Lithuanian Mathematical Journal, *Vol.* 46, *No.* 3, 2006

ON THE LOCAL DISTRIBUTION OF CERTAIN ARITHMETIC FUNCTIONS

J. M. De Koninck[1]

Département de Mathématiques et de Statistique, Université Laval, Québec, Québec G1K 7P4, Canada
(e-mail: jmdk@mat.ulaval.ca)

I. Kátai[2]

Computer Algebra Department, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter Sétány I/C, Hungary
(e-mail: katai@compalg.inf.elte.hu)

Abstract. Let $d(n)$, $\sigma_1(n)$, and $\phi(n)$ stand for the number of positive divisors of n, the sum of the positive divisors of n, and Euler's function, respectively. For each $\nu \in \mathbf{Z}$, we obtain asymptotic formulas for the number of integers $n \leqslant x$ for which $e_n := \frac{\phi(n)}{d(n)^2} = 2^\nu m$ for some odd integer m as well as for the number of integers $n \leqslant x$ for which $e_n = 2^\nu r$ for some odd rational number r. Our method also applies when $\phi(n)$ is replaced by $\sigma_1(n)$, thus, improving upon an earlier result of Bateman, Erdős, Pomerance, and Straus, according to which the set of integers n such that $\frac{\sigma_1(n)}{d(n)^2}$ is an integer is of density $\frac{1}{2}$.

Keywords: arithmetic functions, distribution function.

1. INTRODUCTION

Let $d(n)$, $\sigma_1(n)$, and $\phi(n)$ stand for the number of positive divisors of n, the sum of the positive divisors of n, and Euler's function, respectively. Bateman, Erdős, Pomerance, and Straus [1] have shown that the set of integers n such that $\frac{\sigma_1(n)}{d(n)^2}$ is an integer is of density $\frac{1}{2}$.

In this paper, for each integer ν, we obtain asymptotic formulas for the number of integers $n \leqslant x$ for which $e_n := \frac{\phi(n)}{d(n)^2} = 2^\nu m$ for some odd integer m as well as for the number of integers $n \leqslant x$ for which $e_n = 2^\nu r$ for some odd rational number r. Our method can also be used to obtain similar estimates when $\phi(n)$ is replaced by $\sigma_1(n)$, from which the result of [1] follows. Moreover, in the process of establishing these results, we also investigate the local distributions of certain arithmetic functions.

As a starting point, we mention the following result of Wijsmuller [6]: *For each prime number q, let the completely additive function $\beta = \beta_q$ be defined on the primes p by $\beta(p) = r$, where r is the unique integer such*

[1]Research supported in part by a grant from NSERC.
[2]Research supported by the Applied Number Theory Research Group of the Hungarian Academy of Science and by a grant from OTKA.

Published in Lietuvos Matematikos Rinkinys, Vol. 46, No. 3, pp. 315–331, July–September, 2006. Original article submitted February 12, 2006.

0363–1672/06/4603–0257

that $q^r \| p+1$, *and set*

$$d_1 = \frac{q}{(q-1)^2} \quad \textit{and} \quad d_2 = \frac{q(q+1)}{(q-1)^3}. \tag{1}$$

Then

$$\lim_{x\to\infty} \frac{1}{x} \#\big\{n \leqslant x:\ \beta(n) - d_1 \log\log x \leqslant z\sqrt{d_2 \log\log x}\big\} = \Phi(z),$$

where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-t^2/2}\,\mathrm{d}t.$$

While this result establishes the global central limit distribution of the function β_q, we shall examine its local distribution as well as that of similar functions. In order to do this, we shall use the ideas developed in our earlier paper [2].

2. THE PROBABILISTIC SET UP

As is customary, let φ be the density function of the Gaussian law, precisely,

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-y^2/2}. \tag{2}$$

We start this section with an important result of Esseen [3].

LEMMA 1 (Esseen). *Let* $X_1, X_2, \ldots$ *be independent identically distributed random variables of lattice type such that* $MX_j = 0$ *and* $M|X_j|^\xi < +\infty$ *for some* $\xi \geqslant 3$. *Assume that the values of the* X_j*'s belong to the set* $\{\nu - \mu:\ \nu \in \mathbf{Z}\}$ *for some fixed real number* μ *and that the relation* $P(X_j = s-\mu)\cdot P(X_j = s+1-\mu) \neq 0$ *holds for at least one* s. *Then*

$$P(X_1 + X_2 + \cdots + X_n = k - n\mu) = \frac{1}{\sigma\sqrt{n}}\varphi(z_{n,k}) + \mathrm{O}\Big(\frac{1}{n}\Big),$$

where φ *is defined in* (2) *and where*

$$z_{n,k} = \frac{k - n\mu}{\sigma\sqrt{n}}, \qquad \sigma = MX_j^2.$$

Let q be a fixed prime, and let $\xi_1, \xi_2, \ldots$ be a sequence of identically distributed independent random variables with

$$P(\xi = s) = \rho(s), \quad \text{where } \rho(0) = \frac{q-2}{q-1}, \quad \rho(s) = \frac{1}{q^s} \text{ for each integer } s \geqslant 1.$$

Letting d_1 and d_2 be the constants defined in (1) and setting $X_j = \xi_j - d_1$, we easily see that $MX_j = 0$ and $\sigma^2 = MX_j^2 = d_2$. Now, letting

$$\eta_h = \xi_1 + \cdots + \xi_h, \tag{3}$$

from Lemma 1 it follows that

$$\begin{aligned} P(\eta_h = T) &= P(X_1 + \cdots + X_h + hd_1 = T) = P(X_1 + \cdots + X_h = T - hd_1) \\ &= \frac{1}{\sigma\sqrt{h}}\varphi\Big(\frac{T - hd_1}{\sigma\sqrt{h}}\Big) + \mathrm{O}\Big(\frac{1}{h}\Big). \end{aligned} \tag{4}$$

The notation $u \approx v$ means that $\frac{1}{2} \leqslant \frac{u}{v} \leqslant 2$.

3. NOTATION AND PRELIMINARY OBSERVATIONS

As usual, let $\mathbf{N}_0$, $\mathbf{N}$, and $\mathbf{Z}$ stand for the sets of nonnegative integers, positive integers, and all integers, respectively.

We use the standard notation $x_1 = \log x$, $x_i = \log x_{i-1}$ for $i = 2, 3, \ldots$.

Throughout this paper, p always denotes a prime number, while n and m stand for positive integers. On the other hand, $q \geqslant 2$ stands for a fixed prime. The number c stands for a positive constant, not necessarily the same at each occurrence. Moreover, $\omega(n)$ denotes the number of distinct prime factors of n, while $P(n)$ and $p(n)$ stand for the largest and smallest prime factors of n, respectively. We also let

$$\pi_k(x) := \#\{n \leqslant x \colon \ \omega(n) = k\}.$$

Let $\wp$ stand for the set of all prime numbers. Then, for each integer $r \geqslant 0$, let

$$\wp_r = \{p \in \wp \colon \ q^r \| p - 1\},$$

so that $\wp = \bigcup_{r=0}^{\infty} \wp_r$. Note that, in particular, the prime q itself belongs to $\wp_0$. Given an interval $I \subseteq [0, +\infty[$ and an integer $r \geqslant 0$, we let

$$\pi(I|\wp_r) = \#\{p \in I \cap \wp_r\}.$$

Let $f = f_q$ be the completely additive function defined implicitly by $f(p) = r$ if $p \in \wp_r$. Observe that Wijsmuller's result mentioned in Section 1 also holds when β_q is replaced by f_q with the same constants d_1 and d_2.

Setting

$$e_n := \frac{\phi(n)}{d(n)^2} \quad (n = 1, 2, \ldots),$$

for each integer ν, we let $\mathcal{D}_\nu = \mathcal{B}_\nu \setminus \mathcal{B}_\nu^*$, where

$$\mathcal{B}_\nu = \{n \in \mathbf{N} \colon \ e_n = 2^\nu m_1/m_2, \ \text{where} \ \ m_1 \text{ and } m_2 \text{ are odd positive integers}\},$$
$$\mathcal{B}_\nu^* = \{n \in \mathbf{N} \colon \ e_n = 2^\nu m, \ \text{where } m \text{ is an odd positive integer}\}.$$

In general, given a set C, we denote by $C(x)$ the cardinality of the set $\{n \leqslant x \colon \ n \in C\}$.

Moreover, for each integer $n \geqslant 2$ and each number $y > 2$, we let

$$n_y = \prod_{\substack{p^\alpha \| n \\ p < y}} p^\alpha. \tag{5}$$

We now introduce some notation which is somewhat similar to that we used in [2].

Let x be a fixed large number. Then letting A be a large constant and c_0 a positive constant, we introduce the set $\mathcal{L} = \{\ell_j \colon \ j = 0, 1, 2, \ldots\}$, where

$$\ell_0 = \exp\{x_2^A\}, \qquad \ell_{j+1} = \ell_j + \frac{\ell_j}{(\log \ell_j)^{c_0}} \quad \text{for } j = 0, 1, 2, \ldots.$$

For each positive integer ν, define the interval $I_\nu := [u_\nu, u_\nu + \Delta u_\nu]$, where $u_\nu = \ell_{j_\nu}$ and $\Delta u_\nu = \ell_{j_\nu+1} - \ell_{j_\nu}$, and set $\chi(u_\nu) := \mathrm{li}(u_\nu + \Delta u_\nu) - \mathrm{li}(u_\nu)$, where $\mathrm{li}(x) := \int_2^x \frac{\mathrm{d}t}{\log t}$. From the Prime Number Theorem for arithmetic progressions it follows that, for some fixed constant $c_2 > 0$,

$$\pi(I_\nu|\wp_r) = \rho(r)\chi(u_\nu)\Big(1 + \mathrm{O}\Big(\exp\big\{-c_2\sqrt{\log u_\nu}\big\}\Big)\Big), \tag{6}$$

an estimate valid for all $q^r \leqslant (\log \ell_0)^{c_1}$ with an arbitrarily large fixed number $c_1 > 0$, where

$$\rho(0) = \frac{q-2}{q-1} \qquad \rho(j) = \frac{1}{q^j} \quad \text{for } j = 1, 2, \ldots. \tag{7}$$

Let $x^{1/2} \leqslant Y \leqslant x$. An h-tuple $(u_1, \ldots, u_h)$ is said to be *feasible* if it satisfies both relations

$$\ell_0 \leqslant u_1 < \cdots < u_h \quad \text{and} \quad u_1 \cdots u_h \leqslant Y.$$

Now, consider a feasible h-tuple $(u_1, \ldots, u_h)$ such that $u_{\nu+1} \geqslant 2u_\nu$ for $\nu = 1, 2, \ldots, h-1$ and

$$(u_1 + \Delta u_1) \cdots (u_h + \Delta u_h) \leqslant Y,$$

and then let

$$E_h(u_1, \ldots, u_h) := \sum_{\substack{p_1 p_2 \cdots p_h \\ p_\nu \in [u_\nu, u_\nu + \Delta u_\nu]}} 1 = \prod_{\nu=1}^{h} \sum_{p_\nu \in [u_\nu, u_\nu + \Delta u_\nu]} 1 \tag{8}$$

and

$$S(u_1, \ldots, u_h) := \prod_{\nu=1}^{h} \left(1 + \mathrm{e}^{-c_2\sqrt{\log u_\nu}}\right),$$

where $c_2 > 0$ is a fixed constant. Then we have

$$\frac{1}{S(u_1, \ldots, u_h)} \leqslant \frac{E_h(u_1, \ldots, u_h)}{\prod_{\nu=1}^{h} \chi(u_\nu)} \leqslant S(u_1, \ldots, u_h). \tag{9}$$

But since

$$\begin{aligned} \log S(u_1, \ldots, u_h) &= \sum_{\nu=1}^{h} \log\Big(1 + \exp\big\{-c_2\sqrt{\log u_\nu}\big\}\Big) \\ &\ll \exp\Big\{-\frac{c_2}{2}\sqrt{\log u_1}\Big\} \\ &\ll \exp\Big\{-\frac{c_2}{2}\sqrt{\frac{1}{2}\log \ell_0}\Big\} \\ &\ll \exp\big\{-c_3 x_2^{A/2}\big\} \end{aligned} \tag{10}$$

for some constant $c_3 > 0$, from (9) and (10) it follows that

$$E_h(u_1, u_2, \ldots, u_h) = \prod_{\nu=1}^{h} \chi(u_\nu) \cdot \Big(1 + \mathrm{O}\big(\exp\big\{-c_3 x_2^{A/2}\big\}\big)\Big).$$

An expression of the form $j_1 j_2 \cdots j_t$, where t is a positive integer and each j_i is a nonnegative integer, is called a *word* of length t.

Given a word $\alpha = j_1 j_2 \cdots j_t$, we let

$$\rho(\alpha) := \rho(j_1)\rho(j_2)\cdots\rho(j_t),$$

where each $\rho(j_i)$ is defined by (7).

Let us define the function H on primes and on products of primes as follows: if $p \in \wp$, then $H(p) = f_q(p)$, and if $p_1 < p_2 < \cdots < p_t$ is a sequence of primes, then $H(p_1 p_2 \cdots p_t) = H(p_1)H(p_2)\cdots H(p_t)$.

4. MAIN RESULTS

Given $k \in \mathbf{N}$ and $s \in \mathbf{N}_0$, let

$$\mathcal{E}_{k,s} := \{n \in \mathbf{N}\colon\ \omega(n) = k,\ \ f_q(n) = s\}$$

and set

$$\mathcal{E}_{k,s}(x) := \#\{n \leqslant x\colon\ n \in \mathcal{E}_{k,s}\}.$$

THEOREM 1. *Let* $0 < \delta < \frac{1}{2}$ *be an arbitrary constant. Then, as* $x \to \infty$,

$$\frac{\mathcal{E}_{k,s}(x)}{\pi_k(x)} = P(\eta_k = s) + \mathrm{O}\Big(\frac{x_3}{x_2}\Big),$$

uniformly as

$$|k - x_2| < x_2^{\frac{1}{2}+\delta} \quad \text{and} \quad |s - d_1 x_2| < x_2^{\frac{1}{2}+\delta}, \tag{11}$$

where d_1 *is given in* (1).

THEOREM 2. *For each* $s \in \mathbf{N}_0$, *as* $x \to \infty$,

$$\frac{1}{x}\#\{n \leqslant x\colon\ f_q(n) = s\} = \frac{1}{\sqrt{d_2 x_2}}\varphi\Big(\frac{s - d_1 x_2}{\sqrt{d_2 x_2}}\Big) + \mathrm{O}\Big(\frac{x_3}{x_2}\Big),$$

where d_1 *and* d_2 *are given in* (1).

THEOREM 3. *For each* $\nu \in \mathbf{Z}$, *as* $x \to \infty$,

$$\frac{1}{x}\mathcal{B}_\nu(x) = \sqrt{\frac{2}{3x_2}}\,\varphi\Big(\frac{2\nu}{\sqrt{3x_2}}\Big) + \mathrm{O}\Big(\frac{x_3}{x_2}\Big).$$

THEOREM 4. *For each* $\nu \in \mathbf{Z}$, *as* $x \to \infty$,

$$\frac{1}{x}\mathcal{B}^*_\nu(x) = \sqrt{\frac{2}{3x_2}}\,\varphi\Big(\frac{2\nu}{\sqrt{3x_2}}\Big) + \mathrm{O}\Big(\frac{x_3}{x_2}\Big).$$

Note that the constants implied in the error terms appearing in Theorems 1, 2, 3, and 4 are absolute. On the other hand, the main terms have preponderance only if s (in Theorems 1 and 2) and ν (in Theorems 3 and 4) vary in some intervals.

Clearly, from Theorem 4 it follows, in particular, that the set of positive integers n such that $\frac{\phi(n)}{d(n)^2}$ is an odd integer times a (possibly, negative) power of 2 is of density 1. Moreover, summing up the estimate of Theorem 4 for $\nu = 0, 1, 2 \ldots$, we obtain that $\frac{\phi(n)}{d(n)^2}$ is an integer for about half of the positive integers, meaning that the density of the set of positive integers n for which $\frac{\phi(n)}{d(n)^2}$ is an integer is equal to $1/2$. Moreover, from our method it will become clear that $\phi(n)$ can be replaced by $\sigma_1(n)$. This observation implies the result of [1] mentioned in Section 1.

Remarks. Theorems similar to Theorems 1 and 2 can be proved for a more general class of additive functions by using the techniques developed in this paper combined with those of our earlier paper [2]. More precisely, let the set of primes $\wp$ be subdivided into finite or infinite disjoint sets $\wp_k$. Assume that

$$\begin{aligned}\pi\big([x, x+y]|\wp_k\big) &:= \#\{p \in \wp_k\colon\ p \in [x, x+y]\}\\ &= \rho(k)\big(\mathrm{li}(x+y) - \mathrm{li}(x)\big)\Big(1 + \mathrm{O}\big(\exp\{-c_2\sqrt{\log x}\}\big)\Big),\end{aligned}$$

provided that $\frac{x}{\log^c x} \leqslant y \leqslant x$ for some constant $c > 0$, where all $\rho(k) \geqslant 0$ and $\sum_{k\geqslant 0} \rho(k) = 1$. Now let g be an integer-valued additive function such that $g(p) = h(k)$ for $p \in \wp_k$. Let also ξ_i $(i = 1, 2, \ldots)$ be independent random variables taking the values $h(k)$ with probability $\rho(k)$ so that $P(\xi_i = h(k)) = \rho(k)$, and let $\eta_h = \xi_1 + \cdots + \xi_h$. Assume that $M\xi_i = d_1$, $X_j = \xi_j - d_1$, and $MX_j^2 = d_2$ exist. Assume, furthermore, that the greatest common divisor of $\{h(k)\colon\ \rho(k) > 0\}$ is equal to 1. Then, under these conditions, we can prove analogues of Theorems 1 and 2 with $g(n)$ instead of $f_q(n)$.

5. PRELIMINARY LEMMAS

LEMMA 2. *There exists an absolute constant $c > 0$ such that, given any integer $D \geqslant 3$, for all $x \geqslant 3$,*

$$\sum_{\substack{p\leqslant x \\ p\equiv 1 \pmod D}} \frac{1}{p} < \frac{c}{\phi(D)} \log\log x.$$

Proof. See Lemma 3 of Kátai [4].

LEMMA 3. *Let $G(x)$ be the number of integers $n \leqslant x$ having two prime divisors p_1 and p_2 satisfying $\ell_0 < p_1 < p_2 < 4p_1$. Then*

$$G(x) \ll \frac{x}{\log \ell_0}.$$

Proof. We have

$$G(x) \leqslant \sum_{\ell_0 < p_1 < p_2 \leqslant 4p_1} \frac{x}{p_1 p_2} \leqslant x \sum_{\ell_0 < p_1 \leqslant \sqrt{x}} \frac{1}{p_1} \sum_{p_1 < p_2 \leqslant 4p_1} \frac{1}{p_2}$$
$$\ll x \sum_{\ell_0 < p_1 < \sqrt{x}} \frac{1}{p_1 \log p_1} \ll x \sum_{n > \ell_0} \frac{1}{n \log^2 n} \ll \frac{x}{\log \ell_0},$$

which proves Lemma 3.

LEMMA 4. *Let $b > 0$ be a constant, and let $R = R(x) > bx_3$. Then*

$$\#\{n \leqslant x\colon\ \text{there exists } p|n \text{ such that } q^R|p-1\} \ll \frac{x\,x_2}{q^R}.$$

Proof. Let $S(x)$ be the quantity which is to be estimated. Using Lemma 2, we then have

$$S(x) = \sum_{p-1\equiv 0 \pmod{q^R}} \left[\frac{x}{p}\right] \leqslant x \sum_{p-1\equiv 0 \pmod{q^R}} \frac{1}{p} \ll x \frac{x_2}{\phi(q^R)},$$

which proves Lemma 4.

LEMMA 5. *Let $A_h := \sum_{(u_1,\ldots,u_h)} E_h(u_1, \ldots, u_h)$, where the sum runs over those feasible k-tuples for which $\prod_{\nu=1}^h u_\nu < Y < \prod_{\nu=1}^h (u_\nu + \Delta u_\nu)$. Let d be an arbitrary positive constant. Then*

$$\sum_{h=1}^{[dx_2]} A_h \ll Y \cdot x_2^{-Ac_0+1} + Y \cdot x_2^{-A}.$$

Proof. In view of Lemma 3, it is clear that it suffices to sum over those $(u_1, \ldots, u_h)$ for which $u_{j+1} \geqslant 2u_j$ $(j = 1, \ldots, h-1)$, $(u_0 \geqslant \ell_0)$. If $m \in E_h(u_1, \ldots, u_h)$, $u_1 \cdots u_h \leqslant Y \leqslant \prod_{\nu=1}^{h}(u_\nu + \Delta u_\nu)$, then

$$m \in \mathcal{J} := [Y - Y_1, Y], \tag{12}$$

where $Y_1 \leqslant Y x_2^{-Ac_0+1}$. Note that (12) holds, since

$$\prod_{\nu=1}^{h} u_\nu = \prod_{\nu=1}^{h}(u_\nu + \Delta u_\nu) \prod_{\nu=1}^{h} \frac{1}{1 + \frac{\Delta u_\nu}{u_\nu}} > Y \exp\Big\{-\frac{1}{2}\sum_{\nu=1}^{h} \frac{\Delta u_\nu}{u_\nu}\Big\}$$

and

$$\sum_{\nu=1}^{h} \frac{\Delta u_\nu}{u_\nu} \leqslant \sum_{\nu=0}^{h-1} \frac{1}{(\log 2^\nu \ell_0)^{c_0}} \leqslant \frac{d x_2}{x_2^{Ac_0}}.$$

Hence, the proof of Lemma 5 is complete.

6. PROOF OF THEOREM 1

We first classify the integers $n \in \mathcal{E}_{k,s}$ according to the value of n_{ℓ_0} (recall definition (5)), that is, for each integer K such that $P(K) \leqslant \ell_0$, we let

$$\mathcal{E}_{k,s}^{(K)} := \{n \in \mathcal{E}_{k,s},\ n_{\ell_0} = K\}.$$

Note that from here on, K always denotes an integer whose largest prime factor does not exceed ℓ_0.

Using the well-known estimate

$$\Psi(x, y) := \#\big\{n \leqslant x:\ P(n) \leqslant y\big\} \ll x \exp\Big\{-\frac{\log x}{2\log y}\Big\}$$

(see, for instance, Tenenbaum [5]), one can easily show that

$$\sum_{K > \exp\{x_2^{A+1}\}} \mathcal{E}_{k,s}^{(K)}(x) \ll \frac{x}{x_2^{2A}}. \tag{13}$$

Hence, from (13) it follows that

$$\mathcal{E}_{k,s}(x) = \sum_{K \leqslant \exp\{x_2^{A+1}\}} \mathcal{E}_{k,s}^{(K)}(x) + \mathrm{O}\Big(\frac{x}{x_2^{2A}}\Big). \tag{14}$$

Now, clearly, given any fixed $b > 0$, we have

$$\sum_{\substack{K \leqslant \exp\{x_2^{A+1}\} \\ \omega(K) > bx_3}} \mathcal{E}_{k,s}^{(K)}(x) \ll x \sum_{\substack{K \leqslant \exp\{x_2^{A+1}\} \\ \omega(K) > bx_3}} \frac{1}{K}. \tag{15}$$

In order to estimate this last sum, note that, for each real number $1 < u \leqslant \frac{1}{2}\exp\{x_2^{A+1}\}$,

$$\frac{1}{u} \sum_{\substack{u < K \leqslant 2u \\ K \leqslant \exp\{x_2^{A+1}\} \\ \omega(K) > bx_3}} 1 \ll \frac{2^{-bx_3}}{u} \sum_{\substack{u < K \leqslant 2u \\ K \leqslant \exp\{x_2^{A+1}\}}} d(K) = \frac{2^{-bx_3}}{u} \cdot S_0, \tag{16}$$

say. By Mertens' theorem,

$$S_0 \ll \sum_{\substack{u<K\leqslant 2u \\ K\leqslant \exp\{x_2^{A+1}\}}} \sum_{\substack{K_1|K \\ K_1<\sqrt{2u}}} 1 \ll \sum_{K_1<\sqrt{2u}} \frac{2u}{K_1} \ll u \prod_{p\leqslant \ell_0}\Big(1+\frac{1}{p}\Big) \ll u \log \ell_0. \tag{17}$$

Substituting (17) into (16), we get that

$$\frac{1}{u} \sum_{\substack{u<K\leqslant 2u \\ K\leqslant \exp\{x_2^{A+1}\} \\ \omega(K)>bx_3}} 1 \ll 2^{-bx_3} \log \ell_0 = 2^{-bx_3} x_2^A. \tag{18}$$

Thus, by choosing $u=2^\nu$ with $\nu=1,2,\dots$ such that $2^\nu \leqslant \frac{1}{2}\exp\{x_2^{A+1}\}$, from (18) and (15) it follows that

$$\sum_{\substack{K\leqslant \exp\{x_2^{A+1}\} \\ \omega(K)>bx_3}} \mathcal{E}_{k,s}^{(K)}(x) \ll x\, x_2^{-b\log 2} x_2^A x_2^{A+1} \ll x\, x_2^{-2A}, \tag{19}$$

provided that b is chosen large enough. Hence, substituting (19) into (14), we get that

$$\mathcal{E}_{k,s}(x) = \sum_{\substack{K\leqslant \exp\{x_2^{A+1}\} \\ \omega(K)\leqslant bx_3}} \mathcal{E}_{k,s}^{(K)}(x) + \mathrm{O}\Big(\frac{x}{x_2^{2A}}\Big). \tag{20}$$

Given a constant $c_4>0$, we now investigate the sum

$$S := \sum_{\substack{K\leqslant \exp\{x_2^{A+1}\} \\ f_q(K)>c_4x_3}} \frac{1}{K}.$$

First, let $Q\in(1,q)$ be a fixed number. Then

$$\begin{aligned} S &\leqslant Q^{-c_4x_3} \sum_{K\leqslant \exp\{x_2^{A+1}\}} \frac{Q^{f_q(K)}}{K} \\ &\leqslant Q^{-c_4x_3} \prod_{p<\ell_0}\Big(1+\frac{Q^{f_q(p)}}{p}+\frac{Q^{2f_q(p)}}{p^2}+\cdots\Big) \\ &= Q^{-c_4x_3} U, \end{aligned} \tag{21}$$

say. Now, using Lemma 2, we have

$$\begin{aligned} \log U &\ll \sum_{p<\ell_0} \frac{Q^{f_q(p)}}{p} = \sum_{r=0}^{\infty} Q^r \sum_{\substack{p<\ell_0 \\ p\equiv 1 \pmod{q^r}}} \frac{1}{p} \\ &\ll \sum_{r=0}^{\infty} \frac{Q^r}{\phi(q^r)} \log\log \ell_0 \ll \frac{Ax_3}{1-Q/q}. \end{aligned} \tag{22}$$

Substituting (23) into (22), we get that

$$S \ll Q^{-c_4 x_3} \exp\left\{\frac{A}{1 - Q/q} x_3\right\} \ll \frac{1}{x_2^{2A}}, \tag{23}$$

provided that we choose $Q = \sqrt{q}$ and c_4 sufficiently large with respect to A.

Now, let $\mathcal{T}$ be the set of integers K satisfying the following three conditions:

$$K \leqslant \exp\left\{x_2^{A+1}\right\}, \qquad \omega(K) \leqslant b x_3, \qquad f_q(K) \leqslant c_4 x_3. \tag{24}$$

From (14), (20), and (23) it follows that

$$\mathcal{E}_{k,s}(x) = \mathcal{E}'_{k,s}(x) + \mathrm{O}\left(\frac{x}{x_2^{2A}}\right), \tag{25}$$

where

$$\mathcal{E}'_{k,s}(x) := \sum_{K \in \mathcal{T}} \mathcal{E}_{k,s}^{(K)}(x).$$

So, let $K \in \mathcal{T}$ and set $Y = x/K$. We claim, that in the estimation of $\mathcal{E}_{k,s}(x)$, we may drop the integers $n \leqslant x$ such that $n = Km$ with $p(m) > \ell_0$ which satisfy any of the following three conditions:

(a) m is nonsquarefree (since those integers n with a corresponding m such that m is divisible by a square > 1 only introduce an error term of order at most x/ℓ_0);

(b) n has two "close" prime divisors p_1 and p_2 in the sense that $\ell_0 < p_1 < p_2 < 4p_1$ (since, according to Lemma 3, this only introduces an error term of order at most $x/\log \ell_0$);

(c) n is such that $\max_{p|n} f_q(p) > c_4 x_3$ (in view of (23)).

Now, for each positive integer K, let U_K be the set of integers $m \leqslant x/K$ which remain after having deleted those integers $n = Km$ which satisfy at least one of conditions (a), (b), or (c).

Let $m \in U_K$, so that, since $(K, m) = 1$, for each $n \in \mathcal{E}_{k,s}$, we have

$$\begin{aligned} \omega(m) &= \omega(n) - \omega(K) = k - \omega(K) := h, \\ f_q(m) &= f_q(n) - f_q(K) = s - f_q(K) := t. \end{aligned} \tag{26}$$

We shall now estimate $\mathcal{E}_{k,s}^{(K)}(x)$ for k and s satisfying conditions (11) and for $K \in \mathcal{T}$. Recall also that we only need to count those integers $n = Km$ for which conditions (a), (b), and (c) do not hold.

The function $E_h(u_1, \ldots, u_h)$ having been defined in (8), we further define

$$E_h(u_1, \ldots, u_h | \alpha) := \sum_{\substack{p_1 \cdots p_h \\ p_\nu \in [u_\nu, u_\nu + \Delta u_\nu] \\ H(p_1 \cdots p_h) = \alpha}} 1.$$

We now have to introduce three more definitions, precisely:

$$\begin{aligned} &\Gamma_{h,t} = \left\{\alpha = i_1 i_2 \cdots i_h :\ i_1 + i_2 + \cdots + i_h = t\right\}, \\ &N_h(Y|\ell_0, \alpha) = \#\left\{p_1 \cdots p_h < Y :\ \ell_0 < p_1 < \cdots < p_h,\ H(p_1 \cdots p_h) = \alpha\right\}, \\ &N_h(Y|\ell_0) = \#\left\{p_1 \cdots p_h < Y :\ \ell_0 < p_1 < \cdots < p_h\right\}. \end{aligned}$$

Now, using the Prime Number Theorem for arithmetic progressions, as we did in (6), we obtain that, at least in the case $\max_{\nu=1,\dots,h} H(p_\nu) \leqslant cx_3$,

$$E_h(u_1,\dots,u_h|\alpha) = \rho(\alpha)\prod_{\nu=1}^{h}\chi(u_\nu)\Big(1+\mathrm{O}\big(\exp\{-c_2\sqrt{\log u_\nu}\}\big)\Big). \tag{27}$$

Repeating the argument appearing in (10), at least in the case $u_{\nu+1} \geqslant 2u_\nu$, from (27) it follows that the estimate

$$E_h(u_1,\dots,u_h|\alpha) = \rho(\alpha)E_h(u_1,\dots,u_h)\Big(1+\mathrm{O}\big(\exp\{-c_3x_2^{A/2}\}\big)\Big) \tag{28}$$

holds for feasible h-tuples $(u_1,\dots,u_h)$ and, therefore, that

$$\begin{aligned}\sum_{\alpha\in\Gamma_{h,t}} & E_h(u_1,\dots,u_h|\alpha)\\ &= \Big(1+\mathrm{O}\big(\exp\{-c_3x_2^{A/2}\}\big)\Big)E_h(u_1,\dots,u_h)\sum_{\alpha\in\Gamma_{h,t}}\rho(\alpha) + \text{ Error}_1 \qquad (29)\\ &= \Big(1+\mathrm{O}\big(\exp\{-c_3x_2^{A/2}\}\big)\Big)E_h(u_1,\dots,u_h)P(\eta_h=t) + \text{ Error}_1,\end{aligned}$$

where Error_1 comes from those words α with $H(p_1\cdots p_h)=\alpha$ such that $\max_{\nu=1,\dots,h} H(p_\nu) > c_4x_3$. But the total contribution of such integers $m \leqslant Y$ with $H(m)=\alpha$ does not exceed $\frac{Yx_2}{q^{c_4x_3}}$, as shown in Lemma 4. Therefore, assuming that c_4 is sufficiently large, we can conclude that

$$\text{Error}_1 \ll \frac{Y}{x_2^{2A}}. \tag{30}$$

On the other hand, it is clear that

$$N_h(Y|\ell_0) = \sum_{(u_1,\dots,u_h)}{}^{*}\ E_h(u_1,\dots,u_h) + \text{ Error}_2 \tag{31}$$

and

$$N_h(Y|\ell_0,\alpha) = \sum_{(u_1,\dots,u_h)}{}^{*}\ E_h(u_1,\dots,u_h|\alpha) + \text{ Error}_3(\alpha), \tag{32}$$

where the star on each of the above sums indicates that the sum runs over all feasible numbers $(u_1,\dots,u_h)$ which also satisfy the conditions

$$(u_1+\Delta u_1)\cdots(u_h+\Delta u_h) < Y \quad\text{ and }\quad u_{\nu+1}\geqslant 2u_\nu \quad (\nu=1,\dots,h-1).$$

From Lemma 3 we know that ignoring the integers n having two close prime divisors p_1 and p_2 with $\ell_0 < p_1 < p_2 < 4p_1$ only generates an error $\ll \frac{x}{\log \ell_0} = \frac{x}{x_2^A}$ in each of the sums appearing in (31) and (32) and, thus, an error term which is no larger than that claimed in the statement of Theorem 1.

Similarly, according to Lemma 5, when estimating the sums in (31) and (32), we can ignore the integers $m = p_1\cdots p_h$ with $p_\nu \in [u_\nu, u_\nu+\Delta u_\nu]$ and $u_1\cdots u_h < Y < (u_1+\Delta u_1,\dots,u_h+\Delta u_h)$, since the error generated by counting them is $\ll \frac{Y}{x_2^{Ac_0}}$.

Furthermore, we clearly have that

$$\sum_{\alpha\in\Gamma_{h,t}}\text{Error}_3(\alpha) \ll \text{Error}_2. \tag{33}$$

Thus, taking into account these error terms, from (29) and (32) we get that

$$\sum_{\alpha\in\Gamma_{h,t}} N_h(Y|\ell_0,\alpha) = P(\eta_h = t)\cdot N_h(Y|\ell_0) + \mathrm{O}\Big(\frac{Y}{x_2^{2A}}\Big). \tag{34}$$

Consequently, gathering relations from (29) to (34), we get that

$$\begin{aligned}\mathcal{E}'_{k,s}(x) &= \sum_{K\in\mathcal{T}} \#U_K = \sum_{K\in\mathcal{T}}\sum_{\alpha\in\Gamma_{h,t}} N_h\Big(\frac{x}{K}|\ell_0,\alpha\Big)\\ &= \Big(1+\mathrm{O}\big(\exp\{-c_3x_2^{A/2}\}\big)\Big)\sum_{K\in\mathcal{T}} P(\eta_h=t)N_h\Big(\frac{x}{K}|\ell_0\Big)+\mathrm{O}\Bigg(\sum_{K\in\mathcal{T}}\frac{x}{Kx_2^{2A}}\Bigg)\\ &= \Big(1+\mathrm{O}\big(\exp\{-c_3x_2^{A/2}\}\big)\Big)\sum_{K\in\mathcal{T}} P(\eta_h=t)N_h\Big(\frac{x}{K}|\ell_0\Big)+\mathrm{O}\Big(\frac{x}{x_2^{A-1}}\Big).\end{aligned} \tag{35}$$

Now, using (4) and the fact that $\varphi'(z)$ is bounded on the set of real numbers, we have

$$\big|P(\eta_h=t)-P(\eta_h=s)\big| \ll \frac{|t-s|}{x_2} = \mathrm{O}\Big(\frac{x_3}{x_2}\Big). \tag{36}$$

Using (36) in (36), we obtain

$$\begin{aligned}\mathcal{E}'_{k,s}(x) = \Big(1+\mathrm{O}\big(\exp\big\{-c_3x_2^{A/2}\big\}\big)\Big)\sum_{K\in\mathcal{T}} P(\eta_h=s)N_h\Big(\frac{x}{K}|\ell_0\Big)\\ +\mathrm{O}\Bigg(\frac{x_3}{x_2}\sum_{K\in\mathcal{T}} N_h\Big(\frac{x}{K}|\ell_0\Big)\Bigg).\end{aligned} \tag{37}$$

Recalling (26), we have

$$\sum_{K\in\mathcal{T}} N_h\Big(\frac{x}{K}|\ell_0\Big) \leqslant \#\big\{mK\leqslant x:\ \ p(m)>\ell_0,\ \ \omega(mK)=k\big\} = \pi_k(x),$$

which means that the error term in (38) can be replaced by $\mathrm{O}(\frac{x_3}{x_2}\pi_k(x))$.

On the other hand, in view of (4) and since $h=k+\mathrm{O}(x_3)$, we have

$$\begin{aligned}\big|P(\eta_h=s)-P(\eta_k=s)\big| &\ll \frac{1}{h}+\Bigg|\frac{1}{\sigma\sqrt{h}}\varphi\Big(\frac{s-hd_1}{\sigma\sqrt{h}}\Big)+\frac{1}{\sigma\sqrt{k}}\varphi\Big(\frac{s-kd_1}{\sigma\sqrt{k}}\Big)\Bigg|+\mathrm{O}\Big(\frac{1}{x_2}\Big)\\ &\ll \frac{1}{h}+\Big|\frac{1}{\sqrt{h}}-\frac{1}{\sqrt{k}}\Big|+\frac{1}{\sqrt{h}}|\sqrt{h}-\sqrt{k}|\\ &\ll \frac{1}{h}+\frac{|h-k|}{h},\end{aligned} \tag{38}$$

so that, for any arbitrary fixed constant $c>0$, we have

$$\max_{\substack{h\\ |h-k|<cx_3}} \big|P(\eta_h=s)-P(\eta_k=s)\big| \ll \frac{x_3}{x_2}.$$

Therefore, replacing $P(\eta_h=s)$ by $P(\eta_k=s)$ in (38) only introduces an additional error which is

$$\ll \frac{x_3}{x_2}\sum_{K\in\mathcal{T}} N_h\Big(\frac{x}{K}|\ell_0\Big) \leqslant \frac{x_3}{x_2}\pi_k(x).$$

In view of these last two remarks, we are entitled to replace (38) by

$$\mathcal{E}'_{k,s}(x) = \Big(1 + \mathrm{O}\big(\exp\{-c_3 x_2^{A/2}\}\big)\Big) P(\eta_k = s) \sum_{K\in\mathcal{T}} N_h\Big(\frac{x}{K}|\ell_0\Big) + \mathrm{O}\Big(\frac{x_3}{x_2}\pi_k(x)\Big). \tag{39}$$

Now, noting that

$$\begin{aligned}\pi_k(x) &= \#\big\{n = Km \leqslant x:\ P(K) < \ell_0,\ p(m) > \ell_0,\ \omega(n) = k\big\}\\ &= \sum_{P(K)<\ell_0} \#\big\{m \leqslant x/K:\ p(m) > \ell_0,\ \omega(n) = k - \omega(K)\big\}\\ &= \sum_{K\in\mathcal{T}} \#\big\{m \leqslant x/K:\ m \text{ squarefree}, p(m) > \ell_0,\ \omega(m) = k - \omega(K)\big\} + \mathrm{O}\Big(\frac{x}{x_2^{2A}}\Big),\end{aligned}$$

we get that

$$\sum_{K\in\mathcal{T}} N_h\Big(\frac{x}{K}|\ell_0\Big) = \pi_k(x) + \mathrm{O}\Big(\frac{x}{x_2^{2A}}\Big). \tag{40}$$

Using (40) in (39), we obtain that

$$\mathcal{E}'_{k,s}(x) = \Big(1 + \mathrm{O}\big(\exp\{-c_3 x_2^{A/2}\}\big)\Big) P(\eta_k = s)\bigg(\pi_k(x) + \mathrm{O}\Big(\frac{x}{x_2^{2A}}\Big)\bigg) + \mathrm{O}\Big(\frac{x_3}{x_2}\pi_k(x)\Big). \tag{41}$$

Substituting (41) into (25), Theorem 1 follows.

7. PROOF OF THEOREM 2

As we shall see, Theorem 2 is an easy consequence of Theorem 1.

It is clear that

$$\#\big\{n \leqslant x:\ f_q(n) = s\big\} = \sum_{k=1}^{\infty} \mathcal{E}_{k,s}(x) = \Sigma_1 + \Sigma_2 + \Sigma_3,$$

where, in Σ_1, we sum over $k < x_2 - \kappa$, in Σ_3, over $k > x_2 + \kappa$, and, in Σ_2, over k for which $|k - x_2| \leqslant \kappa$, where we have set $\kappa = \sqrt{x_2}\cdot x_3^2$.

Clearly,

$$\begin{aligned}\Sigma_1 &\leqslant \#\big\{n \leqslant x:\ \omega(n) \leqslant x_2 - \kappa\big\}\\ &\ll \frac{x}{x_1}\frac{x_2^{x_2-1}}{(x_2-1)!}\frac{(x_2-1)\cdots(x_2-\kappa)}{x_2^{\kappa}}\\ &\ll \frac{x}{\sqrt{x_2}}\prod_{j=1}^{\kappa}\Big(1 - \frac{j}{x_2}\Big) \ll \frac{x}{\sqrt{x_2}}\exp\Big\{-\frac{1}{x_2}\sum_{j=1}^{\kappa} j\Big\}\\ &\ll \frac{x}{\sqrt{x_2}}\exp\Big\{-\frac{\kappa^2}{2x_2}\Big\} \ll \frac{x}{x_2^{c}}\end{aligned} \tag{42}$$

for each number $c > 0$. Similarly, we have

$$\Sigma_3 \ll \frac{x}{x_2^{c}}. \tag{43}$$

Hence, in view of (43) and (43), we only need to estimate Σ_2. We shall do this by using Theorem 1 and (39). From (36) it follows that

$$\left|P(\eta_k = s) - P(\eta_{[x_2]} = s)\right| \ll \frac{|k - [x_2]|}{x_2}.$$

Hence,

$$\Sigma_2 = P(\eta_{[x_2]} = s) \sum_{|k-x_2|<\kappa} \pi_k(x) + \mathrm{O}\left(\sum_{|k-x_2|<\kappa} \frac{|k-[x_2]|}{x_2} \pi_k(x) \right) + \mathrm{O}\left(\sum_k \frac{x_3}{x_2} \pi_k(x) \right),$$

which implies that

$$\Sigma_2 = x P(\eta_{[x_2]} = s) + \mathrm{O}\left(x \frac{x_3}{x_2} \right),$$

whence Theorem 2 follows after applying (4).

8. PROOF OF THEOREM 3

We only give an outline of the proof, since it follows by using the same method as in Theorem 1.

First, we subdivide the set of primes $\wp$ into classes

$$Q_{-2},\ Q_{-1},\ Q_0,\ Q_1,\ Q_2,\ Q_3,\ \ldots,$$

in the following way. We first let $Q_{-2} = \{2\}$. Then, for each integer $r \geqslant 1$, we let $p \in Q_{r-2}$ for $2^r \| p-1$. Then, let g be the completely additive function defined implicitly by $g(p) = s$ if $p \in Q_s$, so that $g(p) \in \{-2, -1, 0, 1, 2, 3, \ldots\}$.

If m is a squarefree number, then e_m is clearly of the form "odd rational $\times 2^{g(m)}$."

Repeating the argument used in the proof of Theorem 1, one can obtain an asymptotic formula with a remainder term for the number of integers $m = p_1 \cdots p_h < Y$ with $\ell_0 < p_1 < \cdots < p_h$ and $p_\nu \in Q_{i_\nu}$ for $\nu = 1, \ldots, h$ and for every choice of the word $\alpha = i_1 \cdots i_h$ under the constraints $|i_\nu| \leqslant c_4 x_3$. Proceeding in this way, we derive the asymptotic formula of Theorem 3, noting on the way that, in this case, $\sigma = \sqrt{3/2}$.

9. PROOF OF THEOREM 4

Clearly, it suffices to prove that

$$\mathcal{D}(x) = \sum_{\nu=-\infty}^{\infty} \mathcal{D}_\nu(x) = \mathrm{O}\left(\frac{x}{x_2^2} \right). \tag{44}$$

Now, let us write $n = Km$, where K is the squarefull part of n, and m is the squarefree part of n with $(K, m) = 1$, and, in this case, we have $e_n = e_K \cdot e_m$.

First note that it is clear that the number of positive integers $n \leqslant x$ such that the corresponding value K satisfies $K > x_2^2$ is $\mathrm{O}(\frac{x}{x_2^2})$.

We shall now obtain an upper bound for the number of positive integers $n \leqslant x$ such the corresponding value K satisfies $K \leqslant x_2^2$ and for which $q^{-a} \| e_n$ holds for some odd prime q. Since $q^0 \| e_m$, it follows that $q^{-a} \| e_K$. Now assume that $q^b \| d(K)$. Then, given any $\varepsilon > 0$, if K is sufficiently large, we have $d(k) < K^\varepsilon \leqslant x_2^{4\varepsilon}$ and, in this case, we can write that $q^a \leqslant q^{2b} < x_2^{4\varepsilon}$.

Now, if q^{2b} does not divide $\phi(n)$, then n contains no prime divisor p satisfying $p \equiv 1 \pmod{q^{2b}}$. Noting that, by Lemma 2,

$$\prod_{\substack{p < x^\delta \\ p \equiv 1 \pmod{q^{2b}}}} \left(1 - \frac{1}{p}\right) = \exp\left\{ - \sum_{\substack{p < x^\delta \\ p \equiv 1 \pmod{q^{2b}}}} \frac{1}{p} \right\} \ll \exp\left\{ -\frac{1}{\phi(q^{2b})} x_2 + \mathrm{O}(1) \right\},$$

it follows by Selberg's sieve that the number of $n \leqslant x$ with this property is less than $cx\exp\{-\frac{x_2}{\phi(q^a)}\}$ for some positive constant c. Summing up over all $q^a \leqslant x_2^{4\varepsilon}$, (44) follows, and Theorem 4 is thus established.

REFERENCES

1. P. Bateman, P. Erdős, C. Pomerance, and E. G. Straus, The arithmetic mean of the divisors of an integer, in: *Analytic Number Theory* (Philadelphia, PA, 1980), Lecture Notes in Math., 889, Springer (1981), pp. 197–220.
2. J. M. De Koninck and I. Kátai, On the distribution of subsets of primes in the prime factorization of integers, *Acta Arith.*, **72**, 169–200 (1995).
3. C. G. Esseen, Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law, *Acta Math.*, **77**, 1–125 (1945).
4. I. Kátai, On the number of prime factors of $\varphi(\varphi(n))$, *Acta Math. Hungar.*, **58**(9–2), 211–225 (1991).
5. G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Stud. Adv. Math., Cambridge University Press (1995).
6. M. Wijsmuller, The value distribution of an additive function, *Ann. Univ. Sci. Budapest Sect. Comput.*, **14**, 279–291 (1994).