

# 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  $P_1(n) \geq P_2(n) \geq \dots$  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  $x = (x_1, x_2, \dots)$  endowed with the product topology. It is well known that  $\mathbb{R}^\infty$  is a separable metrizable topological space. Consider the function  $\xi : \mathbb{N} \rightarrow \mathbb{R}^\infty$  defined as follows: if  $n = p_1 \cdots p_t$  with all  $p_i$  primes and  $p_1 \geq \dots \geq p_t$  then

$$\xi(n) = (\log p_1, \dots, \log p_t, 0, 0, \dots).$$

Let  $\nu_x$  denote the uniform distribution on  $\{n \in \mathbb{N} : n \leq x\}$ . The function  $n \mapsto \xi(n)$  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

$$\frac{\xi(n)}{\log n} \rightsquigarrow \eta,$$

here  $\rightsquigarrow$  denotes convergence in distribution (as  $x \rightarrow \infty$ ) 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  $\nu_x^I$  denote the uniform distribution on

$$\mathcal{F}_x^I = \left\{ \frac{m}{n} \in I : n \leq x \right\}.$$

Each element of  $\mathbb{Q}_+$  is represented in the unique way by an irreducible fraction  $m/n$ ; we consider the numerator and denominator of it as random variables on the probability space  $(\mathbb{Q}_+, \nu_x^I)$ , 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  $0 < \gamma \leq 1$

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

Then

$$\left( \frac{\xi(m)}{\log m}, \frac{\xi(n)}{\log n} \right) \rightsquigarrow (\eta, \eta'),$$

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

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

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

and  $R : \Delta \rightarrow \Delta$  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

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

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  $0 < \gamma \leq 1$

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

where  $p \geq 0$ . Then

$$\frac{\xi(mn)}{\log(mn)} \rightsquigarrow RT \left( \frac{p\eta}{p+1}, \frac{\eta'}{p+1} \right),$$

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

$$\begin{aligned} \frac{\xi(mn)}{\log(mn)} &= R \left( \frac{\log p_1}{\log(mn)}, \frac{\log q_1}{\log(mn)}, \frac{\log p_2}{\log(mn)}, \frac{\log q_2}{\log(mn)}, \dots \right) \\ &= RT \left( \frac{\xi(m)}{\log(mn)}, \frac{\xi(n)}{\log(mn)} \right) \\ &= RT \left( \frac{\log m}{\log(mn)} \cdot \frac{\xi(m)}{\log m}, \frac{\log n}{\log(mn)} \cdot \frac{\xi(n)}{\log n} \right). \end{aligned}$$

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

$$\frac{\log n}{\log(mn)} \rightsquigarrow \frac{1}{p+1}.$$

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

## 2 Marginal distributions

Let  $P_1(n) \geq P_2(n) \geq \dots$  be the prime divisors of  $n$  arranged in the non-increasing order. Then the distributions of  $\log P_k(n)/\log n$  converge as  $x \rightarrow \infty$  to the one-dimensional marginal distributions of the Poisson–Dirichlet law. Since  $\log n/\log x \rightsquigarrow 1$ , 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

$$\nu_x\{P_k(n) \leq x^{1/u}\} \rightarrow \rho_k(u), \quad u > 0, x \rightarrow \infty. \tag{2}$$

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:

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

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

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

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

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

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

$$\Phi_k(u_1, \dots, u_k) = \int_0^{u_k} \int_{t_k}^{u_{k-1}} \dots \int_{t_2}^{u_1} \rho \left( \frac{1-t_1-\dots-t_k}{t_k} \right) \frac{dt_1 \dots dt_k}{t_1 \dots t_k}.$$

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

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

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

$$\begin{aligned} P(U_{20}) &= F_2^*(\alpha^{-1}u), & P(U_{02}) &= F_2^*(\beta^{-1}u), & P(U_{11}) &= F_1^*(\alpha^{-1}u)F_1^*(\beta^{-1}u), \\ P(U_{20} \cap U_{02}) &= F_2^*(\alpha^{-1}u)F_2^*(\beta^{-1}u), & P(U_{20} \cap U_{11}) &= F_2^*(\alpha^{-1}u)F_1^*(\beta^{-1}u), \\ P(U_{02} \cap U_{11}) &= F_1^*(\alpha^{-1}u)F_2^*(\beta^{-1}u) \end{aligned}$$

and

$$P(U_{02} \cap U_{20} \cap U_{11}) = F_2^*(\alpha^{-1}u)F_2^*(\beta^{-1}u)$$

hence

$$\begin{aligned} F_2^*(u) &= F_2^*(\alpha^{-1}u) + F_2^*(\beta^{-1}u) + F_1^*(\alpha^{-1}u)F_1^*(\beta^{-1}u) \\ &\quad - F_2^*(\alpha^{-1}u)F_1^*(\beta^{-1}u) - F_1^*(\alpha^{-1}u)F_2^*(\beta^{-1}u). \end{aligned}$$

### 3 Proof of Lemma 1

Let  $F_x$  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:

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

We need to show that  $F_x(z) \rightarrow F(z)$ , as  $x \rightarrow \infty$ , 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  $x \geq x_0$

$$\frac{\log(\epsilon x)}{\log x} > \frac{z}{1-z} \cdot \frac{\log(\lambda_2 x)}{\log x}.$$

Inequalities

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

imply

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

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

$$\nu_x^I \left\{ \epsilon x < n, \frac{\log n}{\log(mn)} \leq z \right\} = 0$$

for  $x \geq x_0$ .

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

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

which yields

$$\nu_x^I\{n \leq \epsilon x\} \leq \frac{\#\mathcal{F}_{\epsilon x}^I}{\#\mathcal{F}_x^I} \rightarrow \epsilon^2,$$

as  $x \rightarrow \infty$ . Hence

$$\overline{\lim}_{x \rightarrow \infty} F_x(z) \leq \epsilon^2$$

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

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

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

for  $x \geq x_0$ . Inequalities

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

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

$$\nu_x^I \left\{ \epsilon x < n, \lambda_1 + \epsilon(\lambda_2 - \lambda_1) < \frac{m}{n}, \frac{\log n}{\log(mn)} > z \right\} = 0$$

for  $x \geq x_0$ . Also

$$\nu_x^I\{n \leq \epsilon x\} \leq \frac{\#\mathcal{F}_{\epsilon x}^I}{\#\mathcal{F}_x^I} \rightarrow \epsilon^2$$

and

$$\nu_x^I \left\{ \lambda_1 < \frac{m}{n} < \lambda_1 + \epsilon(\lambda_2 - \lambda_1) \right\} = \frac{\#\mathcal{F}_x^{(\lambda_1; \lambda_1 + \epsilon(\lambda_2 - \lambda_1))}}{\#\mathcal{F}_x^I} \rightarrow \epsilon.$$

Therefore

$$\overline{\lim}_{x \rightarrow \infty} (1 - F_x(z)) \leq \epsilon + \epsilon^2$$

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

## 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.

- [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 racionaliųjų skaičių pirminių daliklių variacinės eilutės. Įrodoma teorema apie sekos, gautos iš šių eilučių, ribinį skirstinį.

*Raktiniai žodžiai:* racionalieji skaičiai, pirminiai dalikliai, Puasono-Dirichle skirstinys.