SOME ASYMPTOTIC FORMULAS FOR MULTIPLICATIVE FUNCTIONS

P. ERDÖS

The present paper contains several asymptotic formulas for the sum of multiplicative functions. A function $f(n)$ is called multiplicative if $f(a \cdot b) = f(a) \cdot f(b)$ for $(a, b) = 1$. We assume $f(n) > 0$. In this paper $f(n)$, $f_1(n)$ will always denote multiplicative functions. First we prove the following theorem.

THEOREM 1.[1] *Assume that the two series*

$$(1) \qquad -\sum_{p,\alpha} \frac{f(p^\alpha) - 1}{p^\alpha}, \qquad \sum_{p,\alpha} \frac{(f(p^\alpha) - 1)^2}{p^\alpha},$$

converge; then $f(n)$ has a mean value, that is,

$$\lim_{n\to\infty} \frac{1}{n} \sum f(m)$$

exists and is not equal to zero.

This result was conjectured in a slightly more special form at the end of my paper *Some remarks on additive and multiplicative functions.*[2]

REMARK. The convergence of (1) is the necessary and sufficient condition for the existence of the distribution function of $f(n)$.[3]

For the sake of simplicity we assume $f(p^\alpha) = f(p)$. Then we prove

$$(2) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{m \leq n} f(m) = \prod_p \left(1 + \frac{f(p) - 1}{p}\right).$$

It easily follows from (1) that the product on the right side of (2) converges and thus the value of the limit is not 0.

We easily obtain from (1) that for every $\epsilon > 0$

$$\sum_{|f(p)-1|>\epsilon} \frac{1}{p} < \infty.$$

Received by the editors May 27, 1946, and, in revised form, December 11, 1946.

[1] This result generalizes a result of Wintner, Amer. J. Math. vol. 67 (1945) pp. 481–485.

[2] Bull. Amer. Math. Soc. vol. 52 (1946) pp. 527–537.

[3] P. Erdös and A. Wintner, Amer. J. Math. vol. 61 (1939) pp. 713–721. See also footnote 2.

Thus by arguments used in previous papers[4] we can assume, for the sake of simplicity (without loss of generality), that $f(p)\to 1$ as $p\to\infty$.

Define

$$f_k(m) = \prod_{p\mid m,\, p\leq p_k} f(p).$$

Also $M(x, n)$ and $M_k(x, n)$ (x large) denote the number of integers $m\leq n$ for which $f(m)\geq x$, $f_k(m)\geq x$, respectively. Further let

$$A(m) = \prod_{p^{\alpha}\|m}{}' p^{\alpha}, \qquad B(m) = \prod_{p^{\alpha}\|m}{}'' p^{\alpha} \qquad\qquad (m = A(m)\cdot B(m)).$$

Here $p^{\alpha}\| m$ means that $p^{\alpha}\mid m$ and $p^{\alpha+1}\nmid m$, and the prime denotes that the product is extended over the $p\leq n^{1/x^{10}}$, and the double prime denotes that the product is extended over the $p>n^{1/x^{10}}$.

First we have to prove some lemmas.

LEMMA 1. *The number N of integers $m\leq n$ with*

$$A(m) \geq n^{1/2}$$

is $o(n/x^4)$ (we assume that $x\to\infty$).

In the product $\prod_{m\leq n}A(m)$, the prime p occurs as many times as p divides n (and $p\leq n^{1/x^{10}}$). Hence,

$$\prod_{m\leq n} A(m) < \prod_{p}{}' p^{n/(p-1)} < \exp\left(\frac{cn\log n}{x^{10}}\right)$$

since $\sum_{p\leq x}\log p/p < c\log x$ (the prime means $p\leq n^{1/x^{10}}$). Multiplying together the inequalities $n^{1/2}\leq A(m)$ we have

$$n^{N/2} \leq \prod_{m\leq n} A(m) < \exp\left(\frac{cn\log n}{x^{10}}\right),$$

or

$$N < c\frac{2n}{x^{10}},$$

which proves the lemma.

LEMMA 2. *The number N' of integers $m\leq n$ with $f(B(m))\geq x^{1/2}$* is *$o(n/x^4)$.*

[4] P. Erdös, J. London Math. Soc. vol. 10 (1935) p. 124. (By the method used there we can show that without loss of generality we can neglect a sequence of primes p_i with $\sum 1/p_i<\infty$.)

Since $f(p)\to 1$ and all prime factors of $B(m)$ are greater than $n^{1/x^{10}}$, we obtain that if $p\mid B(m)$, $f(p)<1+\epsilon$. Thus if $f(B(m))\geqq x^{1/2}$ we conclude that m has at least 100 log x prime factors greater than $n^{1/x^{10}}$. Thus on the one hand

$$\sum^*_{f(B(m))\geqq x^{1/2}} \frac{n}{B(m)} > N' ;$$

on the other hand, where the star indicates that each $B(m)$ is counted only once,

$$\sum^*_{f(B(m))\geqq x^{1/2}} \frac{1}{B(m)} < \left(\sum{}' \frac{1}{p}\right)^u \Big/ u!,$$

where $u=[100 \log x]$ and the prime indicates that $n^{1/x^{10}}<p\leqq n$. Now

$$\sum{}' \frac{1}{p} < \log\log n - \log\log (n^{1/x^{10}}) + o(1) < 11 \log x;$$

hence

$$N' < n \frac{(11 \log x)^u}{u!} = o\left(\frac{n}{x^4}\right),$$

which proves the lemma.

Lemma 3. *There exists an absolute constant c (independent of k) such that for all x and n*

$$M(x, n) < cn/x^3 \quad \textit{and} \quad M_k(x, n) < cn/x^3.$$

Remark. Lemma 3 is not trivial only for large x and n. It will be clear from the proof that the lemma is true with an arbitrary t instead of 3. It will be clear from the proof that it suffices to consider $M(x, n)$.

Suppose the lemma is false. Then we clearly can assume that there exist infinite sequences x_i and n_i such that

$$(3) \qquad M(x_i, n_i) > 1/x_i^4, \qquad x_i\to\infty,\ n_i\to\infty.$$

Let $a_1<a_2<\cdots<a_k\leqq n_i$ be the integers not greater than n_i with $f(a_j)\geqq x$. For simplicity of notation we replace x_i by x and n_i by n where there is no danger of confusion. We obtain from Lemmas 1 and 2 that there exist at least $n/2x^4$ a's for which

$$(4) \qquad f(A(a_j)) > x^{1/2}, \quad A(a_j) < n^{1/2} \quad (\text{since } f(m) = f(A(m))f(B(m))).$$

Denote these a's by a^+. Consider the integers $m\leqq n$ for which $A(m)=t$. We must have $m=tv\leqq n$, where v is not divisible by any $p\leqq n^{1/x^{10}}$, since these primes are included in t, and $v \leqq n/t$. The number of such integers v is equal to the number of integers $m \leqq n$ with $A(m)=t$. Since we are later going to let t run through the $A(a_j^+)$, it is permissible to assume $t < n^{1/2}$ by (4). Then Brun's method[5] yields the result

$$\sum_{A(m)=t} 1 \leqq c\frac{n}{t}\prod_p{}'\left(1-\frac{1}{p}\right) < c\,\frac{nx^{10}}{t\log n}; \tag{5}$$

the prime indicates that $p \leqq n^{1/x^{10}}$. Now letting t run through the $A(a_j^+)$ we obtain that the number of a_j^+'s satisfying (4) is less than

$$\sum{}'(c/A(a_j^+))(nx^{10}/\log n) \tag{6}$$

(the prime indicates that each $A(a_j^+)$ is counted only once); on the other hand the number of a_j^+'s is equal to or greater than $n/2x^4$ as stated in the lines preceding (4). Hence

$$\sum{}'\frac{c}{A(a_j^+)}\,\frac{nx^{10}}{\log n} > \frac{1}{2}\,\frac{n}{x^4}. \tag{7}$$

Let now N tend to infinity. As above we have that the number of integers $m \leqq N$ for which $A(m)=t$ equals

$$(1+o(1))\,\frac{N}{t}\prod_p{}'\left(1-\frac{1}{p}\right) > \frac{cN}{t}\,\frac{x^{10}}{\log n},$$

with a new value for the constant c. Thus we obtain from (7) that the number of integers $m \leqq N$ with $A(m)=A(a_j^+), j=1, 2, \cdots$, is greater than cN/x^4. But for these integers we have by (4)

$$f_k(m) = f(A(m)) \geqq x^{1/2} \qquad (k = \pi(n^{1/x^{10}})). \tag{8}$$

Thus for all sufficiently large N

$$M_k(x_i^{1/2}, N) > cN/x_i^4 \qquad (x_i = x) \tag{9}$$

where c is an absolute constant and x_i and k are independent of N. From (9) we obtain by a simple calculation that

$$\sum_{m=1}^{N} (f_k(m))^{10} > cx_iN \tag{10}$$

and x_i can assume arbitrarily large values. Now we shall show that (10) is false; in fact we prove the following lemma.

[5] V. Brun, *Le crible d'Eratosthene et le théorème de Goldbach*, Skrifter udgivne af Videnskabs selskabet Kristiania, Si Matematisk-Naturvidenskapelig Klasse vol. 3 (1920).

LEMMA 4. *Put* $h_t(p) = (f(p))^t - 1$. We have

$$\sum_{m=1}^{N} f_k(m)^t = (1 + o(1))N \prod_{p \leq p_k} \left(1 + \frac{h_t(p)}{p}\right) < cN, \tag{11}$$

where $c = c(t)$ *is independent of* k.

We have

$$\sum_{m=1}^{N} f_k(m)^t = \sum_{m=1}^{N} \prod_{p \mid m, p \leq p_k} (1 + h_t(p)) = \sum_d{}' \left[\frac{N}{d}\right] \prod_{p \mid d} h_t(p)$$
$$= N \sum_d{}' \frac{\prod_{p \mid d} h_t(p)}{d} + O(1) = (1 + o(1))N \prod_{p \leq p_k} \left(1 + \frac{h_t(p)}{p}\right),$$

where the dash indicates that d is squarefree and that all prime factors of d are not greater than p_k, and the error term $O(1)$ depends on k but not on N (the number of terms in $\sum'$ is bounded, and the bound depends on k but not on N). The second inequality of (11) follows easily from (1). This proves Lemma 4, and since (11) contradicts (10), Lemma 3 is also proved.

Now we can prove Theorem 1. For $t = 1$ we obtain from (11)

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} f_k(m) = \prod_{p \leq p_k} \left(1 + \frac{f(p) - 1}{p}\right). \tag{12}$$

Further from (1) we have

$$\lim_{k \to \infty} \prod_{p \leq p_k} \left(1 + \frac{f(p) - 1}{p}\right) = \prod_{p=2}^{\infty} \left(1 + \frac{f(p) - 1}{p}\right). \tag{13}$$

Thus to prove (2) (that is, Theorem 1) it will suffice to show that for every ϵ there exists a k_0 so that for $k > k_0$ and all sufficiently large n

$$\frac{1}{n} \left| \sum_{m=1}^{n} f(m) - f_k(m) \right| < \epsilon. \tag{14}$$

Write

$$\sum_{m=1}^{n} (f(m) - f_k(m)) = \sum\nolimits_1 + \sum\nolimits_2, \tag{15}$$

where in $\sum_1$ the summation is extended over the $m \leq n$ for which $f(m) \leq t$ and $f_k(m) \leq t$ (t fixed), and in $\sum_2$ the summation is extended over the remaining $m \leq n$.

Consider $\sum_2$. Divide the integers m appropriate to $\sum_2$ into classes

m_i $(i=t, t+1, \cdots)$ such that $i \leqq f(m_i) < i+1$. Then applying Lemma 3, we get

$$(16)\qquad \begin{aligned} \sum\nolimits_2 &\leqq \sum_{f(m)\geqq t} f(m) + \sum_{f_k(m)\geqq t} f_k(m) = \sum_{i=t}^{\infty} f(m_i) + \sum_{k=t}^{\infty} f_k(m_i) \\ &< 2cn \sum_{i=t}^{\infty} \frac{i+1}{i^3} < \frac{\epsilon}{2} n, \end{aligned}$$

for t sufficiently large. We now fix t_0 so that for all $t > t_0$, the last inequality of (16) holds.

Now we estimate $\sum_1$. Let

$$\sum\nolimits_1 = \sum\nolimits_1' + \sum\nolimits_1'',$$

where in $\sum_1'$ we impose the condition $|f(m)-f_k(m)| > \epsilon/4$, and in $\sum_1''$ we require $|f(m)-f_k(m)| \leqq \epsilon/4$. Then in $\sum_1''$ there are at most n summands each numerically less than $\epsilon/4$, so this sum is numerically less than $\epsilon n/4$. In $\sum_1'$ each summand exceeds $\epsilon/4$ in absolute value but is less than t by our definition of $\sum_1$. Moreover by taking k sufficiently large we can insure that the number of summands in $\sum_1'$ will be less than $\epsilon n/4t$. For it is well known that both $f_k(m)$ and $f(m)$ have asymptotic distribution functions and that the asymptotic distribution function of $f_k(m)$ tends to that of $f(m)$ as $k \to \infty$.[6] But this implies that $f_k(n)$ tends to $f(n)$ in relative measure, that is, that the number of integers not exceeding n for which $|f(n)-f_k(n)| > \epsilon$ tends to zero with $1/k$. Hence, we have the estimate

$$\left|\sum\nolimits_1\right| = \left|\sum\nolimits_1'\right| + \left|\sum\nolimits_1''\right| < \frac{\epsilon n}{4t}\cdot t + \frac{\epsilon n}{4} < \frac{\epsilon n}{2},$$

and

$$(17)\qquad \left|\sum_{m=1}^{n} (f(m) - f_k(m))\right| = \left|\sum\nolimits_1\right| + \left|\sum\nolimits_2\right| < \epsilon n.$$

This completes the proof.

It seems possible that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} f(m)$$

exists if we only assume that $\sum f(p)/p$ converges (some assumption like $|f(p)| < c$ might also be necessary). At present I am unable to decide this question. The present proof breaks down since Lemma 4

[6] Ibid. footnote 3.

is false if $\sum f(p)/p$ converges and $\sum (f(p))^2/p$ diverges.

Ramanujan[7] conjectured that for any $\alpha>0$, $\beta>0$

$$\sum_{\nu=1}^{n} \sigma_\alpha(\nu)\sigma_\beta(n-\nu)$$
$$= (1+o(1))\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}\frac{\zeta(\alpha+1)\zeta(\beta+1)}{\zeta(\alpha+\beta+2)}\sigma_{\alpha+\beta+1}(n),$$

where $\sigma_\alpha(n)$ denotes the sum of the αth power of the divisors of n.

Ingham[8] proved this conjecture; he also found analogous asymptotic formulas for $\sum_{\nu=1}^{n}\sigma_\alpha(\nu)\sigma_\alpha(\nu+k)$, $\sum_{\nu=1}^{n}\phi(\nu)\phi(\nu+k)$, $\sum_{\nu=1}^{n}d(\nu)d(n-\nu)$, $\sum_{\nu=1}^{n}d(\nu)d(\nu+k)$.

We are going to generalize these results. First we prove the following theorem.

THEOREM 2. *Assume that* $f^{(1)}(m)$ *and* $f^{(2)}(m)$ *satisfy* (1), *also* $f^{(i)}(p^\alpha)=f^{(i)}(p)$, $i=1, 2$. *Put* $f^{(i)}(p)-1=g^{(i)}(p)$, $i=1, 2$. *Then*

$$\sum_{\nu=1}^{n} f^{(1)}(\nu)f^{(2)}(n-\nu) = (1+o(1))An,$$

$$\sum_{\nu=1}^{x} f^{(1)}(\nu)f^{(2)}(\nu+n) = (1+o(1))Ax,$$

where

$$A = \prod_{p\nmid n}\left(1+\frac{g^{(1)}(p)+g^{(2)}(p)}{p}\right)$$
$$\cdot\prod_{p\mid n}\left(1+\frac{g^{(1)}(p)+g^{(2)}(p)+g^{(1)}(p)g^{(2)}(p)}{p}\right).$$

REMARKS. (1) It clearly follows from (1) that the product for A converges. (2) The assumption $f(p^\alpha)=f(p)$ can be omitted without any difficulty but the expression for A then becomes much more complicated.

We have by a simple calculation ($g_k(p)=f_k(p)-1$)

$$\sum_{\nu=1}^{n} f_k^{(1)}(\nu)f_k^{(2)}(n-\nu) = \sum_{\nu=1}^{n}\left(\prod_{p\mid\nu}(1+g_k^{(1)}(p))\prod_{p\mid n-\nu}1+g_k^{(2)}(n-\nu)\right)$$
$$= \sum_{d_1,d_2\geq 1}\left[\frac{n}{d_1d_2}\right]'\prod_{p\mid d_1}g_k^{(1)}(p)\prod_{p\mid d_2}g_k^{(2)}(p),$$

[7] *Collected papers*, p. 137.

[8] J. London Math. Soc. vol. 2 (1927) pp. 202–208.

where $[n/d_1d_2]$ denotes the number of solutions of $n=d_1x+d_2y$ in positive integers x and y. Clearly $|n/d_1d_2-[n/d_1d_2]'|\leqq 1$. Thus a simple calculation shows that

$$\sum_{\nu=1}^{n} f_k^{(1)}(\nu)f_k^{(2)}(n-\nu) = (1+o(1))n \prod_{p\nmid n,\, p\leqq p_k}\left(\frac{1+g^{(1)}(p)+g^{(2)}(p)}{p}\right) \prod_{p\mid n,\, p\leqq p_k}\left(1+\frac{g^{(1)}(p)+g^{(2)}(p)+g^{(1)}(p)g^{(2)}(p)}{p}\right).$$

Since the product for A converges our proof will be complete if we show that for sufficiently large k

$$\left|\sum_{\nu=1}^{n}\{f^{(1)}(\nu)f^{(2)}(n-\nu)-f_k^{(1)}(\nu)f_k^{(2)}(n-\nu)\}\right| < \epsilon n. \tag{18}$$

We suppress the proof of (18) since it is almost identical with that of (14); this completes the proof of the first half of Theorem 2. The proof of the second half is similar.

By the same method we can prove

$$\begin{aligned} \sum F^{(1)}(\nu)F_k^{(2)}(n-\nu) &= (1+o(1))A\,\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)}\,n^{\alpha+\beta+1}, \\ \sum_{\nu=1}^{x} F^{(1)}(\nu)F^{(2)}(\nu+n) &= (1+o(1))\,\frac{A}{\alpha+\beta+1}\,x^{\alpha+\beta+1}; \end{aligned} \tag{19}$$

where $F^{(1)}(n)=n^{\alpha}f^{(1)}(n)$, $F^{(2)}(n)=n^{\beta}f^{(2)}(n)$, $\alpha>0$, $\beta>0$ are arbitrary real numbers. (19) contains all the results of Ingham except those on $d(n)$ (we of course have to drop the assumption $f(p^{\alpha})=f(p)$).

Carlitz proved the following theorem:[9] Let

$$f(p) = 1+o(p^{-1/2+\epsilon}).$$

Then

$$\sum{}' f(n_1)f(n_2)\cdots f(n_v) = Cn^{v-1}+O(n^{v-9/8+\epsilon}); \tag{20}$$

the prime means that the summation is extended over all partitions of n into v summands. The value of C is given by a complicated expression.

By the method we used in proving Theorem 2 we can prove the following theorem.

THEOREM 3. *Let $f^{i}(n)$ satisfy* (1), $i=1, 2, \cdots, \nu$. *Then*

---

[9] Quart. J. Math. Oxford Ser. vol. 2 (1931) pp. 97–106; see also vol. 3 (1932) pp. 273–290.

$$\sum{}' f^{(1)}(n_1) f^{(2)}(n_2) \cdots f^{(\nu)}(n_\nu) = (1 + o(1)) D n^{\nu - 1},$$

also

$$\sum_{m=1}^{n} f^{(1)}(m + k_1) f^{(2)}(m + k_2) \cdots f^{(\nu)}(m + k_\nu) = (1 + o(1)) E n,$$

D and E are given by a complicated expression.

STANFORD UNIVERSITY