

SÉMINAIRE DELANGE-PISOT-POITOU. THÉORIE DES NOMBRES

JONAS KUBILIUS

On the distribution of number-theoretic functions

Séminaire Delange-Pisot-Poitou. Théorie des nombres, tome 11, n° 2 (1969-1970),
exp. n° 23, p. 1-11

http://www.numdam.org/item?id=SDPP_1969-1970__11_2_A9_0

© Séminaire Delange-Pisot-Poitou. Théorie des nombres
(Secrétariat mathématique, Paris), 1969-1970, tous droits réservés.

L'accès aux archives de la collection « Séminaire Delange-Pisot-Poitou. Théorie des nombres » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ON THE DISTRIBUTION OF NUMBER-THEORETIC FUNCTIONS

by Jonas KUBILIUS

1. - A real or complex-valued function $f(m)$, defined for all positive integers m , is called additive if

$$f(mn) = f(m) + f(n) \quad ,$$

provided m and n are coprime. Analogously, a sequence of real or complex numbers $g(m)$ ($m = 1, 2, \dots$) is called a multiplicative function if

$$g(mn) = g(m) g(n) \quad ,$$

whenever m and n are relatively prime.

As it is well known, many classical number-theoretic functions are additive or multiplicative, and many classical problems of number theory are closely connected with the behaviour of these functions.

In general, the values of such functions are distributed very erratically. Nevertheless it turns out that, in the large, the distribution of values of many of these functions are subject to certain simple laws, which can be formulated and proved by using ideas and methods of probability theory.

We naturally arrive at the concept of asymptotic local and integral distribution laws. In the first case, it is a matter of finding an asymptotic expression for the proportion of natural $m \leq n$ for which a real arithmetic function $h(m)$ assumes a given value k ,

$$\frac{1}{n} N \{m \leq n, h(m) = k\} = v_n \{h(m) = k\} \quad .$$

In the second case, we seek asymptotic expression for

$$\left[\frac{1}{n} N \{m \leq n, h(m) < x\} = v_n \{h(m) < x\} \right] \quad ,$$

where x is any real number.

2. - In case of additive functions, it is very convenient to use the theory of Fourier transforms. If $f(m)$ is a real additive function, then the characteristic function of the distribution function $F_n(x) = v_n \{f(m) < x\}$,

$$\varphi_n(t) = \int_{-\infty}^{\infty} e^{itx} dF_n(x) = \frac{1}{n} \sum_{m=1}^n e^{itf(m)} \quad .$$

If $\varphi_n(t)$ converges, as $n \rightarrow \infty$, for all real t to some function $\varphi(t)$, continuous at the point $t = 0$, then $F_n(x)$ converges to some distribution function $F(x)$ at each of its points of continuity, and $\varphi(t)$ is the characteristic function of $F(x)$. From the rate of convergence of $\varphi_n(t)$ to $\varphi(t)$, we can obtain some information about the rate of convergence of $F_n(x)$ to $F(x)$.

In case of multiplicative functions, a modification of Mellin transform is more convenient. Let $G(x)$ be a distribution function. We introduce the pair of characteristic transforms

$$w_k(t) = \int_{-\infty}^{\infty} |x|^{it} \operatorname{sgn}^k x \, dG(x) \quad (k = 0, 1),$$

where the dash means that the point $x = 0$ is withdrawn from the path of the integration.

There exists a one-to-one correspondence between distribution functions and pairs of their characteristic transforms. This correspondence is continuous in the following sense.

Given a sequence of distribution functions $G_n(x)$ ($n = 1, 2, \dots$), and the sequence of corresponding characteristic transforms $w_{kn}(t)$ ($k = 0, 1; n = 1, 2, \dots$), a necessary and sufficient condition for the convergence of the sequence $G_n(x)$ to a distribution function $G(x)$ at every of its points of continuity and for $G_n(0) \rightarrow G(0)$, $G_n(+0) \rightarrow G(+0)$ (if $x = 0$ is not a continuity point of $G(x)$), is that, for every t , the sequences $w_{kn}(t)$ ($k = 0, 1$) converge to limits $w_k(t)$, which are continuous at $t = 0$. When this condition is satisfied, the limits $w_k(t)$ are identical with the characteristic transforms of the limiting distribution function $G(x)$.

Now, let $g(m)$ be a real-valued multiplicative function. The characteristic transforms of the distribution function $G_n(x) = \nu_n(g(m) < x)$ are the sums

$$w_{kn}(t) = \frac{1}{n} \sum_{\substack{m=1 \\ g(m) \neq 0}}^n |g(m)|^{it} \operatorname{sgn}^k g(m) \quad (k = 0, 1).$$

Instead of the function $G_n(x)$, we can consider the sums $w_{kn}(t)$.

3. - Thus in both cases of additive and multiplicative functions, the problem is to investigate the asymptotic behaviour of the sum

$$(1) \quad \frac{1}{n} \sum_{m=1}^n h(m),$$

where $h(m)$ is a complex-valued multiplicative function satisfying $|h(m)| \leq 1$

and depending on a parameter.

The sums of multiplicative functions were considered by many authors. In particular, H. DELANGE [2] obtained necessary and sufficient conditions in order that the mean value (1) tends to a non-zero limit. From this result, it follows the well-known Erdős-Wintner theorem ([3]).

The distribution function $\nu_n (f(m) < x)$ of a real-valued additive number-theoretic function, $f(m)$, converges to a limit distribution function $F(x)$ at its points of continuity if, and only if, there exists a positive constant c (consequently, for every fixed $c > 0$) such that the series

$$(2) \quad \sum_{|f(p)| > c} \frac{1}{p}, \quad \sum_{|f(p)| < c} \frac{f(p)}{p}, \quad \sum_{|f(p)| < c} \frac{f^2(p)}{p},$$

where the sums are taken over all primes p with the indicated properties, converge. The characteristic function of the limiting distribution function $F(x)$, whenever it exists, is equal to

$$\varphi(t) = \prod_p \left(1 - \frac{1}{p}\right) \sum_{\alpha=0}^{\infty} \frac{e^{itf(p^\alpha)}}{p^\alpha}.$$

In case of integer-valued additive functions, the conditions (2) are equivalent to

$$(3) \quad \sum_{f(p) \neq 0} \frac{1}{p} < \infty.$$

It turns out that this condition is necessary and sufficient for the following local theorem.

THEOREM. - Let $f(m)$ be an integer-valued additive function. For each integer k ,

$$\nu_n(f(m) = k)$$

tends to some λ_k with the property

$$\sum_{k=-\infty}^{\infty} \lambda_k = 1$$

uniformly in k , as $n \rightarrow \infty$, if, and only if, (3) is true.

The proof ([6]) is based on the fact that, for integer-valued additive functions, this statement is equivalent to the integral theorem of Erdős and Wintner.

4. - The case of multiplicative functions is more difficult ([4]).

If for a given real-valued multiplicative function $g(m)$, there exists a constant

$c > 1$ such that the series

$$(4) \quad \sum_{g(p) \leq 1/c} \frac{1}{p}, \quad \sum_{g(p) > c} \frac{1}{p}, \quad \sum_{1/c < g(p) < c} \frac{\ln g(p)}{p}, \quad \sum_{1/c < g(p) < c} \frac{\ln^2 g(p)}{p}$$

converge, then $v_n(g(m) < x)$, as $n \rightarrow \infty$, converges to a limiting distribution function $G(x)$ at its continuity points and at $x = 0$ (in the sense mentioned above). The characteristic transforms of $G(x)$ equal

$$w_k(t) = \prod_p \left(1 - \frac{1}{p}\right) \sum_{\substack{\alpha=0 \\ g(p^\alpha) \neq 0}}^{\infty} \frac{|g(p^\alpha)|^{it} \operatorname{sgn}^k g(p^\alpha)}{p^\alpha} \quad (k = 0, 1) .$$

A distribution function $G(x)$ is said to be symmetric if

$$G(x) = 1 - G(-x + 0) ,$$

for all x .

In case of non-symmetric limiting distribution laws, necessary and sufficient conditions can be given. So, $v_n(g(m) < x)$ tends to a non-symmetric distribution function if, and only if, there exists a constant $c > 1$ such that the series (4) converge and $g(2^\alpha) \neq -1$ for at least one $\alpha = 1, 2, \dots$.

Though necessary and sufficient conditions for the convergence of $v_n(g(m) < x)$ to a limiting symmetric distribution function are unknown, however if $v_n(g(m) < x)$ tends to a non-degenerate distribution function $G(x)$, then $G(x)$ is symmetric if, and only if, at least one of the following conditions is satisfied:

- 1° The series $\sum_{g(p) < 0} \frac{1}{p}$ diverge;
- 2° $g(2^\alpha) = -1$ for all $\alpha = 1, 2, \dots$.

5. - Let us return to the additive functions $f(m)$. If the series

$$\sum_p \frac{f^2(p)}{p}$$

converge, then Erdős-Wintner theorem gives necessary and sufficient conditions for the existence of the asymptotic integral distribution law. The situation is more complicated if these series diverge. In this case, we need to introduce normalization factors. Thus we need to consider the asymptotic laws for

$$v_n(f(m) < A_n + xD_n) ,$$

with some A_n and D_n .

There are some methods for solving such problems ([4], [9]). They consist, in

general, in the following. Instead of the function $f(m)$, we consider a "truncated" function

$$\psi(m) = f(m_r) + v\left(\frac{m}{m_r}\right),$$

where m_r contains only such prime factors which do not exceed a slowly increasing function $r = r(n)$:

$$r(n) \rightarrow \infty, \quad \frac{\ln r(n)}{\ln n} \rightarrow 0.$$

The function $v(m)$ is either identical to 0, or has the property that

$$\left(\sum_{\substack{p \leq n \\ v(p) < u}} \ell_p \right) / \left(\sum_{p \leq n} \ell_p \right),$$

for a given sequence ℓ_p , converges, as $n \rightarrow \infty$, to a distribution function at each of its continuity points.

It turns out that the function $\psi(m)$ can be approximated by a sum of suitably chosen independent random variables. Thus we are in a position to investigate the distribution of the function $\psi(m)$ by means of the theory of summation of independent random variables. An application of an analogue of the law of large numbers leads to the limit theorems for the function $f(m)$ itself.

There is a number of theorems proved by this way ([4], [9]). The class of all possible limit distribution laws contains all stable laws. As an example, 2 mention only one of such theorems([4]).

Let $f(m)$ be a real-valued additive function. Denote

$$A_n = \sum_{p \leq n} \frac{f(p)}{p},$$

$$D_n^2 = \sum_{p^\alpha \leq n} \frac{f^2(p^\alpha)}{p^\alpha},$$

where the summation is taken over prime powers p^α , $\alpha = 1, 2, \dots$. If for every fixed $\varepsilon > 0$,

$$\frac{1}{D_n^2} \sum_{\substack{p^\alpha \leq n \\ |f(p^\alpha)| > \varepsilon D_n}} \frac{f^2(p^\alpha)}{p^\alpha}$$

(an analogue of Lindeberg's condition), then

$$\psi_n(f(m) < A_n + xD_n) \rightarrow \Phi(x),$$

as $n \rightarrow \infty$, tends to the normal law

$$(5) \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du .$$

Similar theorems can be proved for multiplicative functions (4).

6. - In this method, the transition from the "truncated" function $\psi(m)$ to the function $f(m)$ itself introduces certain error, and as result the accuracy of the limit theorems for $f(m)$ is reduced in comparison with the corresponding theorems of probability theory. This is very important if we wish to estimate the rate of convergence to the limit law.

In order to avoid this reduction in accuracy, we need to forego truncation. It turns out that for a certain class of additive and multiplicative functions, the difficulties can be circumvented by using methods of analytic number theory.

For the estimation of the sum

$$(6) \quad \sum_{m \leq x} h(m) ,$$

where $h(m)$ is multiplicative function, $|h(m)| \leq 1$, we consider the generating Dirichlet series

$$Z(s) = \sum_{m=1}^{\infty} \frac{h(m)}{m^s} ,$$

where s is the complex variable. The sum (6) can be represented by a contour integral of $Z(s)$ multiplied by some simple function. In order to evaluate this integral, we need the analytic continuation of $Z(s)$ at least into the half-plane $\operatorname{Re} s > 1$. It can be done in some cases ([5], [7]).

Suppose that the sum over all primes p ,

$$\sum_p |h(p) - x| \frac{\ln p}{p} < c ,$$

for some x not depending on p , $|x| \leq c_1$, where c and c_1 are constants. In this case, the function $Z(s)$ can be represented in the form

$$Z(s) = \zeta^x(s) H(s) ,$$

where $\zeta(s)$ is Riemann's zeta-function, and $H(s)$ is continuous for $\operatorname{Re} s > 1$ and has the first derivative. This leads to the asymptotic formula

$$(7) \quad \sum_{m \leq x} h(m) = \frac{x(\ln x)^{x-1}}{\Gamma(x)} \prod_p \left(1 - \frac{1}{p}\right)^x \left(1 + \sum_{\alpha=1}^{\infty} \frac{h(p^\alpha)}{p^\alpha}\right) + o\left(x \left(\frac{\ln \ln x}{\ln x}\right)^{1/2}\right), \quad x \geq 3 .$$

Here $\Gamma(x)$ denotes the gamma-function. The constant in the symbol \mathcal{O} depends only on c and c_1 .

This formula can be generalized ([7]). Introducing stronger conditions, we can give a formula for (6) containing more terms with decreasing powers of $\ln x$.

Formula (7) permits to estimate error terms in integral and local theorems for a class of additive and multiplicative functions ([7], [8]).

7. - If $f(m)$ is an integer-valued additive function, satisfying the conditions

$$(8) \quad \sum_{f(p) \neq 1} \frac{\ln p}{p} < \infty ,$$

$$(9) \quad \sum_{f(p) \neq 1} \frac{|f(p)|}{p} < \infty , \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{|f(p^\alpha)|}{p^\alpha} < \infty ,$$

then uniformly for all integers k and $n \geq 3$,

$$(10) \quad v_n(f(m) = k) = \frac{\varphi(y)}{\sqrt{\ln \ln n}} + \mathcal{O}\left(\frac{1}{\ln \ln n}\right) ,$$

where

$$y = \frac{k - \ln \ln n}{\sqrt{\ln \ln n}} , \quad \varphi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2} .$$

If

$$|k - \ln \ln n| > (1 + \delta)(2 \ln \ln n \cdot \ln \ln \ln n)^{1/2}$$

or

$$|k - \ln \ln n| < (1 + \delta)(2 \ln \ln n \cdot \ln \ln \ln n)^{1/2} ,$$

where δ is any fixed positive number, then the principal term in (10) is less or greater than the remainder correspondingly. If besides the condition (8), the function $f(m)$ satisfies a stronger condition than (9), we can enlarge the region of the validity of the local theorem. Let us suppose that there exists an integer $r \geq 1$ such that

$$\sum_{f(p) \neq 1} \frac{|f^r(p)|}{p} < \infty , \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{|f^r(p^\alpha)|}{p^\alpha} < \infty .$$

Then uniformly in k and $n \geq 3$,

$$v_n(f(m) = k) = \sum_{\ell=0}^{r-1} P_\ell(y) e^{-y^2/2} (\ln \ln n)^{-(\ell+1)/2} + \mathcal{O}((\ln \ln n)^{-(r+1)/2}) ,$$

where $P_\ell(y)$ is a polynomial of degree 3ℓ with coefficients depending only on the function $f(m)$.

If

$$|k - \ln \ln n| < (1 - \delta)(2r \ln \ln n \cdot \ln \ln \ln n)^{1/2}$$

or

$$|k - \ln \ln n| > (1 + \delta)(2r \ln \ln n \cdot \ln \ln \ln n)^{1/2},$$

then the principal term of the last formula is greater or less than the remainder respectively.

Further restriction on the function $f(m)$ makes it possible to enlarge the region of the validity of the local theorem to $k = \ln \ln n + o(\ln \ln n)$.

Let $f(m)$ satisfy the condition (7), and for some $\gamma > 0$,

$$\sum_{f(p) \neq 1} \frac{e^{\gamma|f(p)|}}{p} < \infty, \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{e^{\gamma|f(p^\alpha)|}}{p^\alpha}.$$

Then for all integers k , satisfying $k = \ln \ln n + o(\ln \ln n)$,

$$\begin{aligned} v_n(f(m) = k) &= (2\pi \ln \ln n)^{-1/2} (\ln n)^{-1} \left(\frac{e \ln \ln n}{k} \right)^k \times \left\{ 1 + O\left(\frac{|k - \ln \ln n| + (\ln \ln \ln n)^3}{\ln \ln n} \right) \right\}. \end{aligned}$$

8. - We consider now an additive function $f(m)$ assuming any real values. Suppose that there exist two constants $c > 0$ and $\lambda \neq 0$ such that

$$\sum_{\substack{a_p < c \\ p}} \frac{\ln p}{p} < \infty,$$

$$(11) \quad \sum_{\substack{a_p > c \\ p}} \frac{a_p}{p} < \infty, \quad \sum_{\substack{a_p > c \\ p}} \frac{\ln p}{p} < \infty, \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{|f(p^\alpha)|}{p^\alpha} < \infty,$$

where $a_p = |f(p) - \lambda|$. Then uniformly in x and $n \geq 3$,

$$v_n(f(m) < \lambda \ln \ln n + x|\lambda| \sqrt{\ln \ln n}) = \Phi(x) + O((\ln \ln n)^{-1/2}).$$

The constant in the symbol O depends only on the function $f(m)$, and $\Phi(x)$ is the normal law (5).

Replacing (11) by

$$\sum_{a_p > c} \frac{a_p^r}{p} < \infty, \quad \sum_{a_p > c} \frac{\ln p}{p} < \infty, \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{|f^r(p^\alpha)|}{p^\alpha} < \infty,$$

where $r \geq 1$ is an integer, we obtain

$$v_n(f(m) < \lambda \ln \ln n + x|\lambda| \sqrt{\ln \ln n}) = \sum_{\ell=0}^{r-1} Q_\ell(-\Phi)(\ln \ln n)^{-\ell/2} + O((\ln \ln n)^{-r/2}).$$

Here $Q_\ell(\mathbf{z})$ is a polynomial of degree 3ℓ with coefficients depending only on the function $f(m)$, and $Q_\ell(-\Phi)$ is obtained from $Q_\ell(-iz)$ by replacing all powers $(iz)^q$ ($q = 0, 1, \dots$) by $\Phi^{(q)}(x)$.

Finally, replacing (11) by

$$\sum_{a_p > c} \frac{e^{\gamma a_p} \ln p}{p} < \infty, \quad \sum_p \sum_{\alpha=2}^{\infty} \frac{e^{\gamma} |f(p^\alpha)|}{p^\alpha},$$

with some $\gamma > 0$, and supposing $x = o(\sqrt{\ln \ln n})$, we obtain that

$$v_n(f(m) < \lambda \ln \ln n + x|\lambda| \sqrt{\ln \ln n})$$

for $x \leq 0$, and

$$v_n(f(m) > \lambda \ln \ln n + x|\lambda| \sqrt{\ln \ln n})$$

for $x \geq 0$, are equal to

$$e^{K_n(x)} \Phi(-|x|) \left(1 + O\left(\frac{|x| + 1}{\sqrt{\ln \ln n}}\right)\right),$$

where

$$K_n(x) = \frac{x^2}{2} + \{\xi - (1 + \xi) \ln(1 + \xi)\} \ln \ln n,$$

$$\xi = \frac{x \operatorname{sgn} \lambda}{\sqrt{\ln \ln n}}.$$

9. - Now let $g(m)$ be a real-valued multiplicative function. If for some constants $c > 0$ and $\lambda \neq 0$,

$$(12) \quad \sum_{\substack{g(p) > 0 \\ b_p < c}} \frac{b_p \ln p}{p} < \infty, \quad \sum_{\substack{g(p) > 0 \\ b_p > c}} \frac{b_p}{p} < \infty, \quad \sum_{\substack{g(p) > 0 \\ b_p > c}} \frac{\ln p}{p} < \infty,$$

where $b_p = |\ln |g(p)| - \lambda|$, and

$$(13) \quad \sum_{g(p) \leq 0} \frac{\ln p}{p} < \infty ,$$

$$\sum_p \sum_{\substack{\alpha=2 \\ g(p^\alpha) \neq 0}}^{\infty} \frac{|\ln |g(p^\alpha)| |}{p^\alpha} < \infty ,$$

then uniformly in x and $n \geq 3$,

$$v_n(g(m) < |x| |\lambda| (\ln \ln n)^{1/2} \ln^\lambda n \cdot \text{sgn } x)$$

is equal to

$$1 - \frac{\omega_0 + \omega_1}{2} \Phi(-\ln x) + O((\ln \ln n)^{-1/2})$$

for $x > 0$, and to

$$\frac{\omega_0 - \omega_1}{2} \Phi(-\ln |x|) + O((\ln \ln n)^{-1/2})$$

for $x < 0$. Here

$$\omega_k = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \sum_{\substack{\alpha=1 \\ g(p^\alpha) \neq 0}}^{\infty} \frac{\text{sgn}^k g(p^\alpha)}{p^\alpha}\right) \quad (k = 0, 1) .$$

An analogous result is true, if we replace $g(p) > 0$ by $g(p) < 0$ in the inequalities (12), and $g(p) \leq 0$ by $g(p) \geq 0$ in the inequality (13). In this case,

$$v_n(g(m) < |x| |\lambda| (\ln \ln n)^{1/2} \ln^\lambda n \cdot \text{sgn } x)$$

is equal to

$$\left(1 - \frac{\omega_0}{2}\right) \Phi(-\ln x) + O((\ln \ln n)^{-1/2})$$

for $x > 0$, and to

$$\frac{\omega_0}{2} \Phi(-\ln |x|) + O((\ln \ln n)^{-1/2})$$

for $x < 0$.

BIBLIOGRAPHY

- [1] BAKŠTYS (A.). - On the asymptotic distribution of multiplicative number-theoretic functions [in russian, with lithuanian and german summaries], Lietuvos matematikos Rinkinys, t. 8, 1968, p. 5-20, 201-219, 643-680.
- [2] DELANGE (H.). - Sur les fonctions arithmétiques multiplicatives, Ann. scient. Ec. Norm. Sup., 3e série, t. 78, 1961, p. 273-304.
- [3] ERDŐS (P.) and WINTNER (A.). - Additive arithmetical functions and statistical independence, Amer. J. of Math., t. 61, 1939, p. 713-721.
- [4] KUBILIUS (J.). - Probabilistic methods in the theory of numbers. - Providence, American mathematical Society, 1964 (Translations of mathematical Monographs, 11).
- [5] KUBILIUS (J.). - On local theorems for additive number-theoretic functions, Abhandlungen aus Zahlentheorie und Analysis : Zur Erinnerung an Edmund Landau (1877-1938), p. 175-191. - Berlin, VEB Deutscher Verlag der Wissenschaften, 1968.
- [6] KUBILIUS (J.). - A local theorem for additive arithmetic functions, Lietuvos matematikos Rinkinys, t. 11, 1971 (to appear).
- [7] KUBILIUS (J.). - The method of generating Dirichlet series in the theory of the distribution of additive arithmetic functions, Lietuvos matematikos Rinkinys, t. 11, 1971 (to appear).
- [8] KUBILIUS (J.) and JUŠKYS (Z.). - On the distribution of multiplicative arithmetic functions, Lietuvos matematikos Rinkinys, t. 11, 1971 (to appear).
- [9] LEVIN (B.) and FAINLEIB (A.). - Integral limit theorems for some classes of additive arithmetic functions [in russian], Trudy Moskovskogo Matematičeskogo Obščestva, t. 8, 1968, p. 19-54.

(Texte reçu le 29 juin 1970)

Jonas KUBILIUS
 Université de Vilnius
 VILNIUS, Lithuanie
 (Union soviétique)
