ISSN 1064–5624, Doklady Mathematics, 2009, Vol. 79, No. 1, pp. 87–89. © Pleiades Publishing, Ltd., 2009.
Original Russian Text © A.V. Ustinov, 2009, published in Doklady Akademii Nauk, 2009, Vol. 424, No. 4, pp. 459–461.

MATHEMATICS

On the Statistical Properties of Elements of Continued Fractions

A. V. Ustinov

Presented by Academician Ya.G. Sinai August 7, 2008

Received September 3, 2008

DOI: 10.1134/S1064562409010268

For real α, the canonical continued fraction expansion is denoted by square brackets:

$$\alpha = [a_0; a_1, a_2, \ldots, a_s, \ldots] = a_0 + \cfrac{1}{a_1 + \ddots + \cfrac{1}{a_s + \ldots}},$$

where $a_0 = [\alpha]$ (the integer part of α) and $a_1, a_2, \ldots, a_s, \ldots$ are incomplete quotients (positive integers). By $\frac{p_n}{q_n} = [a_0; a_1, a_2, \ldots, a_n]$, we denote the convergents of α.

Assume that $R > 1$ is an increasing parameter and $s \geq 1$ is an integer such that $q_{s-1} \leq R < q_s$. In various problems in the theory of dynamical systems and number theory, it is necessary to know the joint limiting distribution of $\frac{q_{s-1}}{R}$, $\frac{R}{q_s}$, a_{s-K}, $\ldots$, a_{s+K} (K is a constant) when α is a random number in the interval (0, 1) (see [1, 4, 5]). The existence of this distribution was proved by ergodic theory methods in [4]. It is also important to know that this distribution is also obtained when α is a random rational number of the form $\frac{a}{b}$, where $1 \leq a \leq b \leq R^2$ and $(a, b) = 1$ (see [1]).

In this paper, this distribution is written explicitly and we prove that it has the same form in the continuous and discrete cases.

Denote by M the set of all integer matrices $S = \begin{pmatrix} P & P' \\ Q & Q' \end{pmatrix}$ with the determinant $\det S = \pm 1$ such that $1 \leq Q \leq Q'$, $0 \leq P \leq Q$, and $1 \leq P' \leq Q'$.

Let $\alpha \in (0, 1)$ be a real number. Then the rational numbers $\frac{P}{Q}$ and $\frac{P'}{Q'}$ for which $S = \begin{pmatrix} P & P' \\ Q & Q' \end{pmatrix} \in M$ are successive convergents of α (other than α) if and only if

$$0 < \frac{Q'\alpha - P'}{-Q\alpha + P} = S^{-1}(\alpha) < 1$$

(see [3, Lemma 1]). Moreover, if $\alpha = [0; a_1, a_2, \ldots]$, then, for some $s \geq 1$,

$$\begin{aligned} \frac{P}{Q} &= [0; a_1, a_2, \ldots, a_{s-1}], \\ \frac{P'}{Q'} &= [0; a_1, a_2, \ldots, a_s], \\ \frac{Q}{Q'} &= [0; a_s, a_{s-1}, \ldots, a_1], \\ \frac{Q'\alpha - P'}{-Q\alpha + P} &= [0; a_{s+1}, a_{s+2}, \ldots]. \end{aligned} \tag{1}$$

Therefore, the incomplete quotients $a_{s-K}, \ldots, a_{s+K}$ are described by the Gauss–Kuzmin distributions of the fractions $\frac{Q}{Q'}$ and $\frac{Q'\alpha - P'}{-Q\alpha + P}$.

For real α, x_1, x_2, y_1, $y_2 \in (0, 1)$, we denote by $N_{x_1, x_2, y_1, y_2}(\alpha, R)$ the number of solutions to the system of inequalities

$$\begin{aligned} 0 < S^{-1}(\alpha) \leq x_1, \quad Q \leq x_2 Q', \\ Q \leq y_1 R, \quad R \leq y_2 Q', \end{aligned} \tag{2}$$

in which the unknowns are the coefficients of the matrix $S = \begin{pmatrix} P & P' \\ Q & Q' \end{pmatrix} \in M$. Define

Institute of Applied Mathematics, Khabarovsk Branch, Far East Division, Russian Academy of Sciences, ul. Shevchenko 9, Khabarovsk, 680000 Russia

$$N(R) = N_{x_1, x_2, y_1, y_2}(R) = \int_0^1 N_{x_1, x_2, y_1, y_2}(\alpha, R)d\alpha$$

and

$$F(x_1, x_2, y_1, y_2) = \begin{cases} \frac{2}{\varsigma(2)}\left(\ln(1 + x_1x_2)\ln\frac{y_1y_2}{x_2} - \mathrm{Li}_2(-x_1x_2)\right), \\ \text{if } \ x_2 \le y_1y_2 \\ -\frac{2}{\varsigma(2)}\mathrm{Li}_2(-x_1y_1y_2), \text{ if} \\ x_2 > y_1y_2, \end{cases}$$

where $\mathrm{Li}_2(\cdot)$ is the Euler dilogarithm:

$$\mathrm{Li}_2(z) = \sum_{k=1}^{\infty}\frac{z^k}{k^2} = -\int_0^z \frac{\ln(1-t)}{t}dt.$$

Theorem 1. *Let $R \ge 2$. Then*

$$N(R) = F(x_1, x_2, y_1, y_2) + O\left(\frac{x_1y_2\ln R}{R}\right).$$

Proof. By using (1), each number $\alpha = [0; a_1, a_2, \ldots]$ is associated with a matrix $S \in M$ for which $Q \le R < Q'$. Moreover, the inequalities $0 < S^{-1}(\alpha) \le x_1$ specify an interval $I_{x_1}(S)$ inside (0, 1) of length

$$|I_{x_1}(S)| = \left|\frac{P' + x_1P}{Q' + x_1Q} - \frac{P'}{Q'}\right| = \frac{x_1}{Q'(Q' + x_1Q)}.$$

Therefore,

$$N(R) = \sum_{\left(\begin{smallmatrix} P & P' \\ Q & Q' \end{smallmatrix}\right) \in M} \chi_{[0, x_2Q']}(Q)\chi_{[0, y_1R]}(Q) \times \chi_{[0, y_2Q']}(R)\frac{x_1}{Q'(Q' + x_1Q)},$$

where $\chi_I(\cdot)$ is the characteristic function of the interval I. The (second) row (Q, Q') can be supplemented to be a matrix from M in exactly two ways. Therefore,

$$N(R) = 2\sum_{Q' \ge R/y_2}\sum_{(Q, Q') = 1}\chi_{[0, x_2Q']}(Q) \times \chi_{[0, y_1R]}(Q)\frac{x_1}{Q'(Q' + x_1Q)}. \quad (3)$$

Consider the case of $x_2 \le y_1y_2$. By the Möbius inversion formula,

$$N(R) = 2\sum_{d < R}\frac{\mu(d)}{d^2}\sum_{R/(y_2d) \le Q' < y_1R/(x_2d)}\sum_{Q \le x_2Q'}\frac{x_1}{Q'(Q' + x_1Q)}$$

$$+ 2\sum_{d < R}\frac{\mu(d)}{d^2}\sum_{Q' \ge y_1R/(x_2d)}\sum_{Q \le y_1R/d}\frac{x_1}{Q'(Q' + x_1Q)}$$

$$= \frac{2}{\varsigma(2)}\left(\ln(1 + x_1x_2)\ln\frac{y_1y_2}{x_2} + \int_{\frac{1}{x_1x_2}}^{\infty}\ln\left(1 + \frac{1}{t}\right)\frac{dt}{t}\right) + O\left(\frac{x_1y_2\ln R}{R}\right),$$

which yields the required equality. The case of $x_2 > y_1y_2$ is treated similarly.

Consider the sum

$$L(R) = L_{x_1, x_2, y_1, y_2}(R) = \sum_{b \le R^2}\sum_{\substack{a \le b \\ (a, b) = 1}} N_{x_1, x_2, y_1, y_2}\left(\frac{a}{b}, R\right).$$

Theorem 2. *Let $R \ge 2$. Then*

$$\frac{2\varsigma(2)}{R^4}L(R) = F(x_1, x_2, y_1, y_2) + O\left(x_1\frac{(y_1 + y_2)\ln^2 R}{R}\right).$$

Proof. Given a number $\frac{a}{b}$ and a solution to system (2), the integers m and n are defined by the equalities $mP + nP' = a$ and $mQ + nQ' = b$. Then system (2) can be rewritten as

$$mP + nP' = a, \quad mQ + nQ' = b, \quad 0 < \frac{m}{n} \le R^2,$$

$$0 < \frac{Q}{Q'} \le x_2, \quad Q \le y_1R, \quad R \le y_2Q'.$$

Summing its solutions over a and b, we find that the sum $L(R)$ is equal to the number of solutions to the system

$$mQ + nQ' \le R^2, \quad 0 < \frac{m}{n} \le x_1,$$

$$0 < \frac{Q}{Q'} \le x_2, \quad Q \le R < Q',$$

where $\begin{pmatrix} P & P' \\ Q & Q' \end{pmatrix} \in M$, $0 \le m \le n$, and $(m, n) = 1$. Again, given Q and Q', the values P and P' can be found in exactly two ways. Moreover, the number of solutions to the resulting system with respect to m and n is (see [2, Ch. 2, Ex. 21, 22]

$$\frac{R^4}{2\varsigma(2)}\frac{x_1}{Q'(Q' + x_1Q)} + O\left(\frac{x_1R^2\ln R}{Q'}\right).$$

Thus, we obtain a sum similar to (3):

$$L(R) = \frac{R^4}{\varsigma(2)} \sum_{\frac{R}{y_2} \le Q' \le R^2} \sum_{\substack{Q \le \min\{y_1 R, x_2 Q'\} \\ (Q, Q') = 1}} \frac{x_1}{Q'(Q' + x_1 Q)} + O(x_1 y_1 R^3 \ln^2 R).$$

Therefore,

$$L(R) = \frac{R^4}{\varsigma(2)} N(R) + O(x_1 y_1 R^3 \ln^2 R)$$

and Theorem 2 follows from Theorem 1.

Remark. When $x_2 = y_1 = y_2 = 1$, we obtain the distribution function

$$F(x) = F(x, 1, 1, 1) = -\frac{2}{\varsigma(2)} \mathrm{Li}_2(-x),$$

which differs from the Gauss–Kuzmin distribution $\log_2(1 + x)$. Moreover, $F(x)$ (together with the remainders in Theorems 1 and 2) decreases linearly as $x \to 0$: $F(x) \sim \frac{2x}{\varsigma(2)}$. Therefore, the expectation of the incomplete quotient a_s (defined by the conditions $q_{s-1} \le R < q_s$) is infinite.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 07-01-00306) and by the Far East Division of the Russian Academy of Sciences (project no. RAS 06-III-A-01-017) and the Russian Science Support Foundation.

REFERENCES

1. J. Bourgain and Ya. G. Sinai, Usp. Mat. Nauk **62** (4), 77–90 (2007).
2. I. M. Vinogradov, *Fundamentals of Number Theory* (Nauka, Moscow, 1972) [in Russian].
3. A. V. Ustinov, Zap. Nauchn. Semin. POMI **322**, 186–211 (2005).
4. Ya. G. Sinai and C. Ulcigrai, Ergodic Theory Dyn. Syst. **28**, 643–655 (2008).
5. Ya. G. Sinai and C. Ulcigrai, in *Probabilistic and Geometric Structures in Dynamics (Contemporary Mathematics)* (Am. Math. Soc., Providence, R.I., 2008).