

On random multiplicative functions

By I. KÁTAI in Budapest

1. We call $f(n)$ a completely multiplicative (c.m.) function, if $f(mn) = f(m)f(n)$ holds for all pairs m, n of positive integers. Let \mathcal{F} be the set of those c.m. functions which take the values $+1$ and -1 only.

We say that a function $f(n) \in \mathcal{F}$ is of *normal type*, if

$$(1.2) \quad \lim_x x^{-1} N\{n \leq x; f(n+i) = \varepsilon_i, i=0, \dots, k\} = \frac{1}{2^{k+1}}$$

for $k=0, 1, 2, \dots$ and for all choices of $\varepsilon_0 = \pm 1, \dots; \varepsilon_k = \pm 1$.

It would be interesting to give a necessary and sufficient condition for $f(n)$ to be of normal type. Recently E. WIRSING [1] proved that a function $f(n) \in \mathcal{F}$ satisfies (1.1) with $k=0$ if and only if

$$(1.2) \quad \sum_{f(p)=-1} \frac{1}{p} = \infty.$$

As is easy to see, the validity of (1.2) is not sufficient for normality. Let for example $f(n)$ be defined as follows: $f(2)=1$, and for an odd prime p let $f(p)=1$ or -1 according as $p \equiv 1$ or $-1 \pmod{4}$. Then, by an easy calculation we have

$$\sum_{n \leq x} f(n)f(n+4) = \frac{x}{4} + o(x);$$

hence it follows that $f(n)$ is not a normal function.

We shall see in the following section that almost all multiplicative functions are of normal type. One would think that the Liouville function $\lambda(n)$ is normal. However we can only prove that the system $\lambda(n) = \varepsilon_1, \lambda(n+1) = \varepsilon_2$ has infinitely many solutions for an arbitrary choice of $\varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1$. This is a special case of the assertions which we shall prove in the section 3.

2. Let c, c_1, c_2, \dots denote suitable positive constants; let $\varepsilon, \varepsilon_1, \varepsilon_2, \dots$ be arbitrary small positive constants not necessarily the same at every occurrence. Let $d_k(n)$ denote the number of solutions of the equation $n = x_1, \dots, x_k$ in positive integers x_1, \dots, x_k , and let $d_2(n) = d(n)$.

Let p_n denote the n th prime number. Let (Ω, \mathcal{A}, P) be a probability space and $\xi_n = \xi_n(\omega)$ ($n=1, 2, \dots$) be a sequence of independent random variables with the distribution $P(\xi_n = 1) = P(\xi_n = -1) = \frac{1}{2}$. Let $f(n; \omega)$ be a completely multiplicative function which we define on the set of primes by $f(p_n; \omega) = \xi_n(\omega)$.

We have

Theorem 1. *Almost all $f(n; \omega)$ are of normal type.*

For the proof we need some lemmas.

Lemma 1. *For positive integers C, D let $N(z; D, C)$ denote the number of solutions of the diophantine equation*

$$(2.1) \quad x^2 - Dy^2 = C$$

in positive integers x, y satisfying $x \leq z$. Then

$$(2.2) \quad N(z; D, C) \leq c_1 d(C^2) \log 2Dz.$$

Perhaps this lemma is known, but I was unable to find a reference to it. We prove now (2.2).

Without any restriction we can assume that D is a square-free number. For $D=1$ inequality (2.2) obviously holds, therefore we assume that $D>1$.

Let $K(\sqrt{D})$ denote the quadratic extension field over the rational number-field generated by \sqrt{D} . Let R denote the ring of the algebraic integers in $K(\sqrt{D})$, and for a general $\gamma \in R$ let (γ) denote the principal ideal generated by γ .

For a general solution x, y of (2.1) let $\alpha = x + \sqrt{D}y$, $\beta = x - \sqrt{D}y$. Let $(C) = \pi_1^{r_1} \dots \pi_r^{r_r}$, where π_1, \dots, π_r are different prime ideals. Using the fact that the norm of the ideals is a multiplicative function and that $N((C)) = C^2$, furthermore that $N(\pi_i)$ is a prime number or a square of a prime number we have $\prod_{i=1}^r (\gamma_i + 1) \leq d(C^2)$. Since $\alpha\beta = C$ and $\alpha, \beta \in R$, therefore $(\alpha)(\beta) = (C)$ and so $(\alpha) | (C)$. Hence it follows that all the solutions can be classified into at most $d(C^2)$ classes; where two solutions $x, y; x_1, y_1$ belong to the same class if and only if $(\alpha) = (x + \sqrt{D}y) = (\alpha_1) = (x_1 + \sqrt{D}y_1)$. Now we prove that the number of solutions of (2.1) belonging to a fixed class does not exceed $c_1 \log 2Dz$, whence (2.2) immediately shall follow.

Let (x_v, y_v) $v=0, 1, \dots, M$ be the all solutions in a class satisfying $1 \leq x_0 \leq x_1 \leq \dots \leq x_M \leq z$, $y_v \geq 0$ and let $\alpha_v = x_v + y_v \sqrt{D}$, $\beta_v = x_v - y_v \sqrt{D}$. We have $(\alpha_0) = (\alpha_1) = \dots = (\alpha_M)$. Therefore $\alpha_v = \alpha_\mu \varepsilon_{v\mu}$, $\beta_v = \beta_\mu \varrho_{v\mu}$, where $\varepsilon_{v\mu}, \varrho_{v\mu}$ are units in R . Since $C = \alpha_v \beta_v = \alpha_\mu \beta_\mu \varrho_{v\mu} \varepsilon_{v\mu} = \varrho_{v\mu} \varepsilon_{v\mu} C$, we have $\varrho_{v\mu} = \varepsilon_{v\mu}^{-1}$. Using the Dirichlet theorem concerning the form of the units we see that all units have form $\pm \varepsilon_0^n$ ($n=0, \pm 1, \pm 2, \dots$), where $\varepsilon_0 = \frac{u_0 + \sqrt{D}v_0}{2}$, and u_0, v_0 are suitable positive

integers satisfying $u_0^2 - Dv_0^2 = 4$. Hence $\varepsilon_0 > \frac{\sqrt{D}}{2}$ and we can assume that $\alpha_n = \alpha_0 \varepsilon^n$.

Using that $x_n \leq z$ and that by (2.1) $y_n \leq \sqrt{\frac{C+z^2}{D}} \leq \frac{Cz}{D}$, we have $\alpha_n \leq (C+1)z$ ($n=0, \dots, M$). On the other hand, by $\alpha_0 \beta_0 = C$, $0 < \beta_0 < \alpha_0$ we have $\alpha_0 > 1$. Hence $\varepsilon^n < (C+1)z$, whence $M \leq \frac{\log(C+1)z}{\log \varepsilon_0} \leq c_1 \log 2Cz$ follows. This completes the proof of Lemma 1.

Corollary. For positive integers A, B, C let $N(z; A, B, C)$ denote the number of solutions of

$$(2.3) \quad Ax^2 - By^2 = C$$

in positive integers x, y , $x \leq z$. Then

$$N(z; A, B, C) \leq N(Az; AB, AC) \leq c_1 d(A^2 C^2) \log 2A^2 Bz.$$

This is obvious. If (x, y) is a solution of (2.3) then (Ax, y) is a solution of $X^2 - ABY^2 = AC$ which proves the Corollary.

Lemma 2. (Borel—Cantelli) Let A_1, A_2, \dots be an infinite sequence of sets in (Ω, A, P) and let $\sum_{j=1}^{\infty} P(A_j) < \infty$. Then almost all ω in Ω are belonging to finitely many A_i only.

Proof of Theorem 1. Let $0 < i_1 < i_2 < \dots < i_k$ be arbitrary but fixed integers. For a general integer n let $\bar{n} = (n + i_1) \dots (n + i_k)$. Let us introduce the notation

$$(2.4)-(2.5) \quad \eta_N(\omega) = \sum_{n=1}^N f(\bar{n}, \omega); \quad M_{I, N} = \int_{\Omega} (\eta_N(\omega))^I dP.$$

First we give a non-trivial estimation for $M_{4, N}$, whence by using the Borel—Cantelli lemma we deduce that $\lim_{N \rightarrow \infty} \eta_N(\omega)/N = 0$ for almost all $\omega \in \Omega$.

It is obvious, that

$$M_{4, N} = \sum_{n_1, n_2, n_3, n_4} \int_{\Omega} f(\bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4; \omega) dP,$$

where in the sum n_1, n_2, n_3, n_4 run independently over the values $1, 2, \dots, N$. Using $\int_{\Omega} f(m; \omega) dP = 1$ or 0 according to m is a square-number, or not, we have that $M_{4, N}$ is equal to the number of solutions of the equation

$$(2.6) \quad \bar{n}_1 \bar{n}_2 \bar{n}_3 \bar{n}_4 = X^2$$

in unknowns n_1, n_2, n_3, n_4, X , satisfying $1 \leq n_i \leq N$ ($i=1, 2, 3, 4$).

For a fixed square-free integer $E(>0)$ let $H(E)$ denote the number of solutions of the equation

$$\bar{n}_1 \bar{n}_2 = EY^2; \quad 1 \leq n_1 \leq N, \quad 1 \leq n_2 \leq N$$

in unknowns n_1, n_2, Y .

It is obvious that if n_1, n_2, n_3, n_4 is a solution of (2.6) then the square-free parts of the numbers $\bar{n}_1 \bar{n}_2, \bar{n}_3 \bar{n}_4$ are the same. Hence we have

$$M_{4,N} = \sum_E H^2(E),$$

and consequently

$$(2.7) \quad M_{4,N} \leq (\max_E H(E)) \sum_E H(E).$$

Observing that $\sum_E H(E) = N^2$ (since the number of the choice of all pairs $n_1, n_2, 1 \leq n_i \leq N$ is N^2) we have

$$(2.8) \quad M_{4,N} \leq N^2 \max_E H(E).$$

Now we estimate $H(E)$. For a general positive square-free A let $G(A)$ denote the number of $n \leq N$ which can be written in the form

$$(2.9) \quad \bar{n} = AZ^2,$$

where Z is a suitable integer. Then we have

$$(2.10) \quad H(E) \leq \sum_{E_1 E_2 = E} \sum_U G(E_1 U) G(E_2 U),$$

where in the right hand side E_1 runs over the divisors of E and U over the set of all square-free integers coprime to E .

For $k=1$ we evidently have $G(A) \leq \sqrt{N/A}$. Consequently by (2.10)

$$H(E) \leq \sum_{E_1 E_2 = E} \frac{N}{\sqrt{E}} \cdot \sum_{U \leq N} \frac{1}{U} \leq \frac{N \log N}{\sqrt{E}} d(E) \leq cN \log N,$$

and hence by (2.8)

$$(2.11) \quad M_{4,N} \leq cN^3 \log N.$$

Assume now that $k \geq 2$. Consider the solutions of $\bar{n} = AZ^2$. Since the numbers $n+i_{j_1}, n+i_{j_2}$ have no common prime-divisors greater than $i_{j_2}-i_{j_1}$ if $j_1 \neq j_2$, for an n satisfying (2.9) we have

$$(2.12) \quad n+i_j = R_j C_j Z_j^2 \quad (j=1, 2, \dots, k),$$

where R_j, C_j are square-free numbers, the prime factors of R_j are not greater than

$i_k - i_1$ and the prime factors of C_j are greater than $i_k - i_1$ and $\prod_{j=1}^k C_j | A$. If n is a solution of (2.12), then

$$(2.13) \quad i_2 - i_1 = R_2 C_2 Z_2^2 - R_1 C_1 Z_1^2$$

holds with suitable $Z_1, Z_2 \leq N$. Using the Corollary to Lemma 1 we have that the number of solutions of (2.13) with $Z_1, Z_2 \leq N$ is at most $c_1 d((R_1 C_1 (i_2 - i_1))^2) \log N \leq c_1 N^{\epsilon_1}$.

The number of all possible pairs of R_1, R_2 occurring in (2.12) is bounded for fixed i_1, i_2, \dots, i_k . The number of couples (R_1, R_2) is at most $d^2(A) \leq cN^{\epsilon_2}$, since $C_1 C_2 | A$. Therefore

$$(2.14) \quad G(A) \leq cN^{\epsilon}.$$

Using (2.10) and the fact that the number of those A which occur as the square-free part of a number \bar{n} for some $n \leq N$ is at most N , we have

$$H(E) \leq cN^{1+\epsilon}.$$

Hence by (2.8)

$$(2.15) \quad M_{4,N} \leq cN^{3+\epsilon}$$

follows.

Using (2.11) or (2.15) according as $k=1$ or $k \geq 2$, we have

$$(2.16) \quad P(|\eta_N| > N^{\alpha}) \leq \int_{\Omega} \frac{|\eta_N|^4}{N^{4\alpha}} dP < cN^{3-4\alpha+\epsilon}.$$

Let $N_m = m^5$ and $\alpha = \frac{4}{5} + \epsilon$. By (2.16) we have

$$\sum_{m=1}^{\infty} P(|\eta_{N_m}| > N_m^{\frac{4}{5}+\epsilon}) \leq c \sum_{m=1}^{\infty} m^{-1-\epsilon} < \infty.$$

Consequently by Lemma 2 we have

$$(2.17) \quad \lim_{m \rightarrow \infty} \frac{\eta_{N_m}(\omega)}{N_m^{\frac{4}{5}+2\epsilon}} = 0$$

for all fixed $\epsilon > 0$ and for almost all $\omega \in \Omega$. Since for $N_m \leq N < N_{m+1}$

$$(2.18) \quad |\eta_N - \eta_{N_m}| \leq N - N_m \leq N_{m+1} - N_m \leq cm^4 \leq cN_m^{\frac{4}{5}} < cN^{\frac{4}{5}},$$

therefore by (2.17)

$$\lim_{N \rightarrow \infty} \frac{\eta_N(\omega)}{N^{\frac{4}{5}+2\epsilon}} = 0$$

for all $\epsilon > 0$ and almost all $\omega \in \Omega$.

Finally we remark, that a function $f(n) \in \mathcal{F}$ is of normal type if and only if $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\bar{n}) = 0$ for all choice of $k=1, 2, \dots$ and of (i_1, \dots, i_k) . This completes the proof of the theorem.

3. Theorem 2. *Let $f(n)$ be a completely multiplicative function, all values of which are $+1$ or -1 . Assume that there exist at least two primes p_1, p_2 for which $f(p_1)=f(p_2)=-1$. Then for arbitrary $\varepsilon_1, \varepsilon_2$ ($\varepsilon_1 = \pm 1, \varepsilon_2 = \pm 1$) there exist infinitely many n satisfying $f(n)=\varepsilon_1, f(n+1) = \varepsilon_2$.*

For $(\varepsilon_1, \varepsilon_2)=(+1, +1)$ or $(-1, 1)$ we can prove a stronger assertion. This is stated in Theorems 3 and 4.

Theorem 3. *Assuming that the series*

$$(3.1) \quad \sum_{f(p)=-1} \frac{1}{p}$$

diverges we have

$$(3.2) \quad \liminf_{x \rightarrow \infty} x^{-1} N_f(x, 1, 1) \cong \frac{1}{12},$$

$$(3.3) \quad \liminf_{x \rightarrow \infty} x^{-1} N_f(x, -1, -1) \cong \frac{1}{12},$$

where $N_f(x, \varepsilon_1, \varepsilon_2)$ denotes the number of those n not exceeding x for which $f(n)=\varepsilon_1, f(n+1) = \varepsilon_2$. Consequently

$$(3.4) \quad \liminf_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x} f(n)f(n+1) \cong -\frac{5}{6}.$$

Let \mathcal{P} be the set of those primes p for which $f(p) = -1$.

Theorem 4. *Suppose that \mathcal{P} contains at least two elements and that the series $\sum_{p \in \mathcal{P}} \frac{1}{p}$ converges. Then for both values $\varepsilon=1, -1$ we have*

$$(3.5) \quad \lim_{x \rightarrow \infty} x^{-1} N_f(x, \varepsilon, \varepsilon) = \frac{1}{4} \left(1 + 2\varepsilon \prod_{p \in \mathcal{P}} \frac{p-1}{p+1} + \prod_{p \in \mathcal{P}} \frac{p-3}{p+1} \right) \stackrel{(\text{def})}{=} A.$$

The number standing on the right-hand side of (3.5) is positive.

Proof of Theorems 2, 3, and 4. First we prove Theorem 2 for $(\varepsilon_1, \varepsilon_2) = (1, -1)$ and $(-1, 1)$. The remaining two cases will follow from Theorems 3 and 4.

The assertion of Theorem 2 for $(\varepsilon_1, \varepsilon_2)=(1, -1)$ and $(-1, 1)$ is equivalent to saying that $f(n)$ assumes both of the values $+1$ and -1 infinitely many times. For

$+1$ this is true since $f(n^2) = +1$ for all n . To show this for -1 let p be a prime for which $f(p) = -1$. Then $f(p^{2k+1}) = -1$ for all k .

To prove Theorem 3 we need a theorem due to E. WIRSING [1], which we state as

LEMMA 3. *If $f(n) = \pm 1$ and the series (3.1) diverges, then*

$$(3.6) \quad x^{-1} \sum_{n \leq x} f(n) \rightarrow 0 \quad \text{as } x \rightarrow \infty.$$

Let $n_1 < n_2 < \dots < n_L \leq x$ be the sequence of those integers for which $f(n_i) = -1$. Let $m_1 < m_2 < \dots < m_R \leq x$ denote the complementary sequence, i.e. for which $f(m_i) = +1$. Let $\varrho_k(x)$ denote the number of those n_i for which $n_{i+1} - n_i = k$, $n_i \leq x$. Similarly, let $\tau_k(x)$ denote the number of m_i 's satisfying $m_{i+1} - m_i = k$, $m_i \leq x$. From (3.6) we easily deduce

$$(3.7) \quad L + O(1) = \sum_{k=1}^{\infty} \varrho_k(x) = \frac{x}{2} + o(x), \quad R + O(1) = \sum_{k=1}^{\infty} \tau_k(x) = \frac{x}{2} + o(x)$$

$$(3.8) \quad \sum_{k=1}^{\infty} k \varrho_k(x) = x + o(x), \quad \sum_{k=1}^{\infty} k \tau_k(x) = x + o(x).$$

Hence

$$(3.9) \quad \sum_{k=3}^{\infty} (k-2) \varrho_k(x) = \varrho_1(x) + o(x), \quad \sum_{k=3}^{\infty} (k-2) \tau_k(x) = \tau_1(x) + o(x)$$

follow. Consequently

$$(3.10) \quad \sum_{k \neq 2} k \varrho_k(x) \leq 4 \varrho_1(x) + o(x), \quad \sum_{k \neq 2} k \tau_k(x) \leq 4 \tau_1(x) + o(x),$$

$$(3.11) \quad \sum_{k \neq 2} \varrho_k(x) \leq 2 \varrho_1(x) + o(x), \quad \sum_{k \neq 2} \tau_k(x) \leq 2 \tau_1(x) + o(x).$$

Now we prove that $\lim_{x \rightarrow \infty} \inf x^{-1} \varrho_1(x) \geq \frac{1}{12}$. The proof of the relation

$\lim_{x \rightarrow \infty} \inf x^{-1} \tau_1(x) \geq \frac{1}{12}$ is similar, and so we omit it. Since from (3.6)

$$\frac{1}{x} \sum_{2n \leq x} f(2n) \rightarrow 0$$

follows, among the n 's the number of even numbers is $\frac{x}{4} + o(x)$. Hence by (3.11)

we have that there are at least $\frac{x}{4} - 2 \varrho_1(x) - o(x)$ even n 's satisfying $n_{i+1} - n_i = 2$.

Let \mathcal{S} denote the set of these n 's.

We distinguish two cases.

Case a). $f(2) = 1$. Then for $n_i \in \mathcal{S}$ the integers $n_i/2$ and $n_{i+1}/2$ are consecutive numbers, and furthermore $f\left(\frac{n_i}{2}\right) = f\left(\frac{n_{i+1}}{2}\right) = -1$, $\frac{n_i}{2} \equiv \frac{x}{2}$. Thus we have

$$\varrho_1\left(\frac{x}{2}\right) \equiv \frac{x}{4} - 2\varrho_1(x) - o(x),$$

whence $3\varrho_1(x) \equiv \frac{x}{4} - o(x)$, i.e. $\liminf \frac{\varrho_1(x)}{x} \equiv \frac{1}{12}$, follows.

Case b). $f(2) = -1$. Then, for $n_i \in \mathcal{S}$, $\frac{n_i}{2}$ and $\frac{n_{i+1}}{2}$ are consecutive integers, and moreover $f\left(\frac{n_i}{2}\right) = f\left(\frac{n_{i+1}}{2}\right) = +1$, $\frac{n_i}{2} \equiv \frac{x}{2}$. Consequently

$$(3.12) \quad \tau_1\left(\frac{x}{2}\right) \equiv \frac{x}{4} - 2\varrho_1(x) + o(x).$$

Since the interval $[m_i, m_{i+1}]$ for $m_{i+1} - m_i = k$, $k \equiv 3$ contains $(k-1)$ elements from among the n 's, we deduce that

$$\varrho_1(x) \equiv \sum_{k=3}^{\infty} (k-2)\tau_k(x);$$

hence by (3.12)

$$(3.13) \quad \varrho_1(x) \equiv \tau_1(x) + o(x)$$

follows. From here by (3.12) we obtain

$$3\varrho_1(x) \equiv \frac{x}{4} + o(x),$$

$$\text{i.e.} \quad \lim_{x \rightarrow \infty} x^{-1} \varrho_1(x) \equiv \frac{1}{12}.$$

Now we prove Theorem 4. For this we need the following

Lemma 4. [2] If $h(n)$ is a complex-valued completely multiplicative function satisfying the conditions: a) $|h(n)| \leq 1$ ($n = 1, 2, \dots$), and b) $\sum_p \frac{h(p)-1}{p}$ converges, then

$$\lim_{x \rightarrow \infty} x^{-1} \sum_{n \leq x} h(n)h(n+1) = \prod_p \left(1 + 2 \sum_{a=1}^{\infty} \frac{h(p^a) - h(p^{a-1})}{p^a} \right),$$

$$\lim_{x \rightarrow \infty} x^{-1} \sum_{n \leq x} h(n) = \prod_p \left(1 + \sum_{a=1}^{\infty} \frac{h(p^a) - h(p^{a-1})}{p^a} \right).$$

Observing that the conditions of Lemma 4 are satisfied for $h(n)=f(n)$ and that

$$4N(x, \varepsilon, \varepsilon) = \sum_{n \leq x} (f(n) + \varepsilon)(f(n+1) + \varepsilon) = \sum_{n \leq x} f(n)f(n+1) + 2\varepsilon \sum_{n \leq x} f(n) + x + O(1),$$

by Lemma 4 we obtain (3.5).

Finally we prove the positivity of A . If $3 \in \mathcal{P}$, then

$$A \cong \frac{1}{4} \left(1 - 2 \cdot \frac{2}{4} \prod_{\substack{p \in \mathcal{P} \\ p \neq 3}} \frac{p-1}{p+1} \right).$$

Since $\prod_{\substack{p \in \mathcal{P} \\ p \neq 3}}$ is not an empty product, it must be smaller than 1; so indeed $A > 0$. If

$3 \notin \mathcal{P}$, $2 \in \mathcal{P}$, then

$$A \cong \frac{1}{4} \left(1 - \frac{2}{3} \prod_{p \in \mathcal{P}, p > 3} \frac{p-1}{p+1} - \frac{1}{3} \prod_{p \in \mathcal{P}, p > 3} \frac{p-3}{p+1} \right).$$

Using the fact that the products on the right hand side are not empty, we again have $A > 0$. If 2, 3 are not belonging to \mathcal{P} , then

$$A \cong \frac{1}{4} \left(1 - 2 \prod_{p \in \mathcal{P}, p > 3} \frac{p-1}{p+1} + \prod_{p > 3} \frac{p-3}{p+1} \right).$$

Using the relation $\frac{p-3}{p+1} < \left(\frac{p-1}{p+1} \right)^2$ for $p \geq 3$,

$$A \cong \frac{1}{4} \left(1 - \prod_{p \in \mathcal{P}} \frac{p-1}{p+1} \right)^2 > 0$$

also in this case.

References

- [1] E. WIRSING, Das asymptotische Verhalten von Summen über multiplikative Funktionen. II, *Acta Math. Acad. Sci. Hung.*, **18** (1967), 411—467.
- [2] I. KÁTAI, On the distribution of arithmetical functions, *ibidem* (to appear).

(Received January 20, 1969)