Note on the prime divisors of Farey fractions

Vytautas Kazakevičius, Vilius Stakėnas

Department of Mathematics and Informatics, Vilnius University
Naugarduko 24, LT-03225 Vilnius
E-mail: vytautas.kazakevicius@mif.vu.lt; vilius.stakenas@mif.vu.lt

Abstract. Let \( P_1(n) \geq P_2(n) \geq \cdots \) be the prime divisors of a natural number \( n \) arranged in the non-increasing order. The limit distribution of the sequences \( (\log P_i(mn)/\log(mn)) \), \( i \geq 1 \) for \( m/n \in (\lambda_1; \lambda_2), n \leq x \), are considered. It is proved that under some conditions on \( \lambda_i \) the limit distribution of the sequences exists and is closely related to the Poisson–Dirichlet distribution.

Keywords: rational numbers, prime divisors, Poisson-Dirichlet distribution.

1 Introduction and the main result

Let \( \mathbb{N} \) denote the set of natural numbers and \( \mathbb{R}^\infty \) the linear space of all real sequences \( x = (x_1, x_2, \ldots) \) endowed with the product topology. It is well known that \( \mathbb{R}^\infty \) is a separable metrizable topological space. Consider the function \( \xi : \mathbb{N} \to \mathbb{R}^\infty \) defined as follows: if \( n = p_1 \cdots p_t \) with all \( p_i \) primes and \( p_1 \geq \cdots \geq p_t \) then

\[
\xi(n) = (\log p_1, \ldots, \log p_t, 0, 0, \ldots).
\]

Let \( \nu_x \) denote the uniform distribution on \( \{n \in \mathbb{N} : n \leq x\} \). The function \( n \mapsto n \) is a random variable on the probability space \( (\mathbb{N}, \nu_x) \); we denote it by the same letter \( n \). Then \( \xi(n) \) is a random element of \( \mathbb{R}^\infty \) defined on \( (\mathbb{N}, \nu_x) \).

It was proved by P. Billingsley in [1] that

\[
\frac{\xi(n)}{\log n} \rightarrow \eta,
\]

here \( \rightarrow \) denotes convergence in distribution (as \( x \to \infty \)) and \( \eta \) is a random element of \( \mathbb{R}^\infty \) distributed accordingly to the so-called Poisson–Dirichlet law. The new proof of this fact was given by P. Donnelly and G. Grimmett in [4].

We set the analogous problem of convergence of probabilistic measures, related to rational numbers.

Let \( \mathbb{Q}_+ \) denote the set of positive rational numbers, \( I \subset (0; \infty) \) and \( \nu'_x \) denote the uniform distribution on

\[
F^I_x = \left\{ \frac{m}{n} \in I : n \leq x \right\}.
\]

Each element of \( \mathbb{Q}_+ \) is represented in the unique way by an irreducible fraction \( m/n \); we consider the nominator and denominator of it as random variables on the probability space \( (\mathbb{Q}_+, \nu'_x) \), denoted by the same letters \( m \) and \( n \). The following theorem was proved by the second author in [7] using the proof in [4] as a model.
Theorem 1. Let $I = (\lambda_1; \lambda_2)$, where $0 \leq \lambda_1 < \lambda_2 < \infty$ satisfy the condition: for an arbitrary $0 < \gamma \leq 1$

$$(1 + \lambda_1)^{\gamma - 1}(\lambda_2 - \lambda_1)x^\gamma \to \infty, \ x \to \infty.$$ 

Then

$$\frac{\xi(m)}{\log m}, \frac{\xi(n)}{\log n} \overset{\mathcal{L}}{\sim} (\eta, \eta'),$$

where $\eta'$ is an independent copy of $\eta$.

In this paper we consider the limit distribution of $\xi(mn)\log(mn)$. Let

$$\Delta = \left\{ x \in \mathbb{R}^\infty: \forall i \ x_i \geq 0, \sum_{i \neq 1} x_i = 1 \right\}$$

and $R : \Delta \to \Delta$ be the ranking function of Billingsley (see [2, Chapter 1, Section 4]). It omits the zero components of the infinite tuple and rearranges the positive ones into non-increasing order; if the resulting tuple is finite, the infinite tail of zeros is added. Let $T$ denote the map from $\mathbb{R}^\infty \times \mathbb{R}^\infty$ to $\mathbb{R}^\infty$, defined by

$$T(x, y) = (x_1, y_1, x_2, y_2, \ldots).$$

Our main result is the following theorem.

Theorem 2. Let $I = (\lambda_1; \lambda_2)$, where $0 \leq \lambda_1 < \lambda_2 < \infty$ satisfy the condition: for an arbitrary $0 < \gamma \leq 1$

$$(1 + \lambda_1)^{\gamma - 1}(\lambda_2 - \lambda_1)x^\gamma \to \infty \quad \text{and} \quad \frac{\log(\lambda_2x)}{\log x} \to p, \quad \text{as} \ x \to \infty,$$

where $p \geq 0$. Then

$$\frac{\xi(mn)}{\log(mn)} \overset{\mathcal{L}}{\sim} RT\left(\frac{\eta}{p + 1}, \frac{\eta'}{p + 1}\right),$$

where $\eta, \eta'$ are the same random elements as in Theorem 1.

Proof. Let $m = p_1 \cdots p_s$, $n = q_1 \cdots q_t$ with all $p_i, q_j$ primes, $p_1 \geq \cdots \geq p_s$ and $q_1 \geq \cdots \geq q_t$. Then

$$\frac{\xi(mn)}{\log(mn)} = R\left(\frac{\log p_1}{\log(mn)}, \frac{\log q_1}{\log(mn)}, \frac{\log p_2}{\log(mn)}, \frac{\log q_2}{\log(mn)}, \ldots\right)$$

$$= RT\left(\frac{\xi(m)}{\log(mn)}, \frac{\xi(n)}{\log(mn)}\right).$$

Since both $R$ and $T$ are continuous, the theorem follows from Theorem 1 and Lemma 1 below, which is proved in Section 3. $\square$

Lemma 1. If conditions (1) are satisfied, then

$$\frac{\log n}{\log(mn)} \overset{\mathcal{L}}{\sim} \frac{1}{p + 1}.$$

It can be shown actually, that only the values $p \geq 1$ can appear in (1).
2 Marginal distributions

Let $P_1(n) \geq P_2(n) \geq \cdots$ be the prime divisors of $n$ arranged in the non-increasing order. Then the distributions of $\log P_k(n)/\log n$ converge as $x \to \infty$ to the one-dimensional marginal distributions of the Poisson–Dirichlet law. Since $\log n/\log x \to 1$, the same is true for the distributions of $\log P_k(n)/\log x$. The marginal distributions of the Poisson–Dirichlet measure in the number-theoretic context were discovered indeed in the form

$$\nu_x \{ P_k(n) \leq x^{1/u} \} \to \rho_k(u), \quad u > 0, \quad x \to \infty.$$  

(2)

The investigation of these asymptotics was initiated by K. Dickman [3]. The properties of the function $\rho(u) = \rho_1(u)$ were investigated by N.G. de Bruijn. It is called Dickman–de Bruijn function and is defined by the following differential-delay equation:

$$\rho(u) = 1 \quad \text{for} \quad 0 \leq u \leq 1, \quad u \rho'(u) + \rho(u - 1) = 0 \quad \text{for} \quad u > 1.$$ 

The papers of Ramaswami [6], Knuth and Trabb Pardo [5] followed, the functions $\rho_k(u)$ were investigated in numerous articles. It was shown, for example, that they are uniquely determined by the following properties: $\rho_k(u) = 1$ for $0 \leq u \leq 1$ and

$$\rho_k(u) = 1 - \int_0^{u-1} (\rho_k(t) - \rho_{k-1}(t)) \frac{dt}{1 + t} \quad \text{for} \quad u > 1, \quad k \geq 2.$$ 

The multidimensional-marginal distributions are described by P. Billingsley [1], [2], see also A. Vershik [8]. They showed that

$$\nu_x \left\{ \frac{\log P_1(n)}{\log n} \leq u_1, \ldots, \frac{\log P_k(n)}{\log n} \leq u_k \right\} \to \Phi_k(u_1, \ldots, u_k),$$ 

where the functions $\Phi_k$ are expressed via the Dickman–de Bruijn function in the following way:

$$\Phi_k(u_1, \ldots, u_k) = \int_0^{u_1} \int_{t_1}^{u_2} \cdots \int_{t_{k-1}}^{u_k} \rho \left( \frac{1 - t_1 - \cdots - t_k}{t_k} \right) dt_1 \cdots dt_k.$$ 

In this section we find limit distributions for $\log P_k(mn)/\log(mn)$, where $m$ and $n$ are random variables on $(\mathbb{Q}_+, \nu'_2)$. Suppose that conditions (1) are satisfied and denote $\alpha = p/(p + 1)$, $\beta = 1 - \alpha$. Let $\eta = (\eta_1, \eta_2, \ldots)$ and $\eta' = (\eta'_1, \eta'_2, \ldots)$ be independent random sequences, distributed accordingly the Poisson–Dirichlet law, and $\zeta = (\zeta_1, \zeta_2, \ldots) = RT(\alpha \eta, \beta \eta')$. Then, by Theorem 2,

$$\frac{\log P_k(mn)}{\log(mn)} \sim \zeta_k.$$ 

Let $F_k$ and $G_k$ denote the distribution functions of $\eta_k$ and $\zeta_k$, respectively. Then $F_k(u) = \rho_k(1/u)$. We show how $G_k$ is expressed via $F_i$ with $i \leq k$.

The case $k = 1$ is the most simple. Since $\zeta_1 = \max(\alpha \eta_1, \beta \eta'_1)$, we have

$$G_1(u) = P\{ \zeta_1 \leq u \} = P\{ \alpha \eta_1 \leq u \} P\{ \beta \eta'_1 \leq u \} = F_1(\alpha^{-1}u)F_1(\beta^{-1}u).$$
In the general case it is more convenient to work with $G^*_k(u) = 1 - G_k(u)$ and $F^*_k(u) = 1 - F_k(u)$. For positive integers $i, j$ define the random events

$$U_{i0} = \{ \alpha \eta_i > u \}, \quad U_{0j} = \{ \beta \eta_j' > u \}, \quad \text{and} \quad U_{ij} = \{ \alpha \eta_i > u, \beta \eta_j' > u \}. $$

The event $\{ \zeta_k > u \}$ occurs if at least one of the events $U_{ij}$ with $i + j = k$ appears. Hence

$$G^*_k(u) = P \left( \bigcup_{i+j=k} U_{ij} \right).$$

The probabilities of the events $U_{ij}$ as well as of their intersections can be expressed via the functions $F^*_k(u)$. Let us consider the case $k = 2$ for example. We have

$$P(U_{20}) = F^*_2(\alpha^{-1}u), \quad P(U_{02}) = F^*_2(\beta^{-1}u), \quad P(U_{11}) = F^*_1(\alpha^{-1}u)F^*_1(\beta^{-1}u),$$

$$P(U_{20} \cap U_{02}) = F^*_2(\alpha^{-1}u)F^*_2(\beta^{-1}u), \quad P(U_{20} \cap U_{11}) = F^*_2(\alpha^{-1}u)F^*_1(\beta^{-1}u),$$

$$P(U_{02} \cap U_{11}) = F^*_1(\alpha^{-1}u)F^*_2(\beta^{-1}u)$$

and

$$P(U_{02} \cap U_{20} \cap U_{11}) = F^*_2(\alpha^{-1}u)F^*_2(\beta^{-1}u)$$

hence

$$F^*_k(u) = F^*_2(\alpha^{-1}u) + F^*_2(\beta^{-1}u) + F^*_1(\alpha^{-1}u)F^*_1(\beta^{-1}u)$$

$$- F^*_2(\alpha^{-1}u)F^*_1(\beta^{-1}u) - F^*_1(\alpha^{-1}u)F^*_2(\beta^{-1}u).$$

### 3 Proof of Lemma 1

Let $F_z$ denote the distribution function of the random variable $\frac{\log n}{\log(mn)}$ and $F$ be that of the random variable which equals $\frac{1}{p+1}$ with probability 1:

$$F_z(z) = \nu_z^I \left\{ \frac{\log n}{\log(mn)} \leq z \right\}, \quad F(z) = \begin{cases} 0 & \text{for } z < \frac{1}{p+1}, \\ 1 & \text{otherwise.} \end{cases}$$

We need to show that $F_z(z) \to F(z)$, as $x \to \infty$, for all $z \in (0; 1)$, $z \neq \frac{1}{p+1}$.

Let $0 < z < \frac{1}{p+1}$. Fix $\epsilon > 0$ and find $x_0$ such that for all $x \geq x_0$

$$\frac{\log(ex)}{\log x} > \frac{z}{1 - z}, \quad \frac{\log(\lambda_2 x)}{\log x} > \frac{z}{1 - z}.$$

Inequalities

$$ex < n \leq x, \quad \lambda_1 < \frac{m}{n} < \lambda_2, \quad \frac{\log n}{\log(mn)} \leq z$$

imply

$$\log(ex) \leq \log n \leq \frac{z}{1 - z} \log m \leq \frac{z}{1 - z} \log(\lambda_2 n) \leq \frac{z}{1 - z} \log(\lambda_2 x),$$

which is impossible if $x \geq x_0$. Therefore

$$\nu_z^I \left\{ ex < n, \quad \frac{\log n}{\log(mn)} \leq z \right\} = 0$$

for $x \geq x_0$. 

Note on the prime divisors of Farey fractions

On the other hand, conditions (1) imply \((\lambda_2 - \lambda_1)x \to \infty\), as \(x \to \infty\). Therefore, by Theorem 1 in \([7]\),

\[
\mathbb{#F}_x^I \sim \frac{3}{\pi^2}(\lambda_2 - \lambda_1)x^2,
\]

which yields

\[
\nu_x^I\{n \leq \epsilon x\} \leq \frac{\mathbb{#F}_x^I}{\mathbb{#F}_x} \to \epsilon^2,
\]

as \(x \to \infty\). Hence

\[
\lim_{x \to \infty} F_x(z) \leq \epsilon^2
\]

with \(\epsilon\) arbitrary small, i.e., \(F_x(z) \to 0\).

Now let \(\frac{1}{p+1} < z < 1\). Fix \(\epsilon > 0\) and find \(x_0\) such that

\[
1 < \frac{z}{1 - z} \frac{\log(\epsilon^2) + \log(\lambda_2 x)}{\log x}
\]

for \(x \geq x_0\). Inequalities

\[
ex < n \leq x, \quad \lambda_1 + \epsilon(\lambda_2 - \lambda_1) < \frac{m}{n} < \lambda_2, \quad \frac{\log n}{\log(mn)} > z
\]

imply

\[
\log x \geq \log n \geq \frac{z}{1 - z} \log m \geq \frac{z}{1 - z} \log(\epsilon\lambda_2 n) \geq \frac{z}{1 - z} \log(\epsilon^2\lambda_2 x),
\]

which is impossible if \(x \geq x_0\). Therefore

\[
\nu_x^I\{ex < n, \lambda_1 + \epsilon(\lambda_2 - \lambda_1) < \frac{m}{n}, \frac{\log n}{\log(mn)} > z\} = 0
\]

for \(x \geq x_0\). Also

\[
\nu_x^I\{n \leq ex\} \leq \frac{\mathbb{#F}_x^I}{\mathbb{#F}_x} \to \epsilon^2
\]

and

\[
\nu_x^I\left\{\lambda_1 < \frac{m}{n} < \lambda_1 + \epsilon(\lambda_2 - \lambda_1)\right\} = \frac{\mathbb{#F}_x^I(\lambda_1; \lambda_1 + \epsilon(\lambda_2 - \lambda_1))}{\mathbb{#F}_x^I} \to \epsilon.
\]

Therefore

\[
\lim_{x \to \infty} (1 - F_x(z)) \leq \epsilon + \epsilon^2
\]

with \(\epsilon\) arbitrary small, i.e., \(F_x(z) \to 1\).

References

[1] P. Billingsley. On the distribution of large prime divisors. *Periodica Mathematica Hungarica*, 2:283–289, 1972.

[2] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1999.

[3] K. Dickman. On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv för Mat.*, Astron. och Fys. A, 22:1–14, 1930.

Liet. mat. rink. LMD darbai, 52:13–18, 2011.
[4] P. Donnelly and G. Grimmett. On the asymptotic distribution of large prime factors. *J. London Math. Soc.*, 47:395–404, 1993.

[5] D.E. Knuth and L. Trabb Pardo. Analysis of simple factorization algorithm. *Theoretical Comput. Sci.*, 3:321–348, 1976/77.

[6] V. Ramaswami. On the number of positive integers less than $x$ and free of prime divisors greater than $x^c$. *Bull. Amer. Math. Soc.*, 55:1122–1127, 1949.

[7] V. Stakėnas. Sieve result for Farey fractions. *Lith. Math. J.*, 39(1):108–127, 1999.

[8] A. Vershik. The asymptotic distribution of factorizations of natural numbers into prime divisors. *Sov. Math. Dokl.*, 34:57–61, 1987.

REZIUMĖ

**Farey trupmenų pirminiai dalikliai**

*V. Kazakevičius, V. Stakėnas*

Nagrinėjamos racionalių skaičių pirminių daliklių variacinės eilutės. Įrodoma teorema apie sekos, gautos iš šių eilučių, ribinį skirstinį.

**Raktiniai žodžiai:** racionaliai skaičiai, pirminiai dalikliai, Puasono–Dirichle skirstinys.