I. KÁTAI

On distribution of arithmetical functions on the set prime plus one

Compositio Mathematica, tome 19, n° 4 (1968), p. 278-289

<http://www.numdam.org/item?id=CM_1968__19_4_278_0>
On distribution of arithmetical functions
on the set prime plus one
by
I. Kátai

1. Introduction

P. Erdős proved the following theorem [1].
Let \( f(n) \) be a real valued additive number-theoretical function, and put

\[
f^*(n) = \begin{cases} f(n) & \text{for } |f(n)| \leq 1, \\ 0 & \text{for } |f(n)| > 1. \end{cases}
\]

Put

\[
F_N(x) = \frac{1}{N} \sum_{n \leq N} \frac{1}{f(n) < x}.
\]

Then the distribution-functions \( F_N(x) \) tend for \( N \to +\infty \) to a limiting distribution function \( F(x) \) at all points of continuity of \( F(x) \), if the following three conditions are satisfied:

1. \[
\sum_p \frac{f^*(p)}{p} \text{ is convergent},
\]
2. \[
\sum_p \frac{(f^*(p))^2}{p} < +\infty,
\]
3. \[
\sum_{|f(p)| > 1} \frac{1}{p} < +\infty.
\]

It has been shown also by P. Erdős that \( F(x) \) is continuous if and only if the series \( \sum_{f(p) \neq 0} 1/p \) diverges.

New proof of this theorem has been given by H. Delange [2] and by A. Rényi [3].

A multiplicative function \( g(n) \) is called strongly multiplicative, if for all primes \( p \) and all positive integers \( k \) it satisfies the condition

\[
g(p^k) = g(p).
\]

H. Delange proved the following theorem [4].
If \( g(n) \) is a strongly multiplicative number-theoretical function such that \( |g(n)| \leq 1 \) for \( n = 1, 2, \ldots \), and such that the series
\[
\sum_p \frac{g(p) - 1}{p}
\]
converges, then
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^N g(n) = M(g)
\]
exists and
\[
M(g) = \prod_p \left(1 + \frac{g(p) - 1}{p}\right).
\]

A new proof of this theorem has been given by A. Rényi [5].

Throughout the paper \( p, q \) denote primes, and \( \sum_p \) and \( \prod_p \) denote a sum and a product, respectively, taken over all primes. Let further \( \text{li} \ x = \int_2^x \frac{du}{\log u} \).

The aim of this paper is to prove the following statement.

**Theorem 1.** Let \( g(n) \) be a complex-valued multiplicative function such that \( |g(n)| \leq 1 \) for \( n = 1, 2, \ldots \), and such that the series
\[
\sum_p \frac{g(p) - 1}{p}
\]
converges. Let \( N(g) \) denote the product
\[
N(g) = \prod_p \left(1 + \sum_{k=1}^{\infty} \frac{g(p^k) - g(p^{k-1})}{p^k(p-1)}\right).
\]
Then
\[
\lim_{x \to \infty} \frac{1}{\log x} \sum_{p \leq x} g(p+1) = N(g).
\]

From this theorem easily follows the

**Theorem 2.** Let \( f(n) \) be a real valued additive number-theoretical function which satisfies the conditions 1, 2, 3, of the theorem of Erdös.

Put
\[
F_N(y) = \frac{1}{\log N} \sum_{f(p+1) < y} 1.
\]

Then the distribution-functions \( F_N(y) \) tend for \( N \to \infty \) to a limiting distribution-function \( F(y) \) at all points of continuity of \( F(y) \).

Further \( F(y) \) is a continuous function if and only if
\[
\sum_{f(p) \neq 0} \frac{1}{p} = \infty.
\]
2. Deduction of Theorem 2 from Theorem 1

In what follows $c, c_1, c_2, \ldots$ denote constants not always the same in different places.

For the proof of Theorem 2 we need to prove only that the sequence of characteristic functions

\begin{equation}
\varphi_N(u) = \frac{1}{\ln N} \sum_{p \leq N} e^{iu\varphi(n)}
\end{equation}

converges to a function $\varphi(u)$, which is a continuous one on the real axis.

It is easy to verify that from the conditions 1/2/3 it follows that

\begin{equation}
\sum_p \frac{e^{iu\varphi(p)} - 1}{p}
\end{equation}

converges for every real $u$. Using now Theorem 1 with $g(n) = e^{iu\varphi(n)}$ we obtain that $\varphi_N(u) \to \varphi(u)$, where

\begin{equation}
\varphi(u) = \prod_p \left(1 + \sum_{k=1}^\infty \frac{e^{iu\varphi(p^k)} - e^{iu\varphi(p^{k-1})}}{p^{k-1}(p-1)}\right).
\end{equation}

The continuity of (2.3) is guaranteed by the continuity of (2.2), which follows from the conditions 1/2/3 evidently.

For the proof of the continuity of $F(x)$ in the case

\[ \sum_{f(p) \neq 0} \frac{1}{p} = \infty \]

we remark the following.

P. Levy proved the following theorem [8].

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent random variables with discrete distribution and suppose that there exists the sum

\[ \sum_{k=1}^\infty X_k = X \]

with probability 1. Let

\[ d_k = \sup_x P(X_k = x). \]

Then the distribution function of $X$ is continuous if and only if

\[ \prod_{k=1}^\infty d_k = 0. \]
Let now the $X_p$’s be independent random variables with characteristic functions

$$1 + \sum_{k=1}^{\infty} \frac{e^{ituf(p^k)} - e^{ituf(p^{k-1})}}{p^{k-1}(p-1)}$$

and let

$$X = \sum_p X_p.$$ 

It is evident from (2.3) that $X$ has the distribution function $F(x)$, and so this is continuous if and only if

$$\prod_{f(p) \neq 0} \left(1 - \frac{1}{p}\right) = 0,$$

i.e.

$$\sum_{f(p) \neq 0} \frac{1}{p} = \infty.$$

3. The proof of Theorem 1

We need the following Lemmas.

**Lemma 1.** Let $g(p)$ be a complex-valued function defined on the primes, for which $|g(p)| \leq 1$ and

$$\sum_p \frac{g(p)-1}{p}$$

converges. Then

$$\sum_{|\arg g(p)| > \epsilon} \frac{1}{p} < +\infty$$

for every positive constant $\epsilon$, further

$$\sum_p \frac{|g(p)-1|^2}{p} < +\infty$$

and

$$\sum_{x^{1/2} < p < x} \frac{|g(p)-1|}{p} \to 0 \quad \text{for } x \to +\infty.$$

**Proof.** The assertion in (3.4) is an immediate consequence of (3.3) since

$$\sum_{x^{1/2} < p < x} \frac{|g(p)-1|}{p} \leq \left(\sum_{x^{1/2} < p < x} \frac{1}{p}\right)^{1/2} \left(\sum_{x^{1/2} < p < x} \frac{|g(p)-1|^2}{p}\right)^{1/2},$$

further $\sum_{x^{1/2} < p < x} 1/p$ is bounded, and the second sum on the right hand side tends to zero because of (3.3).

Let us put

$$|g(p)| = r(p) \quad \text{and} \quad \arg g(p) = \theta(p),$$

where $-\pi < \theta(p) \leq +\pi$, i.e. we suppose that $g(p) = r(p)e^{it(p)}$. 


From the convergence of (3.1) it follows that

\[ \sum_{p} \frac{1 - \text{Re} g(p)}{p} < +\infty \]

converges too. This sum has positive terms and the inequality \( |\arg g(p)| > c \) involves that \( 1 - \text{Re} g(p) > c_1 (> 0) \). Hence (3.2) follows.

From the inequality \( |a + bi|^2 \leq 2(|a|^2 + |b|^2) \) it follows that

\[ \sum_{p} \frac{|g(p) - 1|^2}{p} \leq 2 \sum \frac{\text{Re} (1 - g(p))^2}{p} + 2 \sum \frac{|\text{Im} g(p)|^2}{p}. \]

The first sum on the right hand side evidently converges since

\[ \sum \frac{\text{Re} (1 - g(p))^2}{p} < \sum \frac{1 - \text{Re} g(p)}{p} + 4 \sum \frac{1}{|\theta(p)| > \frac{1}{2} p}. \]

It is sufficient to prove that

\[ (3.5) \quad \sum_{|\theta(p)| \leq \frac{1}{2}} \frac{r^2(p) \sin^2 \theta(p)}{p} < \infty. \]

Using the inequality

\[ r^2(p) \sin^2 \theta(p) \leq c \theta^2(p) \leq 2c \sin^2 \frac{\theta(p)}{2} \leq 1 - \cos \theta(p) \]

\[ \leq 1 - r(p) \cos \theta(p), \]

we have (3.5).

**Lemma 2.** Let \( N_k(x) \) denote the number of solutions of the equation

\[ p + 1 = kq, \quad p \leq X \]

in primes \( p, q \). Then

\[ N_k(X) < c \frac{x}{\varphi(k) \log^2 X/k}. \]

for \( k < x \), where \( c \) is an absolute constant.

For the proof see Prachar’s book [6], Theorem 4.6, p. 51.

Let \( \pi(x, k, l) \) denote the number of primes in the arithmetical progression \( \equiv l (\mod k) \) not exceeding \( x \).

**Lemma 3.** (Brun-Titchmarsh). For all \( k \leq x^{1-\delta} \), \( \delta > 0 \)

\[ \pi(x, k, l) < c_\delta \frac{x}{\varphi(k) \log x}, \]

where \( c_\delta \) is a constant depending on \( \delta \) only.

For the proof see [6].
**Lemma 4.** (E. Bombieri [7]).

\[
\sum_{D \leq Y} \max_{l \equiv (D, y) \equiv 1} \left| \frac{\pi(x, D, l)}{\varphi(D)} - \frac{\text{li} x}{\varphi(D)} \right| < \frac{cx}{(\log x)^A}
\]

where \(Y = x^{\frac{1}{2}} (\log x)^{-B}; \quad B \geq 2A + 23, \) \(A\) arbitrary constant.

Let \(\tau(n)\) be the number of divisors of \(n\).

**Lemma 5.**

\[
\sum_{n = 1}^{y} \frac{\tau^2(n)}{\varphi(n)} < c \left( \log y \right)^4,
\]

where \(c\) is a constant.

The proof is very simple and so can be omitted.

Let us define the multiplicative function \(g_K(n)\) by putting

\[
g_K(p^\alpha) = \begin{cases} 
g(p^\alpha), & \text{if } p^\alpha \leq K, \
1, & \text{if } p^\alpha > K. 
\end{cases}
\]

By other words we put for any natural number \(n\)

\[
g_K(n) = \prod_{p^\alpha \mid n} g(p^\alpha).
\]

Let us put further

\[
h_K(n) = \sum_{d \mid n} \mu \left( \frac{n}{d} \right) g_N(d),
\]

where \(d\) runs over all (positive) divisors of \(n\) and \(\mu(n)\) is the Möbius function. Then \(h_K(n)\) is also a multiplicative function, \(h_K(p^\alpha) = g_K(p^\alpha) - g_K(p^{\alpha - 1})\); \(h_K(p) = 0\) for \(p > K\); \(h_K(p) = 0\) for \(p^{\alpha - 1} > K, \alpha \geq 2\).

Let further \(h(n)\) be defined by

\[
h = \sum_{d \mid n} \mu \left( \frac{n}{d} \right) g(d).
\]

Let us introduce the notation

\[
I_K(x) = \sum_{p \leq x} g_K(p+1); \quad I(x) = \sum_{p \leq x} g(p+1).
\]

Choose now \(K_1 = (\frac{1}{4} - \varepsilon) \log x, \quad K_2 = x^{\frac{1}{4}}, \quad K_3 = x^{1-\delta}\), where \(\varepsilon\) and \(\delta\) are suitable small positive numbers.

We shall prove the following relations:

\[
I_{K_1}(x) = (1+o(1)) \text{li} x N(g) \quad \text{for } x \to \infty,
\]

\[
I_{K_2}(x) - I_{K_1}(x) = o(\text{li} x) \quad \text{for } x \to \infty,
\]

\[
(3.6) \quad I_{K_1}(x) = (1+o(1)) \text{li} x N(g) \quad \text{for } x \to \infty,
\]

\[
(3.7) \quad I_{K_2}(x) - I_{K_1}(x) = o(\text{li} x) \quad \text{for } x \to \infty,
\]
Theorem 1 follows if we choose \( \delta = \delta(x) \) tending to zero so slowly that the right hand side of (3.8) is \( o(\text{li } x) \).

First we prove (3.6). We have

\[
I_{K_1}(x) = \sum_{p \leq x} g_{K_1}(p+1) = \sum_{p \leq x} \sum_{d | p+1} h_{K_1}(a) = \sum_{d} h_{K_1}(d) \text{li}(x, d, -1) = \text{li } x \sum_{d} \frac{h_{K_1}(d)}{\varphi(d)} + R,
\]

where

\[
|R| \leq \sum_{d} |h_{K_1}(d)| \left| \frac{\pi(x, d, -1)}{\varphi(d)} - \frac{\text{li } x}{\varphi(d)} \right| = R_1.
\]

Using the prime number theorem, we obtain that \( h_{K_1}(d) = O \) for \( d \geq x^{1-\frac{\epsilon}{2}} \) because

\[
\prod_{p^e < K_1} p^e < e^{(\frac{1}{2}-\frac{\epsilon}{2}) \log x},
\]

if \( x \) is sufficiently large.

Since \( |g(n)| \leq 1 \), so \( |h_{K_1}(p^e)| \leq 2 \) and \( |h_{K_1}(d)| \leq \tau(d) \).

For the estimation of \( R_1 \) we split all of the \( d \)'s, \( d \leq x^{1-\frac{\epsilon}{2}} \) into two classes \( \mathcal{U}_1, \mathcal{U}_2 \) as follows:

\( d \) belongs to \( \mathcal{U}_1 \) or \( \mathcal{U}_2 \) according to that \( \tau(d) \leq (\log x)^5 \) or \( \tau(d) > (\log x)^5 \), respectively.

Using Lemma 3 and Lemma 5 we have

\[
\sum_{d \in \mathcal{U}_1} |h_{K_1}(d)| \left| \frac{\pi(x, d, -1)}{\varphi(d)} - \frac{\text{li } x}{\varphi(d)} \right| \leq c \text{ li } x \sum_{d \in \mathcal{U}_1} \frac{\tau(d)}{\varphi(d)} \leq c \text{ li } x (\log x)^{-5} \sum_{d \leq x} \frac{\tau^2(d)}{\varphi(d)} < c \frac{x}{\log^2 x}.
\]

Otherwise, using the Bombieri's result (Lemma 4), we have that the sum

\[
\sum_{d \in \mathcal{U}_1} |h_{K_1}(d)| \left| \frac{\pi(x, d, -1)}{\varphi(d)} - \frac{\text{li } x}{\varphi(d)} \right|
\]

not exceed

\[
(\log x)^5 \sum_{d \leq x^{1/4}} \left| \frac{\pi(x, d, -1)}{\varphi(d)} - \frac{\text{li } x}{\varphi(d)} \right| = O \left( \frac{x}{(\log x)^{A-\delta}} \right) = O \left( \frac{x}{\log^2 x} \right),
\]

if \( A \geq 7 \).

Hence

\[ R = O \left( \frac{x}{\log^2 x} \right). \]
Further we have

$$\sum_{d \leq x} \frac{h_{K_1}(d)}{\varphi(d)} = \prod_{p < K_1} \left(1 + \sum_{a=1}^{\infty} \frac{g_{K_1}(p^a) - g_{K_1}(p^{a-1})}{p^{a-1}(p-1)}\right).$$

From the convergence of the series \(\sum (g(p) - 1)/p\) it follows that the product on the right hand side tends to \(N(g)\) for \(x \to +\infty\). So (3.6) is proved.

Let now \(\tilde{g}(n)\) be a multiplicative function defined by

$$\tilde{g}(p^a) = \begin{cases} g(p^a), & \text{if } p \leq K_1, \\ g(p), & \text{if } p > K_1. \end{cases}$$

It is evident that \(\tilde{g}(n) = g(n)\) except eventually those \(n\)'s for which there exists a \(q, q > K_1\), such that \(q^2 \mid n\). So

(3.10)

$$\sum_{p \leq x} |g(p+1) - \tilde{g}(p+1)| \leq 2 \sum_{q > K_1} \sum_{p+1 \equiv 0(q^2)} 1 < 2c \log x \sum_{q < x^{1/2}} \frac{1}{q(q-1)} + x \sum_{q > x^{1/2}} \frac{1}{q^2} = o(\log x).$$

From (3.10)

$$|I_{K_1}(x) - I_{K_1}(x)| \leq \sum_{p \leq x} |\tilde{g}_{K_1}(p+1) - \tilde{g}_{K_1}(p+1)| + o(\log x)$$

$$\leq \sum_{p \leq x} \prod_{q|p+1} g(q)-1 + o(\log x) = V + o(\log x)$$

follows. Using the formulas

$$\log (1+z) = z + O(|z|^2); \quad \exp (z + O(|z|^2)) = 1 + z + O(|z|^2)$$

for \(|z| \leq 1, \ |\arg z| \leq \pi/2\), we have that

(3.11) \(\prod_{q|p+1} g(q)-1 = \sum_{q|p+1} h(g) + O(\sum_{q|p+1} |h^2(g)|)\),

if all primdivisor \(q\) of \(p+1\) in the interval \(K_1 < q \leq K_2\) satisfies the relation \(|\arg g(q)| \leq \pi/2\). Let \(\mathcal{A}_3\) denote the set of the \(p\)'s possessing this property, and \(\mathcal{A}_4\) the other \(p\)'s.

We can easily estimate the sum

$$V_1 = \sum_{p \in \mathcal{A}_4} \prod_{q|p+1} g(q)-1,$$

since
Let

\[ V_1 < 2 \sum_{K_1 < q \leq K_2} \pi(x, q, -1) < c \text{li} \, x \sum_{K_1 < q < K_2} \frac{1}{q} \]

and by (3.2)

\[ V_1 = o(\text{li} \, x). \]

Let

\[ V_2 = \sum_{\mathfrak{p} \in \mathfrak{P}} \prod_{q \mid \mathfrak{p} + 1, q \leq K_2} g(q) - 1. \]

From (3.11) we have that

\[ V_2 \leq \sum_{\mathfrak{p}} \left| \sum_{q \mid \mathfrak{p} + 1, K_1 < q \leq K_2} h(q) + O \left( \sum_{\mathfrak{p}} \sum_{q \mid \mathfrak{p} + 1, K_1 < q \leq K_2} |h^2(q)| \right) = V_3 + O(V_4). \]

Using Bombieri’s result we have that

\[ V_4 < \sum_{K_1 < q \leq K_2} |h^2(q)| \pi(x, q, -1) < c \sum_{q > K_1} \frac{|h^2(q)|}{q - 1} = o(\text{li} \, x). \]

Further, from the Cauchy’s inequality

\[ V_3 < c(\text{li} \, x)^d \left\{ \sum_{K_1 < q_1, q_2 < K_2, q_1 \neq q_2} h(q_1)h(q_2)\pi(x, q_1 q_2, -1) + \sum_{K_1 < q \leq K_2} |h(q)|^2 \pi(x, q, -1) \right\}^\frac{1}{2}. \]

Using Bombieri’s result we have that

\[ V_3 < c(\text{li} \, x)^d \left\{ \sum_{K_1 < q \leq K_2} \frac{h(q)}{q - 1} (\text{li} \, x)^{\frac{d}{2}} + O \left( \frac{x}{\log^2 x} \right) = o(\text{li} \, x), \]

since

\[ \sum_{K_1 < q \leq K_2} \frac{h(q)}{q - 1} = \sum_{K_1 < q \leq K_2} \frac{h(q)}{q} + O \left( \sum_{K_1 < q \leq K_2} \frac{1}{q^2} \right) = o(1). \]

So we proved that

\[ V_2 = V_3 + O(V_4) = o(\text{li} \, x); \quad V_1 = o(\text{li} \, x); \quad V = V_1 + V_2 = o(\text{li} \, x), \]

whence (3.7) follows.

Similarly we have

\[ |I_{K_2}(x) - I_{K_2}(x)| \leq \sum_{p \leq x} \sum_{q \mid p + 1, K_1 < q \leq K_2} h(q) + c \sum_{p \leq x} \sum_{q \mid p + 1, K_2 < q \leq K_2} |h(q)|^2 \]

\[ + c \sum_{K_1 < q \leq K_2} \pi(x, q, -1) = V_5 + cV_6 + cV_7. \]
Using Lemma 3 and (3.4) in Lemma 1 we have that

\[ V_6 \leq \sum_{K_3 < q \leq K_3} |h(q)| \pi(x, q, -1) < c_3 \ln x \sum_{K_3 < q \leq K_3} \frac{|h(q)|}{q} = o(c_3 \ln x). \]

Further using (3.3) and (3.2) we obtain that

\[ V_6 \leq \sum_{K_3 < q \leq K_3} |h^2(q)| \pi(x, q, -1) < c_3 \ln x \sum_{K_3 < q \leq K_3} \frac{|h^2(q)|}{q} = o(c_3 \ln x), \]

\[ V_7 \leq c_4 \ln x \sum_{q > K_3} \frac{1}{q} = o(c_4 \ln x). \]

Hence (3.8) follows.

Finally using Lemma 2 we have

\[ |I(x) - I_{K_3}(x)| \leq 2 \sum_{K_3 < q < x} \pi(x, q, -1) \leq \sum_{j \leq x^\delta} N_j(x) \]

\[ \leq c \sum_{j \leq x^\delta} \frac{x}{\varphi(j) \log^2 x/j} < c \frac{x}{\log^2 x} \sum_{j \leq x^\delta} \frac{1}{\varphi(j)} < c\delta \frac{x}{\log x}, \]

because

\[ \sum_{j \leq y} \frac{1}{\varphi(j)} < c \log y. \]

So the inequality (3.9) is proved, and from (3.6) – (3.9) our theorem follows.

4. Some remarks

1. From our Theorem 2 it follows evidently that if \( g(n) \) is a positive valued multiplicative number-theoretical function such that

\[ \sum_p \frac{((\log g(p))^*)^*}{p} \text{ is convergent,} \]

\[ \sum_p \frac{(\log g(p))^*^2}{p} < +\infty, \]

\[ \sum_{|\log g(p)| > 1} \frac{1}{p} < +\infty, \]

then putting
the distribution functions $F_N(y)$ tend for $N \to +\infty$ to a limiting distribution function $F(y)$ at all points of continuity of $F(y)$.

Hence it follows especially that the functions

$$\frac{\varphi(p+1)}{p+1}, \quad \frac{\sigma(p+1)}{p+1}$$

($\sigma(n)$ denotes the sum of the divisors of $n$) have limiting distribution functions.

2. Recently M. B. Barban, A. I. Vinogradov, B. V. Levin proved that all results of J. P. Kubilius theory (see [9]) are valid for strongly additive arithmetic functions belonging to the class $H$, when the argument runs through “shiffed” primes $\{p-l\}$, (see [10], [11]).

REFERENCES

P. ERDÔS,

H. DELANGE,

A. RÉNYI,

H. DELANGE,

A. RÉNYI,

K. PRACHAR,

E. BOMBIERI,

P. LEVY,

J. P. KUBILIUS,
M. B. Barban,


M. B. Barban, A. I. Vinogradov, and B. V. Levin,


(Oblatum 22-3-'67).