LOCAL DENSITIES OVER INTEGERS FREE OF LARGE PRIME FACTORS

By ALEKSANDAR IVIĆ *and* GÉRALD TENENBAUM

[Received 17th June 1985]

1. Introduction

LET $f(n)$ be a nonnegative, integer valued arithmetical function. For a fixed integer $k \geqslant 0$ the local density d_k of $f(n)$ is defined as

$$d_k = \lim_{x \to \infty} \frac{1}{x} \sum_{n \leqslant x, f(n)=k} 1 \tag{1.1}$$

whenever this limit exists. Many common multiplicative and additive functions possess local densities. As examples we mention $a(n)$, the number of nonisomorphic abelian groups with n elements, $\mu^2(n)$, the characteristic function of squarefree numbers, and $\Omega(n) - \omega(n)$, where $\Omega(n)$ and $\omega(n)$ denote respectively the number of prime factors of n counted with and without multiplicities.

The function $a(n)$ is multiplicative. It satisfies $a(p) = 1$ for every prime p and, in general, $a(p^\nu)$ equals the number of unrestricted partitions of ν. Kendall and Rankin **[13]** established the existence of local densities of $a(n)$ for any given $k \geqslant 1$. The first author **[10]**, **[11]** and Krätzel **[14]** determined more closely local densities of $a(n)$ and of related multiplicative functions, and investigated the error term in the asymptotic formula for the sum in (1.1).

In the case of $\mu^2(n)$, the only relevant value is $k = 1$ and the well-known question of the density of squarefree numbers can be also regarded as a special case of local density for the additive function $\Omega(n) - \omega(n)$, since $\mu^2(n) = 1$ is equivalent to $\Omega(n) - \omega(n) = 0$. Rényi **[17]** was the first to show that all local densities for $\Omega(n) - \omega(n)$ exist and may be computed from the identity

$$\sum_{k=0}^{\infty} d_k z^k = 6\pi^{-2} \prod_p \left(\frac{1 - z/(p+1)}{1 - z/p} \right), \qquad (|z| \leqslant 1), \tag{1.2}$$

where the product is extended to all primes p. In particular, (1.2) yields the familiar value $d_0 = 6\pi^{-2}$.

Every integer $n \geqslant 1$ can be written uniquely as $n = qs$ with $(q, s) = 1$, where $q = q(n)$ is squarefree and $s = s(n)$ is squarefull (meaning $p^2 \mid n$ whenever $p \mid n$). It is remarkable that $a(n)$, $\mu^2(n)$, and $\Omega(n) - \omega(n)$ all depend only on $s(n)$ and not on $q(n)$. Thus these functions may be appropriately called *arithmetical functions with squarefull kernel*, or simply *s-functions*. Henceforth by an s-function we shall mean a nonnegative, integer valued arithmetical function $f(n)$ such that $f(n) = f(s(n))$ for all $n \geqslant 1$. In the present context the interest of this definition lies in the following simple argument showing that any s-function possesses local densities of any given rank (cf. [5]).

Indeed, if f is an s-function we have

$$\sum_{n \leqslant x, f(n)=k} 1 = \sum_{s \leqslant x, f(s)=k} \sum_{q \leqslant x/s, (q,s)=1} 1 \tag{1.3}$$

where, here and in the sequel, the letters s and q denote respectively generic squarefull and squarefree integers. The existence of local densities follows from the fact that the inner sum in (1.3) can be satisfactorily estimated by a classical argument based on the Möbius inversion formula. Namely, uniformly for x, $r \geqslant 1$ we have

$$\sum_{q \leqslant x, (q,r)=1} 1 = 6\pi^{-2} A(r)x + O(B(r)x^{\frac{1}{2}} \log x) \tag{1.4}$$

with

$$A(r) = \prod_{p|r} (1+p^{-1})^{-1}, \qquad B(r) = \prod_{p|r} (1+p^{-\frac{1}{2}}). \tag{1.5}$$

Inserting (1.4) in (1.3) we obtain uniformly for $x \geqslant 1$, $k \geqslant 0$,

$$\sum_{n \leqslant x, f(n)=k} 1 = d_k x + O(x^{\frac{1}{2}} \log^2 x), \tag{1.6}$$

where d_k is given explicitly as

$$d_k = 6\pi^{-2} \sum_{s \geqslant 1, f(s)=k} A(s)s^{-1}. \tag{1.7}$$

In [5], Fainleib obtains a slightly better estimate for the remainder term in (1.6). Here, it simply stems from the elementary bound

$$\sum_{s \leqslant x} B(s)s^{-\frac{1}{2}} \leqslant \prod_{p \leqslant x} \left(1 + B(p) \sum_{\nu=2}^{\infty} p^{-\nu/2}\right) \ll \log x.$$

In this context, it may be worthwhile to keep in mind that an explicit dependence in k can still be obtained by re-inserting the dropped condition $f(s) = k$ in the summation.

We now prove (1.4). Let $\lambda(n; r)$ be the multiplicative function of n defined in the following way: $\lambda(1; r) = 1$ and

$$\lambda(p^\nu; r) = \begin{cases} -1, & \text{if} \quad \nu = 1, p \mid r \quad \text{or} \quad \nu = 2, p \nmid r, \\ 0, & \text{otherwise.} \end{cases}$$

Then we have the identity

$$\sum_{d|n} \lambda(d; r) = \begin{cases} \mu^2(n) & \text{if} \quad (n, r) = 1, \\ 0 & \text{if} \quad (n, r) > 1, \end{cases} \tag{1.8}$$

which can be easily established by noting that both sides are multiplicative functions of n, and we also have

$$\sum_{d=1}^{\infty} \lambda(d;r)d^{-1} = \prod_{p|r}(1-p^{-1})\prod_{p\nmid r}(1-p^{-2}) = 6\pi^{-2}A(r). \tag{1.9}$$

Thus by (1.8) and (1.9) the left-hand side of (1.4) equals

$$\sum_{n\leqslant x}\sum_{d|n}\lambda(d;r) = \sum_{d\leqslant x}\lambda(d;r)[x/d]$$

$$= 6\pi^{-2}A(r)x + O\Big(\sum_{d\leqslant x}|\lambda(d;r)| + x\sum_{d>x}|\lambda(d;r)|\,d^{-1}\Big).$$

Put $\sigma = 1/\log x$. Then the O-term above is clearly

$$\leqslant x^{\frac{1}{2}+\sigma}\sum_{d=1}^{\infty}|\lambda(d;r)|\,d^{-\frac{1}{2}-\sigma} = x^{\frac{1}{2}+\sigma}\prod_{p|r}(1+p^{-\frac{1}{2}-\sigma})\prod_{p\nmid r}(1+p^{-1-2\sigma})$$
$$\ll x^{\frac{1}{2}}B(r)\zeta(1+2\sigma) \ll x^{\frac{1}{2}}\log xB(r).$$

This completes the proof of (1.4).

2. Definitions and statement of results

As seen above, the question of local densities is essentially solved in the case of an s-function. A further insight into the structure of the sequence of those integers n such that $f(n)=k$ may be derived from the study of the joint distribution of $f(n)$ and some other arithmetic function $F(n)$, say, closely linked to the multiplicative nature of n. We choose $F(n)=P(n)$, the largest prime factor of n $(n\geqslant 2)$, $P(1)=1$.

We hence introduce the quantity

$$\Psi_k(x,y;f) = \sum_{n\leqslant x,\,P(n)\leqslant y,\,f(n)=k} 1 \tag{2.1}$$

generalizing the sum in (1.1). The desired new information will be mainly reflected in the comparison of (2.1) with

$$\Psi(x,y) = \sum_{n\leqslant x,\,P(n)\leqslant y} 1,$$

the number of positive integers not exceeding x and free of prime factors larger than y. This function is of constant occurrence in number theory and has been extensively studied by many authors, including Dickman [**4**], de Bruijn [**2**, **3**], Alladi [**1**], Hildebrand [**6**–**8**], and Hildebrand–Tenenbaum [**9**]. Estimates for $\Psi(x,y)$ play a crucial role in the proofs of our results. We discuss in Sections 3 and 5 respectively the known and new results we need.

Thus, our aim is to describe as precisely as possible the effect of varying y on the asymptotic behaviour of the ratio $\Psi_k(x, y; f)/\Psi(x, y)$. It plainly tends to d_k, as given by (1.7), if $y = x \to \infty$ and we want to determine, to within a relatively small factor, the optimal function $y_0(x)$ such that this still holds uniformly for $y_0(x) \leq y \leq x$. We obtain in Theorems 1, 3 below a satisfactory answer to this problem by establishing the existence of a general "critical point" materializing a change of behaviour of $\Psi_k(x, y; f)/\Psi(x, y)$.

Before we proceed to the formulation of our results, it is convenient to introduce some notation. For $\sigma \geq 0$ we define the multiplicative function

$$A(n; \sigma) = \prod_{p|n} (1 + p^{-\sigma})^{-1}, \tag{2.2}$$

and, given any s-function f, the generalized local density

$$d_k(\sigma; f) = \zeta^{-1}(2\sigma) \sum_{s \geq 1, f(s) = k} A(s; \sigma) s^{-\sigma}, \qquad (\sigma > \tfrac{1}{2}). \tag{2.3}$$

Noting that the Dirichlet series associated with the characteristic function of squarefull numbers is $\zeta(2s)\zeta(3s)/\zeta(6s)$, one obtains by a classical convolution argument

$$\sum_{s \leq x} 1 = \zeta(\tfrac{3}{2})\zeta^{-1}(3)x^{\frac{1}{2}} + O(x^{\frac{1}{3}}). \tag{2.4}$$

Hence by partial summation and (2.4) it is easily seen that the series in (2.3) converges for $\sigma > \frac{1}{2}$, and moreover we have $d_k = d_k(1; f)$. Similarly, for

$$B(n) = \prod_{p|n} (1 + p^{-\frac{1}{2}}) \tag{2.5}$$

we set

$$D_k(\sigma; f) = \sum_{s=1, f(s)=k}^{\infty} B(s) s^{-\sigma} \qquad (\sigma > \tfrac{1}{2}). \tag{2.6}$$

For $2 \leq y \leq x$ we shall systematically use the notation

$$u = \frac{\log x}{\log y},$$

and further we define $\xi = \xi(u)$ as the unique positive root of the equation $e^{\xi} = 1 + u\xi$ if $u > 1$, and set $\xi(1) = 0$. Thus $\xi = \log(1 + u\xi) = \log u + \log(\xi + 1/u)$, whence

$$\xi(u) = \log u + \log\log u + O(1). \tag{2.7}$$

Finally we define

$$\beta = \beta(x, y) = 1 - \frac{\xi(u)}{\log y}, \tag{2.8}$$

hence from (2.7) it follows that in the range

$$\log^{2+\varepsilon} x < y \leqslant x, \qquad x \geqslant x_0(\varepsilon) \tag{2.9}$$

we have

$$\beta \geqslant \frac{1}{2} + \frac{\varepsilon}{6} \tag{2.10}$$

for any fixed ε, $0 < \varepsilon < 1$.

We now formulate our main result.

THEOREM 1. *Let f be an s-function. For any fixed ε, $0 < \varepsilon < 1$, and x, y satisfying* (2.9) *we have, uniformly in $k \geqslant 0$,*

$$\begin{aligned}\Psi_k(x, y; f) = \Psi(x, y)\Big\{d_k(\beta; f) + O_\varepsilon\Big(\Big(d_k\Big(\beta - \frac{\varepsilon}{10}; f\Big) \\ + D_k(\beta; f)\Big)\frac{\log\log y}{\log y}\Big)\Big\}.\end{aligned} \tag{2.11}$$

When $f = \mu^2$, only $s = 1$ appears in $d_1(\beta; \mu^2) = \zeta^{-1}(2\beta)$. Therefore we obtain, as a corollary to Theorem 1, that

$$\Psi_1(x, y; \mu^2) = \zeta^{-1}(2\beta)\Psi(x, y)\Big(1 + O_\varepsilon\Big(\frac{\log\log y}{\log y}\Big)\Big) \tag{2.12}$$

holds uniformly in the range (2.9) for $0 < \varepsilon < 1$. The problem of the estimation of $\Psi_1(x, y; \mu^2)$ was considered by the first author in **[12]**, but the result proved there is superseded by (2.12). During the preparation of the manuscript, the authors have been informed by Daboussi that his student M. Naimi obtained independently the asymptotic formula contained in (2.12), for the same range, by a different method.

It was pointed out in **[12]** that the asymptotic formula

$$\Psi_1(x, y; \mu^2) = \sum_{n \leqslant x, P(n) \leqslant y} \mu(n)^2 \sim d_1\Psi(x, y) = 6\pi^{-2}\Psi(x, y) \tag{2.13}$$

cannot hold for $y_0 < y \leqslant \log x$ (by the prime number theorem it is easily seen that the left hand side does not depend on x if $y < (1 - \varepsilon)\log x$). As a complement to this and (2.12) we prove

THEOREM 2. *For any fixed ε, $0 < \varepsilon < 1$, and $y_1(\varepsilon) < y < \log^{2-\varepsilon} x$, we have*

$$\Psi_1(x, y; \mu^2) \ll \Psi(x, y)\exp\{-(\log y)^{\varepsilon/3}\}. \tag{2.14}$$

As will be clear from the proof, we made here no attempt to find the best possible factor on the right hand side of (2.14)—especially with respect to the dependence in the variable x. Combining (2.12) and (2.14) we see that we have an asymptotic evaluation of

$$\frac{1}{\Psi(x, y)} \sum_{n \leq x, P(n) \leq y} \mu^2(n)$$

for all y except when $y = \log^{2+o(1)} x$. This case can be also treated by our methods at the cost of some technical complications, but we shall not consider it here.

In view of (2.10) the series defining $d_k(\beta; f)$ is normally convergent when (2.9) is satisfied, hence $d_k(\beta; f)$ is in this region a continuous function of β, and

$$d_k(\beta; f) \sim d_k = d_k(1; f)$$

holds simultaneously for all s-functions f if and only if $\xi(u)/\log y \to 0$ as $x \to \infty$. By (2.7) this condition is equivalent to

$$\frac{\log y}{\log \log x} \to \infty, \qquad (x \to \infty). \tag{2.15}$$

Substituting in Theorem 1, we thus obtain a general answer to the question raised above concerning the change of behaviour of $\Psi_k(x, y; f)/\Psi(x, y)$. This is

THEOREM 3. *The growth condition* (2.15) *is necessary and sufficient for the asymptotic formula*

$$\Psi_k(x, y; f) \sim d_k \Psi(x, y) \tag{2.16}$$

to hold for every s-function and any fixed integer $k \geq 0$.

The problem of the estimation of $\Psi_k(x, y; f)$ when k is not fixed is, in general, very difficult. This is so even in the special case when $y = x$ and f is a common arithmetical function (not necessarily only an s-function). Thus it is only very recently that satisfactory estimates of $\Psi_k(x, x; \Omega)$ and $\Psi_k(x, x; \omega)$ were obtained by Nicolas **[15]** and Pomerance **[16]** respectively, when k is not assumed to be fixed.

It is a virtue of Theorem 1 that k in (2.11) does not have to be fixed. Naturally, the range for k in (2.11) is *a priori* determined by the maximal order of $f(n)$. For instance

$$\Omega(n) - \omega(n) \leq \frac{\log n}{\log 2} - 1 \qquad (n \geq 2), \tag{2.17}$$

and we have (see [18]) that for any fixed $\eta > 0$

$$a(n) \leq \exp\left\{(1+\eta)\frac{\log 5}{4} \cdot \frac{\log n}{\log \log n}\right\} \qquad (n \geq n_0(\eta)). \tag{2.18}$$

When k becomes large, the function $D_k(\beta;f)$ appearing in (2.11) may grow (as a function of k) much faster than $d_k(\beta;f)$, and (2.11) reduces to an upper bound estimate. We now investigate this in more detail in the special cases when $f(n)=\Omega(n)-\omega(n)$ or $f(n)=a(n)$.

Consider the case $f(n)=\Omega(n)-\omega(n)$ and suppose that (2.9) holds. For all δ, λ, such that $\delta>\frac{1}{2}$, $2^\delta<\lambda<3^\delta$, we have

$$D_k(\delta;\Omega-\omega)=\sum_{\substack{s=1\\ \Omega(s)-\omega(s)=k}}^{\infty} B(s)s^{-\delta}=\lambda^{-k}\sum_{\substack{s=1\\ \Omega(s)-\omega(s)=k}}^{\infty}\lambda^{\Omega(s)-\omega(s)}B(s)s^{-\delta}$$
$$\leqslant \lambda^{-k}\prod_p\Big(1+B(p)\sum_{j=2}^{k+1}\lambda^{j-1}p^{-j\delta}\Big)\ll_{\lambda,\delta}2^{-\delta k}.$$

Hence we infer

$$D_k(\beta;\Omega-\omega)+d_k(\beta-\varepsilon/10;\Omega-\omega)\ll_\varepsilon 2^{-(\beta-\varepsilon/10)k},$$

and thus obtain from Theorem 1

THEOREM 4. *Let ε, $0<\varepsilon<1$, be fixed. We have uniformly for x, y in the range* (2.9) *and* $0\leqslant k\leqslant(\log x/\log 2)-1$

$$\Psi_k(x,y;\Omega-\omega)\ll_\varepsilon 2^{-(\beta-\varepsilon/10)k}\Psi(x,y). \tag{2.19}$$

A similar estimate may be derived in the case $f(n)=a(n)$ by appealing to (2.18). Indeed, let s_1 denote the smallest possible s such that $a(s)=k$. Then (2.18) implies

$$k=a(s_1)\leqslant\exp\left\{(1+\eta)\frac{\log 5}{4}\left(\frac{\log s_1}{\log\log s_1}\right)\right\}$$

whence, for $k>k_0$

$$s_1>\exp\left\{(1-\eta)\frac{4}{\log 5}\log k\cdot\log\log k\right\}=K,$$

say. Therefore

$$D_k(\beta;a)+d_k\Big(\beta-\frac{\varepsilon}{10};a\Big)\ll_\varepsilon\sum_{s>K}B(s)s^{-\beta+\varepsilon/10}\ll_\varepsilon K^{\frac{1}{2}-\beta+(\varepsilon/9)}.$$

The same estimate is obviously valid for $d_k(\beta;a)$ since this is a decreasing function of β. Choosing $\eta=\eta(\varepsilon)$ sufficiently small we hence deduce from Theorem 1 the following

THEOREM 5. *Let ε, $0<\varepsilon<1$, be fixed. We have uniformly for x, y in the range* (2.9) *and* $1\leqslant k\leqslant\exp\left\{\frac{\log 5}{4}\cdot\frac{\log x}{\log\log x}\right\}$

$$\Psi_k(x,y;a)\ll_\varepsilon\exp\left\{-\Big(\beta-\frac{1}{2}-\frac{\varepsilon}{8}\Big)\frac{4}{\log 5}\log k\cdot\log\log k\right\}\Psi(x,y). \tag{2.20}$$

The authors wish to thank here the anonymous referee for pointing out an error in the original manuscript.

3. The relevant background on $\Psi(x, y)$

In this section, we summarize the main results on $\Psi(x, y)$ that will be needed in the proofs of Theorems 1 and 2.

The Dickman–de Bruijn function $\rho(v)$ for $v \geqslant 0$ is defined as the continuous solution of the delay differential equation $v\rho'(v) + \rho(v-1) = 0$, with the initial condition $\rho(v) = 1$ for $0 \leqslant v \leqslant 1$. Strengthening a result of de Bruijn [3], Hildebrand [6] recently showed that the asymptotic formula

$$\Psi(x, y) = x\rho(u)\left(1 + O_\varepsilon\left(\frac{\log(u+1)}{\log y}\right)\right), \qquad \left(u = \frac{\log x}{\log y}\right) \tag{3.1}$$

holds uniformly in the range

$$x \geqslant 2, \qquad \exp\left((\log\log x)^{\frac{5}{3}+\varepsilon}\right) \leqslant y \leqslant x \tag{3.2}$$

for any fixed $\varepsilon > 0$. The constant $\frac{5}{3}$ is the reciprocal of the constant which appears in the error term of the strongest known form of the prime number theorem. The connection between the range of validity of (3.1) and the zero-free region of the Riemann zeta-function was investigated by Hildebrand [7], who proved that (3.1) holds in the range (2.9) if and only if the Riemann hypothesis is true. Since it is known (see [8], [9]) that for $y \leqslant \log^{2-\varepsilon} x$ the behaviour of $\Psi(x, y)$ depends strongly on the irregularities in the distribution of primes, it appears that (2.9) is essentially the best possible range in which $\Psi(x, y)$ is approximable by a smooth function. Note that (2.9) is precisely the range in which Theorem 1 holds.

DeBruijn [2] established an asymptotic formula for $\rho(v)$. His method has recently been refined by Alladi [1] to improve the error term involved. He obtains

$$\rho(v) = \left(1 + O\left(\frac{1}{v}\right)\right)\left(\frac{\xi'(v)}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\gamma - v\xi(v) + \int_0^{\xi(v)} \frac{e^t - 1}{t}\, dt\right\}, \tag{3.3}$$

where $\gamma = 0.577 \cdots$ is Euler's constant and $\xi(v)$ is defined in Section 2. An alternative proof of (3.3) appears in [9].

Using (3.3) Alladi shows [1; Lemma 3] that

$$-\frac{d}{dv}\log\rho(v) = \log(v\xi(v)) + O\left(\frac{1}{v}\right) = \xi(v) + O\left(\frac{1}{v}\right), \tag{3.4}$$

where the second equality comes from the estimate

$$\log(v\xi(v)) = \log(e^{\xi(v)} - 1) = \xi(v) + O\Big(\frac{1}{v}\Big).$$

We remark that the derivative of the function of v in the exponential in (3.3) is exactly $(-\xi(v))$ and that $\xi''(v)/\xi'(v) \sim -1/v$. This leads immediately to (3.4) with the slightly weaker error term $O(\log v/v)$.

By Taylor's formula (3.4) implies

$$\log\Big(\frac{\rho(v-t)}{\rho(v)}\Big) = \int_{v-t}^{v} \xi(w)\,dw + O\Big(\log\Big(\frac{v}{v-t}\Big)\Big) \tag{3.5}$$

uniformly for $0 \leqslant t \leqslant v-1$.

The Dirichlet series associated with $\Psi(x, y)$ is

$$\zeta(\sigma; y) = \sum_{n=1, P(n)\leqslant y}^{\infty} n^{-\sigma} = \prod_{p\leqslant y} (1-p^{-\sigma})^{-1} \qquad (\sigma > 0).$$

Rankin's upper bound method, fully exploited by de Bruijn [**3**], is based on the inequality

$$\Psi(x, y) \leqslant x^{\sigma}\zeta(\sigma; y) \qquad (\sigma > 0). \tag{3.6}$$

The optimal choice for σ is $\sigma = \alpha(x, y)$, where $\alpha = \alpha(x, y)$ is the unique solution of the equation

$$\sum_{p\leqslant y} \frac{\log p}{p^{\alpha} - 1} = \log x. \tag{3.7}$$

In [9] Hildebrand and the second author evaluate the classical Perron integral for $\Psi(x, y)$ by the saddle point method and obtain that

$$\Psi(x, y) = \frac{x^{\alpha}\zeta(\alpha; y)}{\alpha\sqrt{2\pi\varphi_2(\alpha, y)}}\Big(1 + O\Big(\frac{1}{u} + \frac{\log y}{y}\Big)\Big) \tag{3.8}$$

holds uniformly in the range $x \geqslant y \geqslant 2$, where $\alpha = \alpha(x, y)$ and

$$\varphi_2(\alpha, y) = \sum_{p\leqslant y} \frac{p^{\alpha}\log^2 p}{(p^{\alpha} - 1)^2}.$$

It is shown [**9**; formula (7.8)] that the formula

$$\alpha(x, y) = \beta + O_{\varepsilon}(\exp(-2\sqrt{\log y}) + u^{-1}\log^{-2} y), \tag{3.9}$$

where $\beta = \beta(x, y)$ is defined by (2.8), is uniformly valid in the range

$$x \geqslant 2, \qquad \log^{1+\varepsilon} x \leqslant y \leqslant x \tag{3.10}$$

and that [9; formula (7.6)]

$$\alpha(x, y) = \frac{\log(1 + y/\log x)}{\log y}\Big(1 + O\Big(\frac{1}{\log y}\Big)\Big) \tag{3.11}$$

holds for $x \geqslant 2$, $y \leqslant \log^2 x$. Moreover we have in the range (3.10) [**9**; Theorem 2]

$$\frac{x^\alpha \zeta(\alpha; y)}{\alpha\sqrt{2\pi\varphi_2(\alpha, y)}} = x\left(\frac{\xi'(u)}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\gamma - u\xi(u) + \int_0^{\xi(u)} \frac{e^v - 1}{v}\, dv + O_\varepsilon(R(u, y))\right\} \quad (3.12)$$

with

$$R(u, y) = \frac{\log(u+1)}{\log y} + u \exp(-\sqrt{\log y}).$$

If we write the left-hand side of (3.12) as $\exp F(u, y)$, then it may be shown [**9**; formula (6.4)] that

$$\frac{\partial F}{\partial u}(u, y) = \alpha \log y + O\left(\frac{1}{u} + \frac{\log y}{y}\right) \quad (3.13)$$

holds uniformly for $x \geqslant y \geqslant 2$, with the consequence that

$$\Psi\left(\frac{x}{d}, y\right) = \Psi(x, y) d^{-\alpha}\left(1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right)\right) \quad (3.14)$$

whenever $\sqrt{x} \geqslant y \geqslant d \geqslant 1$.

4. Proof of Theorem 2

We may suppose without loss of generality that $y \geqslant y_1(\varepsilon)$. Then (3.11) easily implies that

$$\alpha(x, y) \leqslant \frac{1}{2} - \frac{\varepsilon}{4}. \quad (4.1)$$

Put $z = \frac{1}{4}\log y$, so that by the prime number theorem

$$\prod_{p \leqslant z} p \leqslant y^{\frac{1}{2}}. \quad (4.2)$$

We shall use the inequality

$$\mu^2(n) \leqslant \sum_{d^2 | n,\, P(d) \leqslant z} \mu(d), \qquad (n \geqslant 1), \quad (4.3)$$

which follows from the fact that the right-hand side of (4.3) is equal to $\mu^2(m)$, where m is the largest divisor of n such that $P(m) \leqslant z$. This yields

$$\Psi_1(x, y; \mu^2) \leqslant \sum_{n \leqslant x,\, P(n) \leqslant y} \sum_{d^2 | n,\, P(d) \leqslant z} \mu(d) = \sum_{d \leqslant x^{\frac{1}{2}},\, P(d) \leqslant z} \mu(d) \Psi(xd^{-2}, y).$$

By (4.2) the condition $d \leq x^{\frac{1}{2}}$ is redundant and we can use (3.14) to estimate $\Psi(xd^{-2}, y)$ in the sum above. Since $y_1(\varepsilon) \leq y \leq \log^{2-\varepsilon} x$ we have $u^{-1} \leq y^{-\frac{1}{2}}$, hence we obtain

$$\Psi_1(x, y, \mu^2) \leq \Psi(x, y) \sum_{P(d) \leq z} \mu(d) d^{-2\alpha}(1 + O(y^{-\frac{1}{2}}))$$

$$= \Psi(x, y)\left\{\prod_{p \leq z} (1 - p^{-2\alpha}) + O\left(y^{-\frac{1}{2}} \prod_{p \leq z} (1 + p^{-2\alpha})\right)\right\}. \tag{4.4}$$

But from the prime number theorem we have

$$\sum_{p \leq z} p^{-2\alpha} \sim \frac{z^{1-2\alpha}}{(1-2\alpha)\log z} \geq \log^{\varepsilon/3} y, \qquad (y \geq y_1(\varepsilon)),$$

hence the expression in curly brackets in (4.4) is easily seen to be $\ll \exp(-\log^{\varepsilon/3} y)$ and Theorem 2 follows.

5. New lemmas on the local behaviour of $\Psi(x, y)$

In this section we shall derive several consequences of the results presented in Section 3. We are interested in the behaviour of the ratio $\Psi(x/d, y)/\Psi(x, y)$ when $1 \leq d \leq x$ and x, y lie in the range

$$x \geq 2, \qquad \log^{1+\varepsilon} x \leq y \leq x, \tag{5.1}$$

where ε, $0 < \varepsilon < 1$, is fixed. The material of this section will be used in the proof of Theorem 1. In view of other applications, we actually prove more than is needed for this purpose.

LEMMA 1. *Put $t = \log d/\log y$. We have uniformly for x, y in the range (5.1) and $1 \leq d \leq x/y$*

$$\Psi\left(\frac{x}{d}, y\right) = \frac{1}{d}\Psi(x, y) \exp\left\{\int_{u-t}^{u} \xi(v)\, dv + O_\varepsilon(E(x, y, d))\right\}, \tag{5.2}$$

where

$$E(x, y, d) = \begin{cases} \dfrac{1}{u} + t\exp(-\sqrt{\log y}) + \log\dfrac{u}{u-t}, & \text{if} \quad u > \log y, \\[2ex] \dfrac{\log(u+1)}{\log y} + \log\dfrac{u}{u-t}, & \text{if} \quad u \leq \log y. \end{cases}$$

Proof. By (3.8), (3.9) and (3.13), we have for x, y satisfying (5.1)

$$\Psi\Big(\frac{x}{d}, y\Big) = \Psi(x, y) \exp\Big\{-\int_{u-t}^{u} (\log y - \xi(v)$$

$$+ O_\varepsilon(v^{-1} + \mathrm{e}^{-\sqrt{\log y}}))\,\mathrm{d}v + O\Big(\frac{1}{u-t}\Big)\Big\}$$

$$= \Psi(x, y) \exp\Big\{-\log d + \int_{u-t}^{u} \xi(v)\,\mathrm{d}v$$

$$+ O_\varepsilon\Big(\frac{1}{u-t} + t\mathrm{e}^{-\sqrt{\log y}} + \log\frac{u}{u-t}\Big)\Big\}.$$

This gives the result stated for the range $u > \log y$, since

$$\frac{1}{u-t} + \log\frac{u}{u-t} \ll \frac{1}{u} + \log\frac{u}{u-t} \qquad (0 \leqslant t \leqslant u-1).$$

For $u \leqslant \log y$ we have by (3.1) and (3.5)

$$\Psi\Big(\frac{x}{d}, y\Big) = \frac{1}{d}\Psi(x, y) \exp\Big\{\log\Big(\frac{\rho(u-t)}{\rho(u)}\Big) + O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\}$$

$$= \frac{1}{d}\Psi(x, y) \exp\Big\{\int_{u-t}^{u} \xi(v)\,\mathrm{d}v + O\Big(\log\frac{u}{u-t} + \frac{\log(u+1)}{\log y}\Big)\Big\},$$

and this is exactly the second estimate of the lemma.

LEMMA 2. *We have uniformly for x, y in the range* (5.1) *and* $1 \leqslant d \leqslant y$

$$\Psi\Big(\frac{x}{d}, y\Big) = \Psi(x, y) d^{-\beta}\Big(1 + O_\varepsilon\Big(\frac{\log d}{\log x} + \frac{\log\log y}{\log y}\Big)\Big), \tag{5.3}$$

where $\beta = \beta(x, y)$ is defined by (2.8).

Proof. When $d \leqslant x/y$, we apply Lemma 1 and observe that for $d \leqslant y$ we have $t \leqslant 1$ and

$$E(x, y, d) \ll \frac{\log d}{\log x} + \frac{\log\log y}{\log y},$$

$$\int_{u-t}^{u} \xi(v)\,\mathrm{d}v = t\xi(u) + O\Big(\frac{t}{u}\Big).$$

It remains to prove (5.3) when $1 \leqslant u \leqslant 2$ and $d \geqslant x/y$. In this case we have

$$\Psi\left(\frac{x}{d}, y\right) = \left[\frac{x}{d}\right] = \frac{x}{d}\left(1 + O\left(\frac{1+\log d}{\log x}\right)\right),$$

$$\Psi(x, y) = x\rho(u)\left(1 + O\left(\frac{1}{\log x}\right)\right) = x\left(1 + O\left(u - 1 + \frac{1}{\log x}\right)\right)$$
$$= x\left(1 + O\left(\frac{1+\log d}{\log x}\right)\right),$$

and

$$d^{\beta} = d \exp\left(-\frac{\log d}{\log y}\xi(u)\right) = d\left(1 + O\left(\frac{\log d}{\log x}\right)\right).$$

The required result follows trivially from these estimates.

LEMMA 3. *We have uniformly for x, y in the range* (5.1) *and* $1 \leqslant d \leqslant x$

$$\Psi\left(\frac{x}{d}, y\right) \ll \Psi(x, y) d^{-\beta + c/\log y}, \tag{5.4}$$

where $\beta = \beta(x, y)$ is defined by (2.8) *and c is an absolute constant.*

Proof. Suppose first that $d \leqslant x/y$. Then Lemma 1 is applicable and it may be easily checked that

$$E(x, y, d) \ll 1 + t.$$

This yields

$$\Psi\left(\frac{x}{d}, y\right) \leqslant \frac{1}{d}\Psi(x, y)\exp\{t\xi(u) + O(1+t)\} \ll \Psi(x, y) d^{-\beta + c_1/\log y},$$

where we used the fact that $\xi(u)$ is an increasing function of u.

If $x/y < d \leqslant x$, that is $u - 1 < t \leqslant u$, then $\Psi(x/d, y) = [x/d]$ and we only have to show that for a suitable $c_2 > 0$

$$\Psi(x, y) \gg x d^{\beta - 1 - c_2/\log y} = x \exp\{-t\xi(u) - c_2 t\}.$$

In view of (2.7) it is enough to prove

$$\Psi(x, y) \gg x \exp\{-u\xi(u) - c_3 u\},$$

but this easily follows from (3.8) and (3.12).

6. Proof of Theorem 1

The proof depends on the following auxiliary result which is an analogue of (1.4) for integers free of large prime factors. For any integer

$r \geqslant 1$, we define

$$\Psi_r(x, y) = \sum_{q \leqslant x, P(q) \leqslant y, (q,r)=1} 1 = \sum_{n \leqslant x, P(n) \leqslant y, (n,r)=1} \mu^2(n).$$

LEMMA 4. *We have uniformly for x, y satisfying* (2.9) *and $r \geqslant 1$*

$$\Psi_r(x, y) = \zeta^{-1}(2\beta)\Psi(x, y)\left\{A(r; \beta) + O_\varepsilon\left(B(r)\frac{\log\log y}{\log y}\right)\right\}. \quad (6.1)$$

Proof. We use (1.8) to express the characteristic function of the set of squarefree numbers prime to r and invert summations. This gives

$$\Psi_r(x, y) = \sum_{d \leqslant x, P(d) \leqslant y} \lambda(d; r)\Psi\left(\frac{x}{d}, y\right) = \sum\nolimits_1 + \sum\nolimits_2,$$

say, corresponding to the respective ranges of summation $1 \leqslant d \leqslant Y := \log^{6/\varepsilon} y$, and $Y < d \leqslant x$.

By Lemma 2, we have

$$\begin{aligned}\sum\nolimits_1 &= \Psi(x, y)\sum_{d \leqslant Y} \lambda(d; r)d^{-\beta}\left(1 + O_\varepsilon\left(\frac{\log\log y}{\log y}\right)\right)\\ &= \Psi(x, y)\left\{\zeta^{-1}(2\beta)A(r; \beta)\right.\\ &\quad \left. + O_\varepsilon\left(\sum_{d>Y} |\lambda(d; r)|\, d^{-\beta} + \frac{\log\log y}{\log y}\sum_{d \leqslant Y} |\lambda(d; r)|\, d^{-\beta}\right)\right\}, \quad (6.2)\end{aligned}$$

where we used the identity

$$\sum_{d=1}^{\infty} \lambda(d; r)d^{-\beta} = \prod_{p|r}(1 - p^{-\beta})\prod_{p \nmid r}(1 - p^{-2\beta}) = \zeta^{-1}(2\beta)A(r; \beta).$$

The remainder term in (6.2) is estimated by appealing to the lower bound (2.10) for β, and using the inequality

$$|\lambda(n; r)| \leqslant \sum_{m^2\delta = n, \delta|r} \mu^2(\delta)\mu^2(m),$$

which is easily verified by taking $n = p^\nu$, since both sides are multiplicative functions of n. We have

$$\begin{aligned}\sum_{d>Y} |\lambda(d; r)|\, d^{-\beta} &\leqslant \sum_{\delta|r} \mu^2(\delta)\delta^{-\beta}\sum_{m>(Y/\delta)^{\frac{1}{2}}} \mu^2(m)m^{-2\beta}\\ &\ll_\varepsilon \sum_{\delta|r} \mu^2(\delta)\delta^{-\beta}(Y/\delta)^{-(2\beta-1)/2} = Y^{\frac{1}{2}-\beta}B(r) \leqslant B(r)/\log y,\end{aligned}$$

and similarly

$$\sum_{d\leq Y}|\lambda(d;r)|\,d^{-\beta}\leq\sum_{\delta|r}\mu^2(\delta)\delta^{-\beta}\sum_{m=1}^{\infty}\mu^2(m)m^{-2\beta}\ll_\varepsilon B(r).$$

Thus so far we have proved that

$$\sum\nolimits_1=\zeta^{-1}(2\beta)\Psi(x,y)\left\{A(r;\beta)+O_\varepsilon\left(B(r)\frac{\log\log y}{\log y}\right)\right\},$$

and it is sufficient to show that $\sum_2$ can be absorbed by the error term in (6.1). Indeed, by Lemma 3 we have

$$\begin{aligned}\sum\nolimits_2&\ll\Psi(x,y)\sum_{d>Y}|\lambda(d;r)|\,d^{-\beta+c/\log y}\\&\ll_\varepsilon\Psi(x,y)Y^{-(\beta-\frac{1}{2})+c/\log y}B(r)\ll_\varepsilon\frac{B(r)\Psi(x,y)}{\log y},\end{aligned}$$

where the sum over d has been estimated similarly as the corresponding sum in $\sum_1$, using the fact that $\beta-c/\log y>\frac{1}{2}$ if x, and therefore y, is large enough. This completes the proof of Lemma 4.

Having at our disposal Lemma 4 it is a fairly simple matter to prove Theorem 1. Using the canonical decomposition $n=qs$, $(q,s)=1$, and noting that $f(n)=f(s)$ because f is an s-function, we obtain immediately

$$\Psi_k(x,y;f)=\sum_{s\leq x,P(s)\leq y,f(s)=k}\Psi_s\left(\frac{x}{s},y\right)=\sum\nolimits_1+\sum\nolimits_2,\tag{6.3}$$

say, where in $\sum_1$ we have $s\leq Y:=\log^{20/\varepsilon}y$, and, in $\sum_2$, $Y<s\leq x$.

By successive applications of Lemmas 4 and 2, we find that uniformly for $s\leq Y$

$$\psi_s\left(\frac{x}{s},y\right)=\zeta^{-1}(2\beta)\Psi(x,y)\{A(s;\beta)s^{-\beta}+O_\varepsilon(B(s)s^{-\beta}(\log\log y)/\log y)\},$$

whence

$$\begin{aligned}\sum\nolimits_1&=\Psi(x,y)\left\{d_k(\beta;f)+O_\varepsilon\left(D_k(\beta;f)\frac{\log\log y}{\log y}+\sum_{s>Y,f(s)=k}A(s;\beta)s^{-\beta}\right)\right\}\\&=\Psi(x,y)\left\{d_k(\beta;f)+O_\varepsilon\left(D_k(\beta;f)\frac{\log\log y}{\log y}+d_k\left(\beta-\frac{\varepsilon}{10};f\right)\frac{1}{\log y}\right)\right\}.\end{aligned}$$

Using Lemma 3 we further obtain

$$\sum\nolimits_2\leq\sum_{Y<s\leq x,f(s)=k}\Psi\left(\frac{x}{s},y\right)\ll\Psi(x,y)\sum_{s>Y,f(s)=k}s^{-\beta+c/\log y}.$$

We now remark that

$$A(s;\beta)^{-1} \ll \exp\sqrt{\log s} \leqslant s^{\varepsilon/30}$$

for $s > Y$ and y large enough. Hence

$$s^{-\beta+c/\log y} \leqslant s^{-\beta+\varepsilon/10} A(s;\beta) Y^{-\varepsilon/20}$$

for $s > Y$, and we deduce that

$$\sum\nolimits_2 = O_\varepsilon\left(\Psi(x,y) d_k\left(\beta - \frac{\varepsilon}{10}; f\right) \frac{1}{\log y}\right).$$

Inserting the expressions for $\sum_1$ and $\sum_2$ in (6.3) we obtain the assertion of Theorem 1.

REFERENCES

1. K. Alladi, 'The Turán-Kubilius inequality for integers without large prime factors', *J. reine angew. Math.* 335 (1982), 180–196.
2. N. G. de Bruijn, 'The asymptotic behaviour of a function occurring in the theory of primes', *J. Indian Math. Soc.* (N.S.) 15 (1951), 25–32.
3. N. G. de Bruijn, 'On the number of positive integers $\leqslant x$ and free of prime factors $>y$', *Nederl. Akad. Wetensch. Proc.* Ser. A 54 (=Indag. Math.) 13 (1951), 50–60; II *ibid.* Ser. A 69, 28 (1966), 239–247.
4. K. Dickman, 'On the frequency of numbers containing prime factors of a certain relative magnitude', *Ark. Mat. Astr. Fys.* 22 (1930), 1–14.
5. A. S. Fainleib, 'Local theorems with a remainder term for a certain class of arithmetic functions', *Acta Math. Acad. Sci. Hungar.* 21 (1970), 271–281 (in russian).
6. A. Hildebrand, 'On the number of positive integers $\leqslant x$ and free of prime factors $>y$, *J. Number Theory,* in print.
7. A. Hildebrand, 'Numbers free of large prime factors and the Riemann hypothesis', preprint.
8. A. Hildebrand, 'On the local behaviour of $\Psi(x, y)$', preprint.
9. A. Hildebrand and G. Tenenbaum, 'On integers free of large prime factors', *Trans. Amer. Math. Soc.,* to appear.
10. A. Ivić, 'The distribution of values of the enumerating function of non-isomorphic Abelian groups of finite order', *Arch. Math.* 30 (1978), 374–379.
11. A. Ivić, 'On the number of Abelian groups of a given order and on certain related multiplicative functions', *J. Number Theory* 16 (1983), 119–137.
12. A. Ivić, 'On squarefree numbers with restricted prime factors', *Studia Scien. Math. Hungarica,* in print.
13. D. G. Kendall and R. A. Rankin, 'On the number of Abelian groups of a given order, *Quart. J. Math.* Oxford Ser. (2) 18 (1947), 197–208.
14. E. Krätzel, 'Die Werteverteilung der nicht-isomorphen Abelschen Gruppen endlicher Ordnung und ein verwandtes zahlentheoretisches Problem', Publs. *Inst. Math.* (Belgrade) 31 (45) (1982), 93–101.
15. J.-L. Nicolas, 'Autour de formules dues à A. Selberg, Journées arithmétiques,' Orsay 1982, *Publs. Math.* Orsay 83.04, 122–134 (1983).
16. C. Pomerance, 'On the distribution of round numbers', Lecture Notes 1122, *Number Theory, Proc. Ootacamund India 1984,* (1985), 173–200.

17. A. Rényi, 'On the density of certain sequences of integers', *Publs. Inst. Math.* (Belgrade) 8 (1955), 157–162.
18. W. Schwarz and E. Wirsing, 'The maximal order of non-isomorphic Abelian groups of order *n*', *Arch. Math.* 24 (1973), 59–62.

Katedra Matematike
RGF-a Universiteta u Beogradu
Djušina 7,11000 *Beograd*
Yugoslavia

Département de Mathématiques
Université de Nancy I
54506 *Vandoeuvre les Nancy Cedex*
France