}(\log 2)^2$ and an admissible error term by Lemma 3.

The term $-1/(4(n/d))$ from Lemma 5 for $S(n/d)$ contributes to $P(n)$ (in (3)) the amount

$$-\frac{1}{2n} \sum_{1 \le d \le n/R} \frac{\mu(d)}{d},$$

which by Lemma 3 is an error term of order

$$O\left(\frac{1}{n} \exp\left(-c \frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\right)\right)$$

and under the Riemann hypothesis

$$O\left(\frac{1}{n}\left(\frac{n}{R}\right)^{-1/2+\varepsilon}\right) = O(R^{1/2} n^{-3/2+\varepsilon}).$$

The term $-1/(4r)$ from Lemma 5 for $S(r)$ contributes to $P(n)$ (in (3)) the amount

$$-\frac{1}{2}\sum_{1\le r<R}\frac{1}{r}\sum_{n/(r+1)<d\le n/r}\frac{\mu(d)}{d^2}$$
$$=-\frac{1}{2}\sum_{1\le r<R}\frac{1}{r}\bigg(\sum_{d>n/(r+1)}\frac{\mu(d)}{d^2}-\sum_{d>n/r}\frac{\mu(d)}{d^2}\bigg)$$
$$=-\frac{1}{2}\bigg(\frac{1}{R}\sum_{d>n/R}\frac{\mu(d)}{d^2}-\sum_{d>n}\frac{\mu(d)}{d^2}+\sum_{2\le r\le R}\frac{1}{r(r-1)}\sum_{d>n/r}\frac{\mu(d)}{d^2}\bigg),$$

which is of order

$$O\bigg(\bigg(\frac{1}{R}\frac{R}{n}+\frac{1}{n}+\frac{1}{n}\sum_{2\le r\le R}\frac{1}{(r-1)}\bigg)\exp\bigg(-c\,\frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\bigg)\bigg),$$

which in turn is

$$O\bigg(\frac{1}{n}\exp\bigg(-c'\,\frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\bigg)\bigg)\quad\text{for } 0<c'<c.$$

Under the Riemann hypothesis this error term becomes

$$O\bigg(\frac{1}{R}\Big(\frac{n}{R}\Big)^{-3/2+\varepsilon}+n^{-3/2+\varepsilon}+\sum_{1\le r\le R}\frac{1}{r^2}\Big(\frac{n}{r}\Big)^{-3/2+\varepsilon}\bigg)=O(R^{1/2}n^{-3/2+\varepsilon}).$$

The error term $O(\log^+(n/d)/(n/d)^2)$ from Lemma 5 for $S(n/d)$ contributes to $P(n)$ (in (3)) the amount

$$O\bigg(\sum_{1\le d\le n/R}\frac{\log^+(n/d)}{n^2}\bigg)=O\bigg(\frac{\log n}{nR}\bigg)$$

The error term $O((\log^+(r))/r^2)$ from Lemma 5 for $S(r)$ contributes to $P(n)$ (in (3)) the amount

$$O\bigg(\sum_{1\le r<R}\frac{\log^+(r)}{r^2}\bigg|\sum_{n/(r+1)<d\le n/r}\frac{\mu(d)}{d^2}\bigg|\bigg),$$

which is

$$O\bigg(\log n\sum_{1\le r<R}\frac{1}{r^2}\frac{r+1}{n}\exp\bigg(-c\,\frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\bigg)\bigg),$$

which in turn is

$$O\bigg(\frac{1}{n}\exp\bigg(-c'\,\frac{(\log n)^{3/5}}{(\log\log n)^{1/5}}\bigg)\bigg)\quad\text{for } 0<c'<c.$$

Under the Riemann hypothesis this error term becomes

$$O\bigg(\log n\sum_{1\le r<R}\frac{1}{r^2}\Big(\frac{n}{r+1}\Big)^{-3/2+\varepsilon}\bigg)=O((\log n)R^{1/2-\varepsilon}n^{-3/2+\varepsilon}).$$

We now see that the choice $R \asymp n^{1/3}$ permits us to optimize the sum $R^{1/2}n^{-3/2} + 1/(nR)$ and completes the proof of the theorem.

Acknowledgements. It is a great pleasure for the author to thank Professor A. Schinzel for an interesting discussion on this subject and his comments on the preprint. The author would also like to thank I. Aliev for his valuable remarks on an earlier version of the preprint.

REFERENCES

[1] I. Aliev, S. Kanemitsu and A. Schinzel, *On the metric theory of continued fractions*, Colloq. Math. 77 (1998), 141–146.

[2] P. Kargaev and A. Zhigljavsky, *Asymptotic distribution of the distance function to the Farey points*, J. Number Theory 65 (1997), 130–149.

[3] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig, 1909; reprinted by Chelsea, New York, 1953.

[4] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, revised by D. R. Heath-Brown, second ed., Oxford Univ. Press, New York, 1986.

[5] A. Walfisz, *Weylsche Exponentialsummen in der neueren Zahlentheorie*, Deutscher Verlag Wiss., Berlin, 1963.

Institut Girard Desargues
Université Lyon I
43, boulevard du 11 novembre 1918
69622 Villeurbanne, France
E-mail: rivat@desargues.univ-lyon1.fr

Received 9 February 1998;
revised 3 April 1998