JOURNAL DE THÉORIE DES NOMBRES DE BORDEAUX

ADOLF HILDEBRAND
GERALD TENENBAUM

Integers without large prime factors

Journal de Théorie des Nombres de Bordeaux, tome 5, nº 2 (1993), p. 411-484
<http://www.numdam.org/item?id=JTNB_1993__5_2_411_0>

Journal de Théorie des Nombres
de Bordeaux **5** (1993), 411–484

Integers without large prime factors

by Adolf Hildebrand* and Gérald Tenenbaum

0. Introduction

The purpose of this paper is to give a survey of recent work on the distribution of integers without large prime factors. A similar survey had been published about twenty years ago by K. Norton (1971), but in the intervening time the subject has been considerably advanced and is now in a mature and largely satisfactory state. Moreover, the results have found some new and rather unexpected applications in diverse areas of number theory. We therefore felt it appropriate to give an account of the present state of the subject which may be useful for those interested in studying the subject for its own sake as well as those interested in applying the results. While integers without *small* prime factors may be viewed as approximations to primes and as such form a natural object of study in prime number theory, the reasons for studying integers without *large* prime factors are less obvious. Integers without large prime factors, also called "smooth" integers, are necessarily products of many small prime factors and are, in a sense, the exact opposite of prime numbers. However, it turns out that results on integers without large prime factors play an important auxiliary role in several problems in prime number theory, in particular in the construction of large gaps between primes (Rankin (1938)) and in the analysis of algorithms for factoring and primality testing (Pomerance (1987), Lenstra (1987)). They are also relevant in some other problems in number theory

Manuscrit reçu le 15 mars 1993, version définitive le 28 décembre 1993.
*Supported in part by an NSF Grant while visiting the Institute for Advanced Study at Princeton.

such as bounds on the least kth power non-residues (Vinogradov (1926), Norton (1969,1971)), Waring's problem (Vaughan (1989), Wooley (1992)), and Fermat's conjecture (Lehmer & Lehmer (1941), Granville (1991b)). Most recently, a result on integers of the form $p + 1$ without large prime factors played an essential role in the resolution of a long-standing conjecture of Carmichael (Alford, Granville, & Pomerance (1993)). In addition, the study of integers without large prime factors is an interesting and difficult problem for its own sake that can be attacked by a variety of methods, some of which have led to advances on other problems in analytic number theory.

The principal object of investigation is the function

$$\Psi(x,y) = \#\{1 \le n \le x : P(n) \le y\},$$

where $P(n)$ denotes the largest prime factor of n, with the convention that $P(1) = 1$. The ratio $\Psi(x,y)/[x]$ may be interpreted as the probability that a randomly chosen integer from the interval $[1,x]$ has all its prime factors $\le y$.

Non-trivial estimates for $\Psi(x,y)$ can be obtained by a variety of methods, depending on the relative size of y and x and the nature of the desired result. In Section 1 we shall survey the principal results and discuss some of the methods of proof. Section 2 will be devoted to the Dickman function, a function defined by a differential-difference equation which arises in this connection. In Sections 3 and 4 we shall give complete proofs of two of the main results on $\Psi(x,y)$. In Section 5 we consider the distribution of integers without large prime factors in short intervals and prove a new result in this context (Theorem 5.7) which extends a recent result of Friedlander & Lagarias. Section 6 is devoted to the distribution in arithmetic progressions, and in the final section we survey various other results on integers without large prime factors. We conclude with a comprehensive bibliography of papers on the subject that have appeared since the publication of Norton's memoir. The reader may find another quite thorough list of references in Moree's thesis (1993), which also incudes a clear introduction to the matter as well as interesting new contributions.

Notation. We let $\log_k x$ denote the $k$fold iterated logarithm, defined by $\log_1 x = \log x$ and $\log_k x = \log\log_{k-1} x$ for $k > 1$. Given a complex number s, we denote its real and imaginary parts by σ and τ, respectively. The letter ϵ denotes as usual an arbitrarily small, but fixed constant; other constants will be denoted by c, y_0, etc., and need not be the same at each occurrence. We use the notations $f(x) = O(g(x))$ and $f(x) \ll g(x)$ interchangeably to mean that $|f(x)| \le cg(x)$ holds with some constant c for all

x in a range which will normally be clear from the context. The constant c here is allowed to depend on ϵ (if the functions involved depend on ϵ), but any dependence on other parameters will be explicitly indicated by writing O_k, $\ll_\eta$, etc. We write $f(x) \asymp g(x)$ if both $f(x) \ll g(x)$ and $g(x) \ll f(x)$ hold.

Acknowledgements. We are grateful to Jean-Marie DeKoninck, Heini Halberstam, Aleksandar Ivić, , Pieter Moree, Karl Norton, Carl Pomerance, Eric Saias, J. Sandor, and Jie Wu for their comments on earlier versions of this paper.

1. Estimates for $\Psi(x,y)$ – a survey

We begin by pointing out a natural but false heuristic for the size of $\Psi(x,y)$. It consists of the approximation

$$\Psi(x,y) \approx x \prod_{y<p\le x} \left(1-\frac{1}{p}\right) \tag{1.1}$$

which is suggested by a simple probabilistic argument and also by "extrapolation" from the sieve estimate

$$\#\{n \le x : (n,p) = 1\ (p \in \mathcal{P})\} \asymp x \prod_{p\in\mathcal{P}} \left(1-\frac{1}{p}\right) \tag{1.2}$$

which holds for any set of primes $\mathcal{P}$ contained in $[1, x^{1/2-\epsilon}]$ (see, e.g., Halberstam & Richert (1974)) By Mertens' theorem, the right-hand side of (1.1) is of order x/u with $u = \log x/\log y$. However, as the results below show, $\Psi(x,y)/x$ is in fact exponentially decreasing in u, and therefore of much smaller order of magnitude. The reason for this discrepancy is that the validity of (1.2) depends on certain independence assumptions which are not satisfied if the set $\mathcal{P}$ contains large primes. For example, primes in the interval $(\sqrt{x}, x]$ do not act independently in the sense that if an integer $n \le x$ is divisible by one such prime then it cannot be divisible by any other prime in this interval. We remark that in many applications it is precisely this discrepancy between the expected and the actual size of $\Psi(x,y)$ which is exploited.

The failure of the heuristic (1.1) shows that classical sieve methods are not an appropriate tool for estimating $\Psi(x,y)$. Indeed, these methods would lead to an approximation for $\Psi(x,y)$ consisting of a main term of the order of the right-hand side in (1.1), and an error term which in view of the above remarks would have to be of at least the same order of magnitude as the main term. Thus one cannot hope to obtain a lower bound for $\Psi(x,y)$ in this way. An upper bound by the right-hand side of (1.1) can be deduced, but for most problems this bound is too weak to be useful.

We now derive a simple, but nonetheless useful bound for $\Psi(x,y)$, using a technique known as "Rankin's method"; see Rankin (1938). It is based on the observation that, for any $\sigma > 0$, $x \geq 1$, and $y \geq 2$,

$$\Psi(x,y) \leq \sum_{\substack{n\geq 1\\ P(n)\leq y}} \left(\frac{x}{n}\right)^{\sigma} = x^{\sigma} \sum_{\substack{n\geq 1\\ P(n)\leq y}} \frac{1}{n^{\sigma}}. \tag{1.3}$$

If we take $\sigma = 1 - 1/(2\log y)$, then the last sum may be estimated by

$$\begin{aligned}\sum_{\substack{n\geq 1\\ P(n)\leq y}} \frac{1}{n^{\sigma}} &= \prod_{p\leq y}\left(1-\frac{1}{p^{\sigma}}\right)^{-1} \ll \exp\Big\{\sum_{p\leq y}\frac{1}{p^{\sigma}}\Big\}\\ &\leq \exp\Big\{\sum_{p\leq y}\frac{1}{p} + O\Big((1-\sigma)\sum_{p\leq y}\frac{\log p}{p}\Big)\Big\} \ll \log y.\end{aligned}$$

Substituting this into (1.3) gives the bound

$$\Psi(x,y) \ll xe^{-u/2}\log y, \tag{1.4}$$

where here and in the sequel we set $u = \log x/\log y$. By using a slightly more complicated argument one can remove the factor $\log y$ in (1.4); see Theorem III.5.1 in Tenenbaum (1990a).

An asymptotic formula for $\Psi(x,y)$ was first obtained by Dickman (1930) who proved that for any fixed $u > 0$

$$\lim_{y\to\infty} \Psi(y^u, y)y^{-u} = \rho(u),$$

where the function $\rho(u)$ is defined as the (unique) continuous solution to the differential-difference equation

$$u\rho'(u) = -\rho(u-1) \qquad (u > 1) \tag{1.5}$$

satisfying the initial condition

$$\rho(u) = 1 \qquad (0 \leq u \leq 1). \tag{1.6}$$

(Actually Dickman stated his results in terms of the function $\#\{n \leq x : P(n) \leq n^{1/u}\}$ which, however, is easily seen to be asymptotically equal to $\Psi(x,y)$ when u is fixed.) A more rigorous proof, by modern standards, was later supplied by Chowla and Vijayaraghavan (1947); the first quantitative estimates were obtained by Buchstab (1949) and Ramaswami (1949).

The "Dickman function" $\rho(u)$ is nonnegative for $u > 0$, decreasing for $u > 1$, and satisfies the asymptotic estimate

$$\log \rho(u) = -(1+o(1))u \log u \quad (u \to \infty). \tag{1.7}$$

These and other properties of the Dickman function will be proved in Section 2.

Substantial progress on the problem of estimating $\Psi(x,y)$ was made in the 1950s by de Bruijn. Among other estimates, de Bruijn (1951b) proved a uniform version of Dickman's result; when combined with the sharpest known form of the prime number theorem, his result states that

$$\Psi(x,y) = x\rho(u)\left\{1 + O\left(\frac{\log(u+1)}{\log y}\right)\right\} \tag{1.8}$$

holds (with $u = \log x/\log y$) uniformly in the range

$$y \geq 2, \quad 1 \leq u \leq (\log y)^{3/5-\epsilon}. \tag{1.9}$$

While the error term in (1.8) is best-possible, Hensley (1985) showed that a lower bound of the form $\Psi(x,y) \gg x\rho(u)$ holds in a much larger range, and the range of validity for (1.8) itself has been significantly improved by Maier (unpublished) and Hildebrand (1986a). The latter result gives the largest range in which the asymptotic relation $\Psi(x,y) \sim x\rho(u)$ is known to hold, and we state it formally as a theorem.

THEOREM 1.1. *For any fixed $\epsilon > 0$ the relation* (1.8) *holds uniformly in the range*

$$y \geq 2, \quad 1 \leq u \leq \exp\left\{(\log y)^{3/5-\epsilon}\right\}. \tag{1.10}$$

The upper limit in this range is closely tied to the best known error term in the prime number theorem, and any improvements in the error term would lead to corresponding improvements in the range (1.10). In fact, the correspondence is in both directions; for example, one can show that (1.8) in the form $\Psi(x,y) = x\rho(u)\exp\left\{O(\log(u+1)/\log y)\right\}$ holds uniformly in the range $y \geq 2$, $1 \leq u \leq y^{1/2-\epsilon}$, for any fixed $\epsilon > 0$, if and only if the Riemann Hypothesis is true; see Hildebrand (1984a). The proof of Theorem 1.1 is based on the functional equation

$$\Psi(x,y)\log x = \int_1^x \frac{\Psi(t,y)}{t}\,dt + \sum_{\substack{p^m \leq x \\ p \leq y}} \Psi\left(\frac{x}{p^m}, y\right)\log p, \tag{1.11}$$

which holds for any $x \geq 1$ and $y \geq 2$. This equation is obtained by evaluating the sum $S = \sum_{n \leq x, P(n) \leq y} \log n$ in two different ways. On the one hand, partial summation shows that S is equal to the difference between the left-hand side and the first term on the right of (1.11). On the other hand, writing $\log n = \sum_{p^m | n} \log p$ and inverting the order of summation, we see that S is also equal to the sum on the right of (1.11). The estimate (1.8) is obtained by first showing that it holds in an initial range, say $x \leq y^2$, and then applying (1.11) iteratively to show that it continues to hold in the range $y^{n/2} < x \leq y^{(n+1)/2}$ for $n \geq 4$ as long as (1.10) is satisfied. The limitation in the range comes from the fact that each iteration step involves a small error due to possible irregularities in the distribution of primes. These errors accumulate and after sufficiently many iterations become as large as the main term.

The principal difference between this approach and the earlier approach of de Bruijn lies in the use of (1.11). De Bruijn based his argument on a different functional equation, the "Buchstab identity"

$$\Psi(x, y) = \Psi(x, z) - \sum_{y < p \leq z} \Psi\Big(\frac{x}{p}, p\Big),$$

which holds for all $x \geq 1$ and $z \geq y > 0$. Compared to (1.11), this equation has the disadvantage that the second argument of Ψ is not fixed, so that an induction argument based on this formula is technically more complicated. A more severe limitation of the Buchstab identity is that the right-hand side consists of two terms having opposite signs but which can be nearly equal in size. As a result, the error terms arising from inductive use of the Buchstab identity are much larger than those coming from (1.11), and exceed the main term already for relatively small values of u.

Using the functional equation (1.11), it is relatively easy to justify, at least heuristically, the appearance of the Dickman function in (1.8) if we suppose that a relation of the form $\Psi(x, y) \sim x f(u)$ holds with *some* function $f(u)$. Replacing $\Psi(t, y)$ by $t f(\log t / \log y)$ in (1.11), we obtain under this assumption for $x \geq y \geq 2$

$$x f(u) \log x \approx \int_1^x f\Big(\frac{\log t}{\log y}\Big)\,\mathrm{d}t + x \sum_{\substack{p^m \leq x \\ p \leq y}} \frac{\log p}{p^m} f\Big(u - \frac{\log p^m}{\log y}\Big).$$

Since $\Psi(x, y) \leq x$, we necessarily have $f(u) \leq 1$, so that the first term on the right-hand side is of order $O(x)$. Moreover, the contribution of the terms with $p^m > y$ to the sum on the right is easily shown to be of the same order of magnitude. After dividing by $x\log x$, both of these contributions become $o(1)$ and we obtain

$$\begin{aligned} f(u) &\approx \frac{1}{\log x}\sum_{p^m\le y}\frac{\log p}{p^m}f\Big(u-\frac{\log p^m}{\log y}\Big)\\ &= \frac{1}{\log x}\int_1^y f\Big(u-\frac{\log t}{\log y}\Big)\frac{\mathrm{d}\psi(t)}{t},\\ &\approx \frac{1}{\log x}\int_1^y f\Big(u-\frac{\log t}{\log y}\Big)\frac{\mathrm{d}t}{t} = \frac{1}{u}\int_{u-1}^{u} f(v)\mathrm{d}v, \end{aligned}$$

using the prime number theorem in the form $\psi(t) \sim t$. Assuming the error terms involved all tend to zero as $x, y \to \infty$, we conclude that the function $f(u)$ satisfies the integral equation $f(u)u = \int_{u-1}^{u} f(v)\mathrm{d}v$ for $u > 1$. Differentiating both sides of this equation shows that $f(u)$ is a solution to the differential-difference equation (1.5). Since for $0 < u \le 1$ we have trivially $\Psi(y^u, y) = [y^u] \sim y^u$ as $y \to \infty$, $f(u)$ also satisfies the initial condition (1.6). Therefore, $f(u)$ must be equal to the Dickman function $\rho(u)$.

The range (1.10) is the largest in which an asymptotic approximation for $\Psi(x,y)/x$ by a *smooth* function is known. This range can be considerably increased if instead of an asymptotic formula for $\Psi(x,y)/x$ we only ask for an asymptotic formula for $\log(\Psi(x,y)/x)$, as the following result shows.

THEOREM 1.2. *For any fixed* $\epsilon > 0$ *we have*

$$\log(\Psi(x,y)/x) = \Big\{1 + O\Big(\exp\big\{-(\log u)^{3/5-\epsilon}\big\}\Big)\Big\}\log\rho(u) \tag{1.12}$$

uniformly in the range

$$y \ge 2, \qquad 1 \le u \le y^{1-\epsilon}. \tag{1.13}$$

Moreover, the lower bound in (1.12) *is valid uniformly for all* $x \ge y \ge 2$.

The upper bound in this result is implicit in de Bruijn (1966), whereas the lower bound is due to Hildebrand (1986a). A slightly weaker result had been given by Canfield, Erdős & Pomerance (1983).

Combining (1.12) with the asymptotic formula (1.7) for the logarithm of the Dickman function, we obtain the following simple, but quite useful corollary.

COROLLARY 1.3. *For any fixed* $\epsilon > 0$ *we have*

$$\Psi(x,y) = xu^{-(1+o(1))u},$$

as y *and* u *tend to infinity, uniformly in the range* $u \le y^{1-\epsilon}$.

In terms of the variables x and y, the ranges (1.10) and (1.13) of Theorems 1.1 and 1.2 correspond to

$$(1.10)' \qquad y \geq 2, \qquad \exp\left\{(\log_2 x)^{5/3+\epsilon}\right\} \leq y \leq x$$

and

$$(1.13)' \qquad y \geq 2, \qquad (\log x)^{1+\epsilon} \leq y \leq x,$$

respectively. To give an idea of the size of $\Psi(x,y)$ in various parts of these ranges, we consider the cases $y = \exp\left\{(\log x)^{\alpha}\right\}$ with $0 < \alpha < 1$ and $y = (\log x)^{\alpha}$ with $\alpha > 1$. The corollary yields in the first case

$$\Psi\big(x, \exp\{(\log x)^{\alpha}\}\big) = x \exp\left\{-(\log x)^{1-\alpha+o(1)}\right\},$$

and in the second case

$$(1.14) \qquad \Psi\big(x, (\log x)^{\alpha}\big) = x^{1-1/\alpha+o(1)}.$$

The estimate (1.14) together with the monotonicity of $\Psi(x,y)$ shows that $\big|\log\big(\Psi(x,y)/x\big)\big|$ is of order $\log x$ whenever y is smaller than a fixed power of $\log x$. In this range it is therefore more appropriate to seek an approximation to $\log\big(\Psi(x,y)\big)$ rather than to $\log\big(\Psi(x,y)/x\big)$. Such an estimate has been given by de Bruijn (1966); his result, in a slightly more precise form due to Tenenbaum (1990a, Theorem III.5.2), is as follows.

THEOREM 1.4. *Uniformly for $x \geq y \geq 2$, we have*

$$(1.15) \qquad \log \Psi(x,y) = Z\left\{1 + O\Big(\frac{1}{\log y} + \frac{1}{\log_2 x}\Big)\right\},$$

where

$$(1.16) \qquad Z = Z(x,y) = \frac{\log x}{\log y} \log\Big(1 + \frac{y}{\log x}\Big) + \frac{y}{\log y} \log\Big(1 + \frac{\log x}{y}\Big).$$

The proof of this result is completely elementary. The upper bound is deduced from Rankin's inequality (1.3) with an optimal choice of the parameter σ. The lower bound is based on the elementary inequality

$$(1.17) \quad \Psi(x,y) \geq \binom{k+\ell}{\ell}, \qquad \Big(\ell = \min\{n \in \mathbb{Z} : n \geq u\}, \quad k = \pi(x^{1/\ell})\Big)$$

together with a lower bound for the right-hand side of (1.17) obtained by Stirling's formula. To prove (1.17), it suffices to note that the right-hand side represents the number of k-tuples $(m_1, m_2, \ldots, m_k)$ with $m_i \geq 0$ and $\sum_{i=1}^{k} m_i = \ell$ and that, by the definition of k and the Fundamental Theorem of Arithmetic, the set of such tuples is in one-to-one correspondence with the set of integers composed of exactly ℓ (not necessarily distinct) prime factors $\leq x^{1/\ell} (\leq y)$, and hence with a subset of the set of integers counted in $\Psi(x,y)$.

The estimate (1.15) clearly shows that there is a change of behavior for $\Psi(x,y)$ at $y \approx \log x$. If $y/\log x \to \infty$, then the first term in the definition of Z dominates, whereas for $y = o(\log x)$, Z is asymptotic to the second term in (1.16). The change in behavior is due to the fact that if y is small compared to $\log x$, then many prime factors of a "typical" integer counted in $\Psi(x,y)$ occur to high powers, whereas for larger values of y most prime factors occur only to the first power. For a closer analysis of $\Psi(x,y)$ near the transition point $y \approx \log x$ see Granville (1989).

In the case of very small values of y, it is useful to observe that, by the Fundamental Theorem of Arithmetic, $\Psi(x,y)$ is equal to the number of solutions $(m_p)_{p \leq y}$ in nonnegative integers to the inequality

$$\prod_{p \leq y} p^{m_p} \leq x,$$

or equivalently the linear diophantine inequality

$$\sum_{p \leq y} m_p \log p \leq \log x. \tag{1.18}$$

The number of solutions to this inequality can be estimated very precisely by elementary geometric methods as long as the number of variables (i.e., $\pi(y)$) is not too large. For example, it is easy to see that if $y \geq 2$ is fixed and $x \to \infty$, then the number of solutions is asymptotically equal to the volume of the $\pi(y)$-dimensional simplex defined by the inequalities

$$t_i \geq 0 \quad (i = 1, \ldots, k), \qquad \sum_{i=1}^{k} t_i \log p_i \leq \log x, \tag{1.19}$$

where $k = \pi(y)$ and p_i denotes the ith prime. By a change of variables, the volume of (1.19) is seen to be equal to

$$\mathrm{Vol}\Big(\Big\{(\tau_1, \ldots, \tau_k) \in [0,\infty)^k : \sum_{i=1}^{k} \tau_i \leq 1\Big\}\Big) \prod_{i=1}^{k} \frac{\log x}{\log p_i} = \frac{1}{k!} \prod_{i=1}^{k} \frac{\log x}{\log p_i}.$$

General asymptotic results of this type have been given by Specht (1949), Hornfeck (1959), Beukers (1975), Tenenbaum (1990a) — Theorem III.5.3 — and Granville (1991a). A more careful reasoning leads to the following quantitative result.

THEOREM 1.5. *Uniformly in the range*

$$2 \le y \le (\log x)^{1/2}, \tag{1.20}$$

we have

$$\Psi(x,y) = \frac{1}{\pi(y)!} \prod_{p \le y} \frac{\log x}{\log p} \left\{ 1 + O\left(\frac{y^2}{(\log x)(\log y)} \right) \right\}. \tag{1.21}$$

This result is stated in Ennola (1969), who has also given a similar, but somewhat more complicated formula for the larger range $y \le (\log x)^{3/4-\epsilon}$; for a detailed proof see Theorem III.5.2 of Tenenbaum (1990a).

Theorems 1.1 and 1.5 both give an asymptotic formula for $\Psi(x,y)$; the first result is valid for relatively large values of y and gives an approximation by a smooth function, whereas the second result holds for small y and gives an approximation by a quantity depending on the primes $\le y$. Between the two ranges (1.10) and (1.20), however, there remains a large gap in which the results quoted above give only much weaker estimates. This gap has been closed by the following result of Hildebrand & Tenenbaum (1986), which gives an asymptotic formula for $\Psi(x,y)$ that is valid uniformly in x and y, provided only that $u = \log x / \log y$ and y tend to infinity.

THEOREM 1.6. *Uniformly in the range $x \ge y \ge 2$, we have*

$$\Psi(x,y) = \frac{x^\alpha \zeta(\alpha,y)}{\alpha\sqrt{2\pi\phi_2(\alpha,y)}} \left\{ 1 + O\left(\frac{1}{u} + \frac{\log y}{y} \right) \right\}, \tag{1.22}$$

where

$$\zeta(s,y) = \prod_{p \le y} \left(1 - p^{-s}\right)^{-1}, \tag{1.23}$$

$$\phi(s,y) = \log \zeta(s,y), \qquad \phi_k(s,y) = \frac{\mathrm{d}^k}{\mathrm{d}s^k} \phi(s,y) \quad (k \ge 1), \tag{1.24}$$

and $\alpha = \alpha(x,y)$ is the unique positive solution to the equation

$$-\phi_1(\alpha,y) = \sum_{p \le y} \frac{\log p}{p^\alpha - 1} = \log x. \tag{1.25}$$

The formula (1.22) is similar in nature to the explicit formula in prime number theory. It expresses $\Psi(x,y)$ in terms of the generating Dirichlet series $\zeta(s,y)$ evaluated at a certain point α which plays a role analogous to that of the zeros in the explicit formula. Note that the function $\zeta(\sigma,y)x^\sigma$ is exactly equal to the upper bound (1.3) for $\Psi(x,y)$ obtained by Rankin's method, and that $\alpha(x,y)$ is the unique point on the positive real line which minimizes this function. Thus the denominator in (1.22) measures the ratio between $\Psi(x,y)$ and the upper bound given by Rankin's method with an optimal choice of parameters.

A formula of this type may seem to be of little value at first sight, since it involves the parameter α which is defined only implicitly through an equation involving prime numbers. Nonetheless, from a sufficiently sharp form of the prime number theorem one can derive the estimate

$$\alpha(x,y) = \frac{\log(1+y/\log x)}{\log y}\left\{1+O\Big(\frac{\log\log(1+y)}{\log y}\Big)\right\} \tag{1.26}$$

uniformly in $x \geq y \geq 2$ (see Hildebrand & Tenenbaum (1986)), and using this estimate, one can show that the right-hand of (1.22) is approximated by the right-hand sides of the formulas of Theorems 1.1 – 1.5 in their respective ranges, except in the case of Theorem 1.1 when $u \leq \log y$, say. Thus, Theorem 1.6 implies Theorems 1.2 – 1.5 in their full strength (except for the quality of the error term of (1.21) when $y \ll (\log x)^{1/3}(\log_2 x)^{2/3}$), as well as the statement of Theorem 1.1 for $u \geq \log y$.

We emphasize that Theorem 1.6 does not lead to any improvements in the ranges of validity of the above theorems. For example, it does not yield an asymptotic approximation to $\Psi(x,y)$ by a *smooth* function in the range $u \geq \exp((\log y)^{3/5})$. This is due to our limited knowledge of the distribution of primes which prevents us from obtaining a smooth approximation to the right-hand side of (1.22) in this range. However, even in the range where no smooth approximation to $\Psi(x,y)$ is available, one can use (1.22) to obtain very precise information on the *local* behavior of $\Psi(x,y)$. The following is a typical result of this type, which we quote form Hildebrand & Tenenbaum (1986).

COROLLARY 1.7. *Uniformly for $x \geq y \geq 2$ and $1 \leq c \leq y$, we have*

$$\Psi(cx,y) = \Psi(x,y)c^{\alpha(x,y)}\left\{1+O\Big(\frac{1}{u}+\frac{\log y}{y}\Big)\right\}, \tag{1.27}$$

where $\alpha(x,y)$ is defined as in Theorem 1.6.

Since by (1.26), $\alpha(x,y) = o(1)$ if and only if $y \leq (\log x)^{1+o(1)}$, this shows, for example, that $\Psi(2x,y) \sim \Psi(x,y)$ as $x,y \to \infty$ holds if and only if $y \leq (\log x)^{1+o(1)}$.

In Section 4, we shall give a complete proof of Theorem 1.6 for the range $u \geq (\log y)^4$. The method of proof is analytic; it depends on representing $\Psi(x,y)$ as a complex integral over the function $\zeta(s,y)x^s/s$ and evaluating this integral by the saddle point method. The point $\alpha(x,y)$ is a saddle point for $\zeta(s,y)x^s$, and the expression on the right-hand side of (1.22) arises as the contribution from this saddle point.

The principal deficiency of Theorem 1.6 is that the error term in (1.22) increases as u decreases and for small u becomes larger than the error term of Theorem 1.1. Saias (1989) has shown that one can adapt the saddle point method to deal more effectively with this range. He obtained the following result which is valid in the same range as Theorem 1.1, but gives a much more precise estimate for $\Psi(x,y)$.

THEOREM 1.8. *For any fixed $\epsilon > 0$, and uniformly for $y \geq y_0(\epsilon)$ and $1 \leq u \leq \exp\{(\log y)^{3/5-\epsilon}\}$, we have*

$$\Psi(x,y) = \Lambda(x,y)\Big\{1 + O\Big(\exp\{-(\log y)^{3/5-\epsilon}\}\Big)\Big\}, \tag{1.28}$$

where

$$\Lambda(x,y) = \begin{cases} x\displaystyle\int_{-\infty}^{\infty} \rho(u-v)d([y^v]y^{-v}) & (x \notin \mathbb{N}), \\ \Lambda(x+0,y) & (x \in \mathbb{N}), \end{cases} \tag{1.29}$$

and $\rho(u)$ is defined by (1.5) *and* (1.6) *for $u \geq 0$ and $\rho(u) = 0$ for $u < 0$.*

We shall prove this result in Section 3. The function $\Lambda(x,y)$ has been introduced by de Bruijn (1951b), who obtained an estimate similar to (1.28), but only for a much smaller range. In Lemma 3.1 we shall show that, uniformly in $y \geq y_0(\epsilon)$ and $1 \leq u \leq y^{1-\epsilon}$,

$$\Lambda(x,y) = x\rho(u)\Big\{1 + O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\}. \tag{1.30}$$

Thus the estimate of Theorem 1.8 implies that of Theorem 1.1 if y is sufficiently large. In fact, one can sharpen (1.30), by giving an asymptotic expansion of the O-term in powers of $1/\log y$, resulting in a corresponding sharpening of the estimate of Theorem 1.1; see, for example, Saias (1989).

2. The Dickman function

In this section we investigate the behavior of the Dickman function $\rho(u)$ defined by (1.5) and (1.6). Our principal result is the following asymptotic estimate for $\rho(u)$.

THEOREM 2.1. *For $u > 1$ we have*

$$\rho(u) = \Big\{1 + O\Big(\frac{1}{u}\Big)\Big\}\sqrt{\frac{\xi'(u)}{2\pi}}\exp\{\gamma - u\xi + E(\xi)\}, \tag{2.1}$$

where γ is Euler's constant, $\xi = \xi(u)$ is the unique positive solution to the equation

$$e^{\xi} = 1 + u\xi \tag{2.2}$$

and $E(s)$ is defined by

$$E(s) = \int_0^s \frac{e^t - 1}{t}dt. \tag{2.3}$$

The asymptotic formula implicit in (2.1) has been first proved by de Bruijn (1951a) using contour integration and the saddle point method. Canfield (1982) gave a combinatorial proof of this formula, and an arithmetic proof via the function $\Psi(x, y)$ is contained in Hildebrand & Tenenbaum (1986). The above quantitative result is due to Alladi (1982b) who used de Bruijn's method. It has been sharpened by Smida (1993), who essentially replaces the factor $1 + O(1/u)$ in (2.1) by a function having an explicit asymptotic expansion in terms of (negative) powers of u and powers of $\xi(u)$. Another type of expansion, as a convergent series of analytic functions, may be derived from a general theorem of Hildebrand & Tenenbaum (1993) on the solutions of differential-difference equations. A different, but quite complicated expansion has recently been given by Xuan (1993). The proof we shall give for Theorem 2.1 here is taken from Tenenbaum (1990a).

We remark that (2.1) can be written as

$$\rho(u) = \Big\{1 + O\Big(\frac{1}{u}\Big)\Big\}\sqrt{\frac{\xi'(u)}{2\pi}}\exp\Big\{\gamma - \int_1^u \xi(t)\mathrm{d}t\Big\}, \tag{2.1$'$}$$

a relation which is often more convenient to work with. This follows from the identity

$$u\xi - E(\xi) = \int_1^u \xi(t)\mathrm{d}t,$$

which can be seen on noting that

$$\frac{\mathrm{d}}{\mathrm{d}u}\big(u\xi - E(\xi)\big) = \Big(u - \frac{e^\xi - 1}{\xi}\Big)\xi'(u) + \xi(u) = \xi(u) \quad (u > 1)$$

by (2.2), and

$$\lim_{u\to 1+}\big\{u\xi(u) - E\big(\xi(u)\big)\big\} = 0,$$

since $\xi(u) \to 0$ as $u \to 1+$.

The function $\xi(u)$ is non-elementary, but it is not hard to obtain an expansion for this function in powers of $\log u$ and $\log_2 u$. We shall prove a simple result of this kind.

Lemma 2.2. *For $u > 1$ we have*

$$\xi(u) = \log u + \log_2(u+2) + O\left(\frac{\log_2(u+2)}{\log(u+2)}\right), \tag{2.4}$$

$$\xi'(u) = \frac{1}{u - (u-1)/\xi} = \frac{1}{u}\exp\Big\{O\Big(\frac{1}{\log(u+1)}\Big)\Big\}. \tag{2.5}$$

Proof. Let $f(x) = (e^x - 1)/x = \int_0^1 e^{tx}\mathrm{d}t$. The function $f(x)$ tends to 1 as $x \to 0+$, and it is strictly increasing in x since

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^1 e^{tx}\mathrm{d}t = \int_0^1 te^{tx}\mathrm{d}t > 0.$$

This shows that for each $u > 1$ the equation $f(\xi) = u$ (i.e., (2.2)) has a unique solution $\xi = \xi(u) > 0$. The monotonicity of f together with the observation that we have for any sufficiently large constant c and all $u > 1$, writing $v = \log_2(u+2)/\log(u+2)$,

$$f\Big(\log u + \log_2(u+2) \pm cv\Big) = \frac{u\log(u+2)e^{\pm cv} - 1}{\log u + \log_2(u+2) \pm cv} \gtrless u$$

implies the estimate (2.4).

Differentiating both sides of (2.2) with respect to u gives

$$u\xi'(u) + \xi = e^{\xi}\xi'(u) = (1 + u\xi)\xi'(u)$$

and hence the first equality in (2.5). The second equality follows from the first and the estimate (2.4), provided u is sufficiently large. To complete the proof of (2.5), it suffices to show that $\xi'(u) \asymp 1/u$ holds uniformly in $u > 1$. We have

$$\begin{aligned} u = f(\xi) &= \int_0^1 e^{t\xi}\mathrm{d}t \geq \int_0^1 te^{t\xi}\mathrm{d}t = f'(\xi) \\ &\geq \tfrac{1}{2}\int_{1/2}^1 e^{t\xi}\mathrm{d}t \geq \tfrac{1}{4}\int_0^1 e^{t\xi}\mathrm{d}t = \tfrac{1}{4}u, \end{aligned}$$

and since $\xi'(u) = 1/f'(\xi)$, we obtain $1/u \leq \xi'(u) \leq 4/u$ as required.

Substituting the estimates of the lemma into (2.1)$'$ we obtain the following corollary.

COROLLARY 2.3. *For $u \geq 1$ we have*

$$\rho(u) = \exp\left\{-u\left(\log u + \log_2(u+2) - 1 + O\left(\frac{\log_2(u+2)}{\log(u+2)}\right)\right)\right\}. \tag{2.6}$$

The estimates (2.4) and (2.6) can be refined and one can in principle approximate $\xi(u)$ and $\log(\rho(u)/u)$ for any given $k \geq 1$ by an elementary function to within an error of order $O((\log u)^{-k})$; see, for example, de Bruijn (1951a).

Many applications require estimates for ratios of the form $\rho(u-v)/\rho(u)$ rather than estimates for $\rho(u)$ itself. Theorem 2.1 leads to the following result of this kind.

COROLLARY 2.4. *For $u > 2$ and $|v| \le u/2$ we have*

$$\rho(u-v) = \rho(u)\exp\left\{v\xi(u) + O\Big(\frac{1+v^2}{u}\Big)\right\}. \tag{2.7}$$

Moreover, for $u > 1$ and $0 \le v \le u$ we have

$$\rho(u-v) \ll \rho(u)e^{v\xi(u)}. \tag{2.8}$$

To prove this result, we apply (2.1)$'$ with u and with $u-v$, getting

$$\frac{\rho(u-v)}{\rho(u)} = \sqrt{\frac{\xi'(u-v)}{\xi'(u)}}\exp\left\{\int_{u-v}^{u}\xi(t)\mathrm{d}t + O\Big(\frac{1}{u-v}\Big)\right\} \tag{2.9}$$

for $0 \le v < u-1$. If, in addition, $0 \le v \le u/2$ then $1 < u-v \asymp u$ and hence

$$\begin{aligned}\int_{u-v}^{u}\xi(t)\mathrm{d}t &= v\xi(u) - \int_{u-v}^{u}\xi'(t)(t-u+v)\mathrm{d}t \\ &= v\xi(u) + O\Big(\int_{u-v}^{u}\frac{v}{t}\mathrm{d}t\Big) = v\xi(u) + O\Big(\frac{v^2}{u}\Big)\end{aligned} \tag{2.10}$$

Differentiating the relation $f(\xi) = u$ twice with respect to u, we obtain

$$\xi''(u) = -\xi'(u)^2\frac{f''(\xi)}{f'(\xi)} = -\xi'(u)^3 f''(\xi).$$

By Lemma 2.2 and the inequality $|f''(\xi)| = \int_0^1 t^2e^{t\xi}\mathrm{d}t \le f(\xi) = u$ this implies $\xi''(u) \ll 1/u^2$. It follows that, for $u > 2$ and $0 \le v \le u/2$,

$$\xi'(u-v)/\xi'(u) = \exp\{O(v/u)\} = \exp\big\{O\big((1+v^2)/u\big)\big\}.$$

Together with (2.10) and (2.9) this implies (2.7) in the case $0 \le v \le u/2$, and a similar argument gives (2.7) when $-u/2 \le v \le 0$.

To prove (2.8), we may assume that $0 \le v < u-1$, since the right-hand side is an increasing function of v, while the left-hand side is equal to 1 for $u-1 \le v \le u$. We can therefore apply the first estimate of (2.10). By Lemma 2.2 the integral involving $\xi'(t)$ is of order $\gg \int_{u-v}^{u}(t-u+v)/t\mathrm{d}t \gg v^2/u$, whereas, for $0 \le v < u-1$,

$$\log(\xi'(u-v)/\xi'(u)) = \log(u/(u-v)) + O(1).$$

Thus, for a suitable constant $c > 0$, the right-hand side of (2.9) is bounded by

$$\ll \exp\big\{v\xi(u) - cv^2/u + \log\big(u/(u-v)\big)\big\} \ll e^{v\xi},$$

and (2.8) follows.

We now proceed to prove Theorem 2.1. We begin by establishing some elementary properties of the function $\rho(u)$. It is convenient here to set $\rho(u) = 0$ for $u < 0$, so that ρ is defined on the entire real line.

LEMMA 2.5. *We have*

$$u\rho(u) = \int_{u-1}^{u} \rho(v)\mathrm{d}v \quad (u \in \mathbb{R}), \tag{2.11}$$

$$\rho(u) > 0 \quad (u \geq 0), \tag{2.12}$$

$$\rho'(u) < 0 \quad (u > 1), \tag{2.13}$$

$$0 < \rho(u) \leq \frac{1}{\Gamma(u+1)} \quad (u \geq 0). \tag{2.14}$$

Proof. The first relation holds for $0 \leq u < 1$ since by definition $\rho(v) = 0$ for $v < 0$ and $\rho(v) = 1$ for $0 \leq v \leq 1$. It remains valid for $u > 1$ since $\rho(u)$ is a continuous function and by (1.5) the derivatives of both sides of (2.11) are equal in this range.

Inequality (2.12) is a consequence of (2.11), (1.5), and the continuity of $\rho(u)$. For suppose (2.12) is false, and let $u_0 = \inf\{u \geq 0 : \rho(u) \leq 0\}$, so that $1 \leq u_0 < \infty$. By the continuity of $\rho(u)$ this would imply $\rho(u_0) = 0$ and $\int_{u_0-1}^{u_0} \rho(u)\mathrm{d}v > 0$, which contradicts (2.11).

Inequality (2.13) follows from (2.12) and (1.5).

The last inequality of the lemma is true for $0 \leq u \leq 1$, since $\rho(u) = 1$ and $\Gamma(u+1) < 1$ in this range. Assuming that it holds for $k \leq u < k+1$ for some $k \geq 0$, we deduce by (2.11) and the monotonicity of $\rho(u)$ that for $k+1 \leq u < k+2$, $\rho(u) \leq \rho(u-1)/u \leq 1/u\Gamma(u) = 1/\Gamma(u+1)$. Hence, by induction, (2.14) holds for all $u \geq 0$, and the proof of the lemma is complete.

We next investigate the Laplace transform of $\rho(u)$, defined by

$$\widehat{\rho}(s) = \int_{0}^{\infty} \rho(u)e^{-us}\mathrm{d}u. \tag{2.15}$$

By (2.14), the integral in (2.15) is uniformly convergent in any compact region in the complex s-plane.

LEMMA 2.6. *We have*

$$\widehat{\rho}(s) = e^{\gamma+E(-s)}, \tag{2.16}$$

where γ is Euler's constant and $E(s)$ is defined by (2.3).

Proof. Using (2.11) we obtain

$$-\frac{\mathrm{d}}{\mathrm{d}s}\widehat{\rho}(s) = \int_0^\infty \rho(u)ue^{-su}\mathrm{d}u = \int_0^\infty \Big(\int_{u-1}^u \rho(v)\mathrm{d}v\Big)e^{-su}\mathrm{d}u$$
$$= \int_0^\infty \rho(v)\Big(\int_v^{v+1} e^{-su}\mathrm{d}u\Big)\mathrm{d}v = \frac{1-e^{-s}}{s}\widehat{\rho}(s).$$

The solution to this differential equation is of the form

$$\widehat{\rho}(s) = C\exp\left\{-\int_0^s \frac{1-e^{-t}}{t}\mathrm{d}t\right\} = Ce^{E(-s)},$$

where C is a constant. To determine the value of C, we examine the behavior of $\widehat{\rho}(s)$ and $e^{E(-s)}$ as $s \to \infty$ along the positive real axis. On the one hand, the definition of $\widehat{\rho}(s)$ gives $\lim_{s\to\infty} s\widehat{\rho}(s) = \lim_{u\to 0+}\rho(u) = 1$. On the other hand, the relation

$$E(-s) = -\log s - \gamma - E_1(s), \tag{2.17}$$

where

$$E_1(s) = \int_s^\infty \frac{e^{-t}}{t}\mathrm{d}t,$$

(cf. Abramowitz & Stegun (1964), p. 228), implies $\lim_{s\to\infty} e^{E(-s)}s = e^{-\gamma}$. Hence $C = e^\gamma$, and (2.16) follows.

LEMMA 2.7. *Let $u > 1$, $\xi = \xi(u)$, and $s = -\xi + i\tau$. Then we have*

$$|\widehat{\rho}(s)| \ll \begin{cases} \exp\left\{E(\xi) - \dfrac{\tau^2 u}{2\pi^2}\right\} & (|\tau| \le \pi), \\ \exp\left\{E(\xi) - \dfrac{u}{\pi^2+\xi^2}\right\} & (|\tau| > \pi), \end{cases} \tag{2.18}$$

and

$$\widehat{\rho}(s) = \frac{1}{s}\exp\left\{O\Big(\frac{u\log(u+1)}{|\tau|}\Big)\right\} \qquad (|\tau| > 1). \tag{2.19}$$

Proof. The second estimate follows immediately from Lemma 2.6, (2.17) and the inequality

$$|E_1(s)| = \left|\int_0^\infty \frac{e^{-(s+t)}}{s+t}\mathrm{d}t\right| \le \frac{e^\xi}{|\tau|} = \frac{1+u\xi}{|\tau|} \ll \frac{u\log(u+1)}{|\tau|}.$$

The first estimate is by Lemma 2.6 equivalent to the inequality

$$H(\tau) \geq \begin{cases} \dfrac{\tau^2 u}{2\pi^2} + O(1) & (|\tau| \leq \pi), \\ \dfrac{u}{\pi^2 + \xi^2} + O(1) & (|\tau| > \pi), \end{cases} \tag{2.20}$$

where

$$H(\tau) = E(\xi) - \operatorname{Re} E(\xi - i\tau).$$

A change of variables gives

$$H(\tau) = \int_0^{\xi} \frac{e^t - 1}{t} \mathrm{d}t - \operatorname{Re} \int_0^{\xi - i\tau} \frac{e^t - 1}{t} \mathrm{d}t = \int_0^1 e^{t\xi} \frac{1 - \cos(t\tau)}{t} \mathrm{d}t.$$

For $|\tau| \leq \pi$ and $0 \leq t \leq 1$ we have $1 - \cos(t\tau) \geq 2(t\tau)^2/\pi^2$ and thus

$$H(\tau) \geq \frac{2\tau^2}{\pi^2} \int_0^1 t e^{t\xi} \mathrm{d}t.$$

The integral here is

$$\geq \tfrac{1}{2} \int_{1/2}^1 e^{t\xi} \mathrm{d}t \geq \tfrac{1}{4} \int_0^1 e^{t\xi} \mathrm{d}t = \tfrac{1}{4} u$$

by (2.2). This gives the first inequality in (2.20).

To prove the second inequality in (2.20), we use the estimate

$$\begin{aligned} H(\tau) &\geq \int_0^1 e^{t\xi} \big(1 - \cos(t\tau)\big) \mathrm{d}t = \frac{e^{\xi} - 1}{\xi} - \operatorname{Re} \Big(\frac{e^{\xi + i\tau} - 1}{\xi + i\tau} \Big) \\ &= u - \operatorname{Re} \Big(\frac{(1 + u\xi) e^{i\tau} - 1}{\xi + i\tau} \Big) \geq u \Big(1 - \frac{\xi}{|\xi + i\tau|} \Big) + O(1). \end{aligned}$$

For bounded values of u, (2.20) holds trivially since $H(\tau) \geq 0$. On the other hand, if u, and therefore ξ, are sufficiently large and $|\tau| \geq \pi$, then the main term in the last expression is greater than or equal to

$$u \Big(1 - \frac{1}{\sqrt{1 + (\pi/\xi)^2}} \Big) \geq \frac{u\pi^2}{2\xi^2} + O\Big(\frac{u}{\xi^3} \Big) \geq \frac{u}{\xi^2 + \pi^2} + O(1),$$

and (2.20) follows again.

Proof of Theorem 2.1. Since $\rho(u)$ is continuously differentiable for $u > 1$ and by Lemma 2.5 the Laplace transform $\widehat{\rho}(s)$ is absolutely convergent for any s, the Laplace inversion formula is applicable and gives

$$\rho(u) = \frac{1}{2\pi i} \int_{\kappa - i\infty}^{\kappa + i\infty} \widehat{\rho}(s) e^{us} \mathrm{d}s \tag{2.21}$$

for $u > 1$ and any $\kappa \in \mathbb{R}$. We choose $\kappa = -\xi$ with $\xi = \xi(u)$ defined by (2.2). This choice is suggested by the saddle point method, since the point $s = -\xi$ is a saddle point of the integrand in (2.21), i.e., a zero of the function

$$\frac{\mathrm{d}}{\mathrm{d}s}\big(\log \widehat{\rho}(s) + us\big) = \frac{e^{-s} - 1}{s} + u.$$

We begin by showing that the main contribution to the integral in (2.21) comes from the range $|\tau| \le \delta$, where $\delta = \pi\sqrt{2\log(u+1)/u}$. To this end we divide the remaining range into the three parts $\delta < |\tau| \le \pi$, $\pi < |\tau| \le u\log(u+1)$, and $|\tau| > u\log(u+1)$. Using the bounds for $\widehat{\rho}(s)$ provided by Lemma 2.7, we obtain for the contributions of these ranges the estimates

$$\ll e^{-u\xi + E(\xi)} \int_\delta^\pi e^{-\tau^2 u/2\pi^2} \mathrm{d}\tau \ll \frac{e^{-u\xi+E(\xi)}}{\sqrt{u}} \int_{\log(u+1)}^\infty e^{-t} \frac{\mathrm{d}t}{\sqrt{t}} \ll \frac{e^{-u\xi+E(\xi)}}{u^{3/2}},$$

$$\ll u\log(u+1)\exp\left\{ -u\xi + E(\xi) - \frac{u}{\pi^2+\xi^2}\right\} \ll \frac{e^{-u\xi+E(\xi)}}{u^{3/2}},$$

and

$$\ll \left| \int_{u\log(u+1)}^\infty \frac{e^{-u\xi + i\tau u}}{-\xi + i\tau} \mathrm{d}\tau \right| + u\log(u+1) \int_{u\log(u+1)}^\infty \frac{e^{-u\xi}}{\tau^2} \mathrm{d}\tau \ll e^{-u\xi},$$

respectively. Since the right-hand side of (2.1) is of order $e^{-u\xi+E(\xi)}/\sqrt{u}$ by Lemma 2.2 and since $E(\xi) \asymp (e^\xi - 1)/\xi = u$, each of these contributions is by a factor $\ll 1/u$ smaller than the right-hand side of (2.1) and thus is absorbed by the error term in (2.1).

It remains to deal with the range $|\tau| \le \delta$. Using Lemma 2.6, we write the integrand in (2.21) as

$$\widehat{\rho}(s) e^{us} = e^{\gamma + E(-s) + us} = e^{\gamma + E(\xi - i\tau) - u\xi + i\tau u} \tag{2.22}$$

and expand the function in the exponent in a Taylor series about $\tau = 0$. Since $E'(\xi) = (e^\xi - 1)/\xi = u$, we have

(2.23)
$$E(\xi-i\tau) = E(\xi)-i\tau u-\tfrac{1}{2}\tau^2E''(\xi)+\tfrac{1}{6}i\tau^3E'''(\xi)+O\Big(\tau^4\max_{|t|\le|\tau|}|E^{(4)}(\xi+it)|\Big).$$

From the inequality

$$|E^{(k)}(s)| = \left|\int_0^1 t^{k-1}e^{ts}\mathrm{d}t\right| \le \int_0^1 e^{t\xi}\mathrm{d}t = u \qquad (k\ge 1,\ \operatorname{Re} s = \xi)$$

we see that the two last terms in (2.23) are of order $O(|\tau|^3u)$ and $O(\tau^4u)$, respectively, and hence of order $O(1)$ for $|\tau| \le \delta = \pi\sqrt{2\log(u+1)/u}$. Applying the estimates

$$e^z = 1+z+O(|z|^2) = 1+O(|z|) \qquad (|z|\ll 1)$$

to the exponentials of these terms, we can then write (2.22) as

$$\Big(1+\tfrac{1}{6}i\tau^3E'''(\xi)+O\big(u\tau^4+u^2\tau^6\big)\Big)e^{\gamma+E(\xi)-u\xi-\frac{1}{2}\tau^2E''(\xi)}.$$

Integrating this function over the interval $-\delta\le\tau\le\delta$, we obtain

$$\begin{aligned}(2.24)\qquad &\frac{1}{2\pi i}\int_{\kappa-i\delta}^{\kappa+i\delta}\widehat{\rho}(s)e^{us}\mathrm{d}s\\ &= \frac{1}{2\pi}e^{\gamma-u\xi+E(\xi)}\int_{-\delta}^{\delta}\Big(1+O\big(u\tau^4+u^2\tau^6\big)\Big)e^{-\tau^2E''(\xi)/2}\mathrm{d}\tau,\end{aligned}$$

since the contribution from the term $\frac{1}{6}i\tau^3E'''(\xi)$ to the integral is zero for symmetry reasons. The contribution from the error term in (2.24) is bounded by

$$\ll \frac{e^{-u\xi+E(\xi)}}{\sqrt{E''(\xi)}}\left(\frac{u}{E''(\xi)^2}+\frac{u^2}{E''(\xi)^3}\right),$$

and extending the range of integration in the main term to $(-\infty,\infty)$ introduces an error of order

$$\ll e^{-u\xi+E(\xi)}\int_\delta^\infty e^{-\tau^2E''(\xi)/2}\mathrm{d}\tau \ll \frac{e^{-u\xi+E(\xi)}}{\sqrt{E''(\xi)}}\int_{\delta^2E''(\xi)/2}^\infty \frac{e^{-t}}{\sqrt{t}}\mathrm{d}t.$$

Since by Lemma 2.2

$$\begin{aligned}(2.25)\qquad E''(\xi) &= \frac{\mathrm{d}}{\mathrm{d}\xi}\frac{e^\xi-1}{\xi} = \frac{e^\xi}{\xi}-\frac{e^\xi-1}{\xi^2} = u\Big(1-\frac{1}{\xi}\Big)+\frac{1}{\xi}\\ &= \frac{1}{\xi'(u)} = u\exp\Big\{O\Big(\frac{1}{\log(u+1)}\Big)\Big\}\end{aligned}$$

and $\frac{1}{2}\delta^2 E''(\xi) = \pi^2 \log(u+1)\exp\{O(1/\log(u+1))\}$, the coefficient of $e^{-u\xi+E(\xi)}/\sqrt{E''(\xi)}$ in both of these error terms is of order $\ll 1/u$. Hence, these terms are absorbed by the error term in (2.1). The main term on the right-hand side of (2.24) with the integral taken over $(-\infty,\infty)$ equals

$$\frac{1}{2\pi}e^{\gamma-u\xi+E(\xi)}\int_{-\infty}^{\infty} e^{-\tau^2 E''(\xi)/2}d\tau = \frac{1}{2\pi}e^{\gamma-u\xi+E(\xi)}\sqrt{\frac{2\pi}{E''(\xi)}},$$

which by (2.25) reduces to the main term in (2.1). This completes the proof of Theorem 2.1.

3. $\Psi(x,y)$ for small u

In this section, we will prove Theorem 1.8. We begin with two lemmas giving estimates for the approximating function $\Lambda(x,y)$.

LEMMA 3.1. *For any fixed $\epsilon > 0$ and uniformly in $y \ge y_0(\epsilon)$ and $1 \le u \le y^{1-\epsilon}$, we have $\Lambda(x,y) \gg x\rho(u)$ and*

$$\Lambda(x,y) = x\rho(u)\Big\{1 + O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\}. \tag{3.1}$$

Moreover, uniformly in $x > 1$ and $y \ge 2$ we have

$$|\Lambda(x,y)| \ll x\Big\{(u+1)^{-u} + y^{-u/3}\Big\}. \tag{3.2}$$

Proof. We may assume that x is not an integer, in which case $\Lambda(x,y)$ is given by

$$\Lambda(x,y) = x\int_{-\infty}^{\infty}\rho(u-v)d([y^v]y^{-v}) = -x\int_{0-}^{u}\rho(u-v)d(\{y^v\}y^{-v}), \tag{3.3}$$

where $\{t\} = t - [t]$. Using integration by parts, we obtain

$$\begin{aligned}\int_{0-}^{u}\rho(u-v)d(\{y^v\}y^{-v}) &= \rho(u-v)\{y^v\}y^{-v}\Big|_{0-}^{u} + \int_0^u \rho'(u-v)\{y^v\}y^{-v}dv \\ &= -\rho(u) + \{y^u\}y^{-u} + \int_0^{u-1}\rho'(u-v)\{y^v\}y^{-v}dv,\end{aligned} \tag{3.4}$$

where the last integral is to be interpreted as zero if $0 < u < 1$. By (1.5) and Corollary 2.4 we have, for $0 \le v < u-1$,

$$-\rho'(u-v) = \frac{\rho(u-v-1)}{u-v} \ll \frac{\rho(u)e^{(v+1)\xi(u)}}{u-v} = \rho(u)\big(1+u\xi(u)\big)\frac{e^{v\xi(u)}}{u-v}.$$

By Lemma 2.2, this is $\ll \rho(u)u\log(u+1)y^{v(1-\epsilon/2)}/(u-v)$ in the range $1 \le u \le y^{1-\epsilon}$. Hence the last integral in (3.4) is bounded by

$$\ll \rho(u)u\log(u+1)\int_0^{u-1}\frac{y^{-\epsilon v/2}}{u-v}\mathrm{d}v \ll \rho(u)\frac{\log(u+1)}{\log y}$$

in this range. In the same range we have

$$y^{-u} \le \left(uy^{\epsilon}\right)^{-u} \ll \rho(u)y^{-\epsilon u/2}$$

by Corollary 2.3. Inserting these estimates in (3.3) and (3.4) gives (3.1).

From (3.3) and (3.4) we also obtain $\Lambda(x,y) \ge x\big(\rho(u) - y^{-u}\big)$, and since, by Corollary 2.3, $y^{-u} \le \frac{1}{2}\rho(u)$ for $y \ge y_0(\epsilon)$ and $1 \le u \le y^{1-\epsilon}$, we see that the lower bound $\Lambda(x,y) \gg x\rho(u)$ holds in this range. The bound (3.2) follows from (3.3) and (3.4) for $0 < u \le 1$, and from (3.1) and and Corollary 2.3 for $1 \le u \le \sqrt{y}$. For $u > \sqrt{y}$ we obtain (3.2) if we estimate the last integral in (3.4) as follows, using (1.5), the monotonicity of ρ and Corollary 2.3,

$$\begin{aligned}&\le \int_0^{u-1}\frac{\rho(u-v-1)}{u-v}y^{-v}\mathrm{d}v \ll y^{-u/3} + \int_0^{\min(u-1,u/3)}\rho(u-v-1)\mathrm{d}v\\ &\ll y^{-u/3} + u\rho\big(\max(0,\tfrac{2}{3}u-1)\big) \ll y^{-u/3} + (u+1)^{-2u/3} \ll y^{-u/3}.\end{aligned}$$

LEMMA 3.2. *Uniformly for any $\epsilon > 0$, $y \ge y_0(\epsilon)$, $1 < u \le y^{1/3-\epsilon}$ and $T \ge u^5$, we have*

$$\Lambda(x,y) = \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} F(s,y)\frac{x^s}{s}\mathrm{d}s + O\Big(x^{\sigma}T^{-1/4} + x^{3/4}\Big), \tag{3.5}$$

where $\sigma = 1 - \xi(u)/\log y$,

$$F(s,y) = \zeta(s)(s-1)(\log y)\widehat{\rho}\big((s-1)\log y\big) \tag{3.6}$$

and $\widehat{\rho}(s)$ is defined by (2.15).

Proof. It suffices to prove (3.5) when x is not an integer. Set

$$\lambda_y(u) = \Lambda(y^u, y)y^{-u}.$$

By Lemma 3.1, $\lambda_y(u)$ is bounded on any finite interval in $(0,\infty)$, and satisfies $\lambda_y(u) \ll y^{-u/3}$ for large u. Hence the Laplace transform

$$\widehat{\lambda}_y(s) = \int_0^{\infty}\lambda_y(u)e^{-us}\mathrm{d}u$$

is absolutely convergent for $\operatorname{Re} s > -\frac{1}{3}(\log y)/3$. By the definition of $\Lambda(y^u, y)$ and the convolution theorem for Laplace transforms we have

$$\widehat{\lambda}_y(s) = \widehat{\rho}(s)G_y(s), \qquad \text{where} \qquad G_y(s) = \int_{0-}^{\infty} e^{-us} d\big([y^u]y^{-u}\big).$$

Setting $t = y^u$ and $w = 1 + s/\log y$, we obtain

$$\begin{aligned} G_y(s) &= \int_{1-}^{\infty} t^{-(w-1)} \mathrm{d}\big([t]t^{-1}\big) = \int_{1-}^{\infty} t^{-w}\mathrm{d}[t] - \int_{1}^{\infty} [t]t^{-1-w}\mathrm{d}t \\ &= \int_{1-}^{\infty} t^{-w}\mathrm{d}[t] + \Big[\frac{1}{w}[t]t^{-w}\Big]_{1-}^{\infty} - \frac{1}{w}\int_{1-}^{\infty} t^{-w}\mathrm{d}[t] = \Big(1 - \frac{1}{w}\Big)\zeta(w). \end{aligned}$$

Hence

$$\widehat{\lambda}_y(s) = \widehat{\rho}\big((w-1)\log y\big)\Big(1 - \frac{1}{w}\Big)\zeta(w) = \frac{F(w,y)}{w\log y}.$$

By the Laplace inversion formula it follows that

$$\lambda_y(u) = \frac{1}{2\pi i}\int_{-\xi(u)-i\infty}^{-\xi(u)+i\infty} \widehat{\lambda}_y(s)e^{us}\mathrm{d}s = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{F(w,y)}{w}x^{w-1}dw$$

for any $u \geq 1$ and $\sigma = 1 - \xi(u)/\log y > \frac{2}{3}$ such that $x = y^u$ is not a positive integer. To obtain (3.5) (for $x \notin \mathbb{N}$), it therefore remains to show that the estimate

$$\left|\int_{|\tau|>T} \frac{F(s,y)}{s}x^s\mathrm{d}s\right| \ll x^\sigma T^{-1/4} + x^{3/4} \tag{3.7}$$

holds for $\sigma = 1 - \xi(u)/\log y$ and $T \geq u^5$, since by Lemma 2.2 and the given bounds on y and u, $\xi(u)/\log y < \frac{1}{3}$ if the constant $y_0(\epsilon)$ is large enough. By Lemma 2.7 we have for $\sigma = 1 - \xi(u)/\log y$ and $|\tau| \geq T \geq u^5$

$$\widehat{\rho}((s-1)\log y)(s-1)\log y = \exp\left\{O\Big(\frac{u\log(u+1)}{|\tau|\log y}\Big)\right\} = 1 + O\Big(\frac{1}{|\tau|^{3/4}}\Big)$$

and therefore

$$F(s,y) = \zeta(s)\Big(1 + O\Big(\frac{1}{|\tau|^{3/4}}\Big)\Big).$$

Hence the left-hand side of (3.7) is bounded by

$$\ll \left|\int_{|\tau|>T} \frac{\zeta(s)}{s}x^s\mathrm{d}s\right| + x^\sigma\int_{|\tau|>T} \frac{|\zeta(s)|}{|\tau|^{7/4}}|\mathrm{d}s|.$$

Since $|\zeta(s)| \ll |\tau|^{1/2}$ for $\sigma \geq \frac{1}{2}$ and $|\tau| \geq 1$, the second term here is bounded by $\ll x^\sigma T^{-1/4}$, and hence by the right-hand side of (3.7). Using the approximate functional equation for the zeta function in the form

$$\begin{aligned}\zeta(s) &= \sum_{n \leq |\tau|} n^{-s} + \frac{|\tau|^{1-s}}{1-s} + O\big(|\tau|^{-\sigma}\big) \\ &= \sum_{n \leq |\tau|} n^{-s} + O\big(|\tau|^{-\sigma}\big) \qquad \big(\sigma \geq \tfrac{1}{2}, \tau \neq 0\big)\end{aligned}$$

(see, for example, Corollary II.3.5.1 in Tenenbaum (1990a)), we obtain for the first integral the bound

$$\ll x^\sigma \sum_{n \geq 1} n^{-\sigma} \left| \int_{|\tau| \geq T_n} \left(\frac{x}{n}\right)^{i\tau} \frac{\mathrm{d}\tau}{\sigma + i\tau} \right| + O\Big(x^\sigma \int_T^\infty \frac{\mathrm{d}\tau}{|\tau|^{\sigma+1}}\Big)$$

with $T_n = \max(T, n)$. The last term in this expression is of order $\ll x^\sigma T^{-\sigma} \leq x^\sigma T^{-2/3}$ and hence bounded by the right-hand side of (3.7). In the first term we apply the estimate

$$\left| \int_{|\tau| \geq T_n} \left(\frac{x}{n}\right)^{i\tau} \frac{\mathrm{d}\tau}{\sigma + i\tau} \right| \ll \min\Big(1, \frac{1}{T_n |\log(x/n)|}\Big),$$

which is easily proved by an integration by parts, and split the range of summation into two parts according as $|\log(x/n)| \leq T_n^{-1/4}$ or not. This yields the bound

$$\begin{aligned}&\ll \sum_{|\log(x/n)| \leq T_n^{-1/4}} \left(\frac{x}{n}\right)^\sigma + x^\sigma \sum_{n=1}^\infty \frac{1}{n^\sigma T_n^{3/4}} \\ &\ll x^{3/4} + \frac{x^\sigma}{T^{3/4}} \sum_{n \leq T} \frac{1}{n^{\sigma+3/4}} + x^\sigma \sum_{n > T} \frac{1}{n^{\sigma+1}} \\ &\ll x^{3/4} + x^\sigma T^{1/4-\sigma} \log(T+1) \ll x^{3/4} + x^\sigma T^{-1/4},\end{aligned}$$

and completes the proof of (3.7).

We next prove an analogous inversion formula for the function $\Psi(x, y)$ in terms of the generating Dirichlet series

$$\zeta(s, y) = \prod_{p \leq y} \left(1 - p^{-s}\right)^{-1}.$$

For later use, we state the result in a somewhat more general form than is needed here.

LEMMA 3.3. *Uniformly for $x \ge y \ge 2$, $0 < \sigma \le 2$ and $T \ge 1$, we have*

$$\Psi(x,y) = \frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT} \zeta(s,y)\frac{x^s}{s}\mathrm{d}s + R \tag{3.8}$$

with

$$R \ll 1 + \frac{x^\sigma}{\sqrt{T}}\zeta(\sigma,y) + \min\Big(\frac{x}{\sqrt{T}}, \frac{x^\sigma}{\sqrt{T}}\int_{-\sqrt{T}}^{\sqrt{T}} |\zeta(\sigma+i\tau,y)|\mathrm{d}\tau\Big). \tag{3.9}$$

Proof. Expanding $\zeta(s,y)$ into a Dirichlet series and using the relation

$$\frac{1}{2\pi i}\int_{\sigma-iT}^{\sigma+iT}\Big(\frac{x}{n}\Big)^s\frac{\mathrm{d}s}{s} - \delta\Big(\frac{x}{n}\Big) \ll \Big(\frac{x}{n}\Big)^\sigma \min\Big(1, \frac{1}{T|\log(x/n)|}\Big),$$

where

$$\delta(t) = \begin{cases} 0 & \text{if } t < 1, \\ \frac{1}{2} & \text{if } t = 1, \\ 1 & \text{if } t > 1, \end{cases}$$

(cf. Titchmarsh and Heath-Brown (1986), p. 61) gives (3.8) with

$$R \ll 1 + x^\sigma \sum_{\substack{n \ge 1 \\ P(n) \le y}} \frac{1}{n^\sigma}\min\Big(1, \frac{1}{T|\log(x/n)|}\Big). \tag{3.10}$$

The contribution of the terms with $|\log(x/n)| > 1/\sqrt{T}$ to the series in (3.10) is bounded by $\ll x^\sigma\zeta(\sigma,y)/\sqrt{T}$, and hence by the right-hand side of (3.9). It therefore suffices to estimate the expression

$$R_1 = x^\sigma \sum_{\substack{|\log(x/n)| \le 1/\sqrt{T} \\ P(n) \le y}} \frac{1}{n^\sigma}.$$

We have trivially

$$R_1 \ll \sum_{|\log(x/n)| \le 1/\sqrt{T}} 1 \ll 1 + \frac{x}{\sqrt{T}}.$$

On the other hand, setting $w(t) = (1/2\pi)\big(\sin(t/2)/(t/2)\big)^2$ and noting that

$$\widehat{w}(\tau) = \int_{-\infty}^{\infty} w(t)e^{it\tau}\mathrm{d}t = \begin{cases} 1 - |\tau| & (|\tau| \le 1), \\ 0 & (|\tau| > 1), \end{cases} \tag{3.11}$$

we obtain

$$R_1 \ll x^\sigma \sum_{\substack{n\geq 1\\ P(n)\leq y}} \frac{1}{n^\sigma} w\big(\sqrt{T}\log(x/n)\big) = \frac{x^\sigma}{2\pi} \sum_{\substack{n\geq 1\\ P(n)\leq y}} \frac{1}{n^\sigma} \int_{-\infty}^{\infty} \Big(\frac{x}{n}\Big)^{i\tau\sqrt{T}} \widehat{w}(\tau)\mathrm{d}\tau$$

$$= \frac{x^\sigma}{2\pi\sqrt{T}} \int_{-\infty}^{\infty} x^{i\tau}\zeta(\sigma+i\tau,y)\widehat{w}\Big(\frac{\tau}{\sqrt{T}}\Big)\mathrm{d}\tau \ll \frac{x^\sigma}{\sqrt{T}} \int_{-\sqrt{T}}^{\sqrt{T}} |\zeta(\sigma+i\tau,y)|\mathrm{d}\tau,$$

since by (3.11), $\widehat{w}(\tau/\sqrt{T}) = 0$ for $|\tau| \geq \sqrt{T}$. The required estimate now follows. The next lemma shows that the function $F(s,y)$ closely approximates the function $\zeta(s,y)$.

LEMMA 3.4. *For any given $\epsilon > 0$ the estimate*

$$\zeta(s,y) = F(s,y)\Big(1 + O\Big(\frac{1}{L_\epsilon}\Big)\Big) \tag{3.12}$$

holds uniformly in the range

$$y \geq y_0(\epsilon), \qquad 1-(\log y)^{-2/5-\epsilon} \leq \sigma \leq 2, \qquad |\tau| \leq L_\epsilon, \tag{3.13}$$

where $L_\epsilon = L_\epsilon(y) = \exp\{(\log y)^{3/5-\epsilon}\}$.

Proof. We first note that, in the range (3.13),

$$-\frac{\zeta'}{\zeta}(s,y) = \sum_{n\leq y} \frac{\Lambda(n)}{n^s} + O\Big(y^{1/2-\sigma}\Big), \tag{3.14}$$

since

$$-\frac{\zeta'}{\zeta}(s,y) - \sum_{n\leq y}\frac{\Lambda(n)}{n^s} = \sum_{p\leq y}\sum_{\substack{p^m>y\\ m>1}} \frac{\log p}{p^{sm}}$$

$$\ll \sum_{p\leq\sqrt{y}} \frac{\log p}{y^\sigma} + \sum_{\sqrt{y}<p\leq y} \frac{\log p}{p^{2\sigma}} \ll y^{1/2-\sigma}.$$

The sum on the right of (3.14) is essentially a prime number sum which can be evaluated using complex integration and Vinogradov's zero-free region in the same way as in the analytic proof of the prime number theorem, giving the estimate

$$\sum_{n\leq y} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'}{\zeta}(s) + \frac{y^{1-s}}{1-s} + O\Big(\frac{1}{L_\epsilon^2}\Big) \tag{3.15}$$

for the range (3.13) — cf. Tenenbaum (1990a), p. 419. From (3.14), (3.15), Lemma 2.6, and the definition of $F(s,y)$, it follows that

$$\begin{aligned}\frac{\zeta'}{\zeta}(s,y) &= \frac{\zeta'}{\zeta}(s) - \frac{y^{1-s}}{1-s} + O\Big(\frac{1}{L_\epsilon^2}\Big)\\ &= \frac{\zeta'}{\zeta}(s) - E'\big(-(s-1)\log y\big)\log y + \frac{1}{s-1} + O\Big(\frac{1}{L_\epsilon^2}\Big)\\ &= \frac{F'}{F}(s,y) + O\Big(\frac{1}{L_\epsilon^2}\Big).\end{aligned}$$

Integrating this relation over the straight line path from s to 1 (which is contained in the range (3.13) if s is in this range), we obtain

$$\frac{\zeta(s,y)}{\zeta(1,y)} = \frac{F(s,y)}{F(1,y)}\exp\Big\{O\Big(\frac{1+|s|}{L_\epsilon^2}\Big)\Big\} = \frac{F(s,y)}{F(1,y)}\Big(1+O\Big(\frac{1}{L_\epsilon}\Big)\Big).$$

The result now follows, since

$$F(1,y) = (\log y)\,\widehat{\rho}(0)\lim_{s\to 1}\zeta(s)(s-1) = (\log y)\,\widehat{\rho}(0) = e^\gamma \log y$$

by Lemma 2.6 and the estimate

$$\zeta(1,y) = \prod_{p\le y}\Big(1-\frac{1}{p}\Big)^{-1} = e^\gamma(\log y)\Big(1+O\Big(\frac{1}{L_\epsilon}\Big)\Big),$$

which itself follows from a strong form of Mertens' theorem.

Proof of Theorem 1.8. Let $\epsilon > 0$ and $x \ge y$ with $1 \le u \le L_\epsilon$ be given. We may assume that y is sufficiently large in terms of ϵ and that u is strictly greater than 1. Set

$$\sigma = 1 - \frac{\xi(u)}{\log y}, \quad T = L_{\epsilon/2}.$$

By Lemma 2.2 we then have in the range $y \ge y_0(\epsilon)$, $1 < u \le L_\epsilon$,

$$1 > \sigma = 1 - \frac{\log(u+1)}{\log y} + O\Big(\frac{\log_2(u+1)}{\log y}\Big) \ge 1 - (\log y)^{-2/5-\epsilon/2}$$

and $T \ge u^5$. The hypotheses of Lemmas 3.2 and 3.3, and those of Lemma 3.4 with $s = \sigma + i\tau$, $|\tau| \le T$, and $\epsilon/2$ in place of ϵ, are therefore satisfied, and we obtain

$$\Psi(x,y) = \Lambda(x,y) + O\Big(\sum_{i=1}^{3} R_i\Big)$$

with

$$R_1 = x^\sigma T^{-1/4} + x^{3/4},$$

$$R_2 = 1 + \frac{x^\sigma}{\sqrt{T}}\zeta(\sigma,y) + \min\Big(\frac{x}{\sqrt{T}}, \frac{x^\sigma}{\sqrt{T}}\int_{-\sqrt{T}}^{\sqrt{T}} |\zeta(\sigma+i\tau,y)|\,\mathrm{d}\tau\Big),$$

$$R_3 = \frac{x^\sigma}{T}\int_{-T}^{T} |\zeta(\sigma+i\tau,y)|\frac{\mathrm{d}\tau}{|\sigma+i\tau|}.$$

To prove the desired estimate (1.28), it thus remains to show that each of the terms R_i satisfies

$$R_i \ll \Lambda(x,y)/L_\epsilon. \tag{3.16}$$

Since by Theorem 2.1, Lemma 2.2, and Lemma 3.1,

$$x^\sigma = xe^{-u\xi} \asymp x\rho(u)\frac{e^{-E(\xi)}}{\sqrt{\xi'(u)}} \asymp \Lambda(x,y)e^{-E(\xi)}\sqrt{u},$$

(3.16) is equivalent to

$$R_i \ll \frac{x^\sigma e^{E(\xi)}}{L_\epsilon\sqrt{u}}. \tag{3.17}$$

To prove (3.17), we first observe that

$$\begin{aligned}\frac{e^{E(\xi)}}{\sqrt{u}} &= \exp\Big\{\int_0^\xi \frac{e^t-1}{t}\mathrm{d}t - \tfrac{1}{2}\log u\Big\}\\ &= \exp\Big\{\Big(1+O\Big(\frac{1}{\xi+1}\Big)\Big)\frac{e^\xi-1}{\xi} - \tfrac{1}{2}\log u\Big\}\\ &= \exp\Big\{\Big(1+O\Big(\frac{1}{\log(u+1)}\Big)\Big)u\Big\} \gg e^{u/2}.\end{aligned} \tag{3.18}$$

Since $T^{1/4} = L_{\epsilon/2}^{1/4} \gg L_\epsilon$ and $1 \le x^{3/4} \ll x^\sigma/L_\epsilon$, the term R_1 and the first term in the definition of R_2 satisfy the bound in (3.17). Moreover, by Lemma 3.4, Lemma 2.6, and the definition of σ and T we have

$$\zeta(\sigma,y) \asymp |F(\sigma,y)| \asymp |\zeta(\sigma)(1-\sigma)(\log y)\widehat{\rho}(-\xi)| \asymp (\log y)e^{E(\xi)},$$

which implies (3.17) for the second term in the definition of R_2 in view of the bound

$$u \le L_\epsilon \ll_\delta L_{\epsilon/2}^\delta = T^\delta \quad (\delta > 0). \tag{3.19}$$

Since $|\zeta(\sigma+i\tau,y)| \le \zeta(\sigma,y)$, the same estimates show that the contributions to R_2 and R_3 arising from the range $|\tau| \le u^5$ of the integrals in these terms also satisfy (3.17). The remaining parts of R_2 and R_3 are bounded by

$$R_2' = \min\left(\frac{x}{\sqrt{T}}, 2x^\sigma M\right), \qquad \text{and} \qquad R_3' = 2M\frac{x^\sigma \log(T+1)}{T},$$

respectively, where

$$M = \max_{u^5 \le |\tau| \le T} |\zeta(\sigma+i\tau,y)|.$$

To estimate the quantity M, we note that, by Lemmas 3.4 and 2.6, we have in the range $u^5 \le |\tau| \le T$

$$|\zeta(\sigma+i\tau,y)| \asymp |F(\sigma+i\tau,y)| \asymp |\zeta(\sigma+i\tau)|\log y.$$

The definition of σ and T implies that $(1-\sigma) \gg (\log T)^{-2/3-\eta}$ with a suitable $\eta = \eta(\epsilon) > 0$. By standard bounds for the zeta function in Vinogradov's zero-free region we therefore have $|\zeta(\sigma+i\tau)| \ll \log(T+1) \ll \log y$ for $1 \le |\tau| \le T$. It follows that $M \ll (\log y)^2$. This implies the bound (3.17) for R_3'. By (3.18) the same bound holds for R_2' if $u > 3\log L_\epsilon$. For $1 \le u \le 3\log L_\epsilon$ we have

$$\begin{aligned}\frac{x}{\sqrt{T}} = x^\sigma \frac{e^{u\xi}}{\sqrt{L_{\epsilon/2}}} &= x^\sigma \exp\left\{-\tfrac{1}{2}(\log y)^{3/5-\epsilon/2} + O\Big((\log y)^{3/5-\epsilon}\log_2 y\Big)\right\} \\ &\ll x^\sigma \exp\left\{-\tfrac{1}{3}(\log y)^{3/5-\epsilon/2}\right\} \ll \frac{x^\sigma}{L_\epsilon},\end{aligned}$$

so that (3.17) holds in this case as well. This completes the proof of Theorem 1.8.

4. $\Psi(x,y)$ for large u

In this section we will prove Theorem 1.6 for the range $u \ge (\log y)^4$. In the complementary range $u < (\log y)^4$, the result can be deduced from Theorem 1.8, but the argument is somewhat technical and we shall not present it here. Given $x \ge y \ge 2$, we let $\alpha = \alpha(x,y)$ be defined as in the theorem, we write as usual $u = \log x/\log y$ and set

$$\overline{u} = \min\big(u, y/\log y\big).$$

LEMMA 4.1. *We have*

$$\alpha \asymp \frac{y}{u \log^2 y} \qquad (y \geq 2,\ u \geq y/\log y), \tag{4.1}$$

$$\frac{1}{\log y} \ll \alpha \leq 1 + O\Big(\frac{1}{\log y}\Big) \qquad (y \geq 2,\ 1 \leq u \leq y/\log y), \tag{4.2}$$

$$\frac{y^{1-\alpha} - 1}{(1-\alpha)\log y} \asymp \overline{u} \qquad (y \geq 2,\ u \geq 1), \tag{4.3}$$

where the expression on the left of (4.3) is to be interpreted as 1 *if* $\alpha = 1$.

Proof. By the definition of α we have

$$u \log y = \sum_{p \leq y} \frac{\log p}{p^\alpha - 1} \leq \sum_{p \leq y} \frac{\log p}{\alpha \log p} \asymp \frac{y}{\alpha \log y}. \tag{4.4}$$

For $u \geq y/\log y$ this implies $0 < \alpha \ll 1/\log y$, so that $p^\alpha - 1 \asymp \alpha \log p$ for $p \leq y$. The two middle terms in (4.4) are therefore of the same order of magnitude in the range $u \geq y/\log y$. This proves (4.1).

The lower bound in (4.2) follows from the fact that $\alpha = \alpha(y^u, y)$ is a non-increasing function of u and by (4.1) satisfies $\alpha \gg 1/\log y$ for $u = y/\log y$. The upper bound follows from the inequality

$$\sum_{p \leq y} \frac{\log p}{p^{1+c/\log y} - 1} < \log y \leq u \log y,$$

which is valid for all $y \geq 2$ and $u \geq 1$ with a sufficiently large constant c.

It remains to prove (4.3). For $u > y/\log y$, the right-hand side of (4.3) equals $y/\log y$, and by (4.1) the left-hand side is of the same order of magnitude. Thus (4.3) holds in this case. To deal with the remaining range $1 \leq u \leq y/\log y$, we use the estimate

$$u \log y = \sum_{p \leq y} \frac{\log p}{p^\alpha - 1} \asymp \int_{3/2}^{y} \frac{dt}{t^\alpha - 1},$$

which follows by partial summation and Chebyshev's prime number bounds. In the range $1 \leq u \leq y/\log y$, we have $\alpha \gg 1/\log y$ by (4.2), and it is easily checked that in this case

$$\int_{3/2}^{y} \frac{dt}{t^\alpha - 1} \asymp \int_1^y \frac{dt}{t^\alpha} = \frac{y^{1-\alpha} - 1}{1 - \alpha}.$$

Hence (4.3) remains valid for $1 \le u \le y/\log y$. We remark that the above arguments can be refined to show that the estimate

$$\alpha = \frac{\log(1+y/\log x)}{\log y}\left\{1+O\Big(\frac{\log_2(y+1)}{\log y}\Big)\right\} \tag{4.5}$$

holds uniformly in $x \ge y \ge 2$; see Theorem 2 of Hildebrand & Tenenbaum (1986).

The next lemma gives estimates for the functions $\phi_k(s,y)$ defined in Theorem 1.6.

LEMMA 4.2. *For any fixed positive integer k and $x \ge y \ge 2$ we have*

$$0 < (-1)^k\phi_k(\alpha,y) \asymp_k (u\log y)^k\overline{u}^{1-k}. \tag{4.6}$$

Proof. Setting $f(t) = 1/(e^t-1)$, we have

$$(-1)^k\phi_k(\alpha,y) = (-1)^{k-1}\sum_{p\le y} f^{(k-1)}(\alpha\log p)\log^k p. \tag{4.7}$$

Since for any integer $\ell \ge 0$ and real $t > 0$,

$$\begin{aligned}(-1)^\ell f^{(\ell)}(t) &= (-1)^\ell\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\sum_{n\ge1}e^{-nt} = \sum_{n\ge1}n^\ell e^{-nt}\\ &\asymp_\ell e^{-t}\sum_{n\ge0}\binom{n+\ell}{\ell}e^{-nt} = \frac{e^{-t}}{(1-e^{-t})^{\ell+1}},\end{aligned}$$

it follows that $(-1)^k\phi_k(\alpha,y)$ is positive and satisfies

$$(-1)^k\phi_k(\alpha,y) \asymp_k \sum_{p\le y}\frac{\log^k p}{(1-p^{-\alpha})^k p^\alpha}.$$

Using partial summation, the prime number theorem, and the bounds $0 < \alpha \le 1 + O(1/\log y)$ from Lemma 4.1, it is straightforward to replace the sum over p by an integral, so that

$$(-1)^k\phi_k(\alpha,y) \asymp_k \int_{3/2}^{y}\frac{(\log t)^{k-1}}{(1-t^{-\alpha})^k t^\alpha}\,\mathrm{d}t. \tag{4.8}$$

If $u > y/\log y$, then $\alpha \ll 1/\log y$ by Lemma 4.1, and therefore $t^\alpha \asymp 1$ and $1 - t^{-\alpha} \asymp \alpha\log t$ for $\frac{3}{2} \le t \le y$. The right-hand side of (4.8) then becomes

$$\asymp_k \frac{1}{\alpha^k}\int_{3/2}^{y}\frac{\mathrm{d}t}{\log t} \asymp_k \Big(\frac{u\log^2 y}{y}\Big)^k\int_{3/2}^{y}\frac{\mathrm{d}t}{\log t} \asymp_k (u\log y)^k\overline{u}^{1-k},$$

by another application of Lemma 4.1, and (4.6) follows for the range $u > y/\log y$. In the remaining range $1 \le u \le y/\log y$ we have $\alpha \gg 1/\log y$ and therefore

$$\frac{(\log t)^{k-1}}{(1-t^{-\alpha})^k t^\alpha} \asymp_k \frac{(\log y)^{k-1}}{t^\alpha}$$

for $\sqrt{y} \le t \le y$. Using the bound $\alpha \le 1 + O(1/\log y)$ from Lemma 4.1, it is easily seen that we may replace the range of integration in (4.8) by the interval $[\sqrt{y}, y]$ without changing the order of magnitude of the right-hand side of (4.8). We thus obtain

$$\begin{aligned}(-1)^{k-1}\phi(\alpha,y) &\asymp_k (\log y)^{k-1}\int_{\sqrt{y}}^{y}\frac{dt}{t^\alpha} \asymp (\log y)^{k-1}\int_1^y \frac{dt}{t^\alpha}\\ &= (\log y)^{k-1}\frac{y^{1-\alpha}-1}{1-\alpha} \asymp (\log y)^{k-1}\overline{u}\log y = (u\log y)^k\overline{u}^{1-k},\end{aligned}$$

using the last part of Lemma 4.1. This proves (4.6) for $1 \le u \le y/\log y$.

LEMMA 4.3. *For any fixed $\epsilon > 0$ and uniformly for $y \ge 2$ and $u \ge (\log y)^4$, we have*

(4.9)

$$\frac{|\zeta(\alpha+i\tau,y)|}{\zeta(\alpha,y)} \ll \begin{cases} e^{-c\tau^2\phi_2(\alpha,y)} + \left(1+\dfrac{\tau^2}{\alpha^2}\right)^{-c\pi(y)} & (|\tau| \le 1/\log y),\\ \exp\left\{-c\log^2 y\right\} & (1/\log y < |\tau| \le T_\epsilon(y)),\end{cases}$$

where $T_\epsilon(y) = \exp\left\{(\log y)^{3/2-\epsilon}\right\}$ and c is a positive constant.

Proof. A simple computation gives

$$\frac{|\zeta(\alpha+i\tau,y)|}{\zeta(\alpha,y)} = \prod_{p\le y}\left|\frac{1-p^{-\alpha}}{1-p^{-\alpha-i\tau}}\right| = \prod_{p\le y}\left(1+\lambda_p\right)^{-1/2}$$

with

$$\lambda_p = \frac{2\big(1-\cos(\tau\log p)\big)}{p^\alpha(1-p^{-\alpha})^2}.$$

Thus, to prove (4.9) it suffices to show that in the ranges $|\tau| \le 1/\log y$ and $1/\log y < |\tau| \le T_\epsilon(y)$ we have, respectively,

$$(4.10)\quad \tfrac{1}{2}\sum_{p\le y}\log(1+\lambda_p) \ge c\min\left\{\tau^2\phi_2(\alpha,y),\ \pi(y)\log\left(1+\frac{\tau^2}{\alpha^2}\right)\right\} + O(1)$$

and

$$\tfrac{1}{2}\sum_{p\le y}\log(1+\lambda_p)\ge c\log^2 y+O(1). \tag{4.11}$$

Suppose first that $|\tau| \le 1/\log y$. Using the elementary inequality $1-\cos x \ge \frac{1}{4}x^2$ ($|x|\le 1$), we obtain

$$\lambda_p \ge \frac{\tau^2\log^2 p}{2p^\alpha(1-p^{-\alpha})^2} = \tau^2\mu_p \quad \text{(say)}.$$

If $|\tau|\le\alpha$ (and $|\tau|\le 1/\log y$) then we have

$$\tau^2\mu_p \le \frac{\tau^2\log^2 p}{2p^\alpha(1-p^{-|\tau|})^2} \ll \frac{1}{p^\alpha} \ll 1$$

for $p\le y$, and hence

$$\begin{aligned}\sum_{p\le y}\log(1+\lambda_p) &\ge \sum_{p\le y}\log(1+\tau^2\mu_p) \gg \tau^2\sum_{p\le y}\mu_p\\ &\asymp \tau^2\sum_{p\le y}\frac{\log^2 p}{p^\alpha(1-p^{-\alpha})^2} = \tau^2\phi_2(\alpha,y),\end{aligned}$$

which proves (4.10) in this case. In the remaining case $\alpha<|\tau|\le 1/\log y$, we have $p^\alpha \asymp 1$ and

$$\lambda_p \ge \frac{\tau^2\log^2 p}{2p^\alpha(1-p^{-\alpha})^2} \asymp \frac{\tau^2}{\alpha^2}$$

for $p\le y$, which again implies (4.10).

Assume now that $1/\log y<|\tau|\le T_\epsilon(y)$. For bounded values of y, (4.11) holds trivially, and we may therefore suppose that y is sufficiently large in terms of ϵ. Our starting point is the estimate

(4.12)

$$\begin{aligned}\sum_{p\le y}\log(1+\lambda_p) &\ge \sum_{p\le y}\log\Big(1+\frac{2(1-\cos(\tau\log p))}{p^\alpha}\Big)\\ &\gg \sum_{p\le y}\frac{1-\cos(\tau\log p)}{p^\alpha} \ge \sum_{p\le y}\frac{(1-\cos(\tau\log p))\log p}{p^\alpha\log y}\\ &\ge \frac{1}{\log y}\Big(S(\alpha,y)-\operatorname{Re}S(\alpha+i\tau,y)\Big),\end{aligned}$$

where

$$S(s,y)=\sum_{p\le y}\frac{\log p}{p^s}.$$

A routine argument using complex integration and Vinogradov's zero free region yields that

$$\begin{aligned}S(s,y)&=\int_1^y\frac{dt}{t^s}+O_A\Big(\frac{1}{\log^A y}\int_1^y\frac{dt}{t^\sigma}\Big)+O(1)\\&=\frac{y^{1-s}}{1-s}+O\Big(\frac{1}{1-\sigma}\Big)+O_A\Big(\frac{y^{1-\sigma}}{(1-\sigma)\log^A y}\Big)\end{aligned}$$

holds uniformly in $y\geq y_0(\epsilon)$, $0<\sigma<1$ and $|\tau|\leq T_\epsilon(y)$, for any fixed constant $A>0$. If follows that

(4.13)

$$S(\alpha,y)-\operatorname{Re}S(\alpha+i\tau,y)\geq\eta\frac{y^{1-\alpha}}{1-\alpha}+O\Big(\frac{1}{1-\alpha}\Big)+O\Big(\frac{y^{1-\alpha}}{(1-\alpha)(\log y)^4}\Big),$$

where

$$\eta=1-\frac{1-\alpha}{|1-\alpha-i\tau|}=1-\Big(1+\frac{\tau^2}{(1-\alpha)^2}\Big)^{-1/2}.$$

Now note that $\eta\gg1/\log^2 y$ for $|\tau|\geq1/\log y$ and, by Lemma 4.1 and our assumption $u\geq(\log y)^4$, $1-\alpha\gg1/\log y$ and

$$\frac{y^{1-\alpha}}{1-\alpha}\asymp\frac{y^{1-\alpha}-1}{1-\alpha}\asymp\overline{u}\log y\gg\log^5 y.$$

Hence the right-hand side of (4.13) is bounded by $\gg\log^3 y$ for sufficiently large y, and (4.11) follows.

Proof of Theorem 1.6. Let $x\geq y\geq2$ be given, and suppose that $u=\log x/\log y\geq(\log y)^4$. By Lemmas 4.1 and 4.2, we have

$$\alpha\sqrt{\phi_2(\alpha,y)}\ll\sqrt{\phi_2(\alpha,y)}\ll\sqrt{u\log^2 y}\ll\sqrt{y\log y}\quad\text{for}\quad u\leq y/\log y,$$

$$\alpha\sqrt{\phi_2(\alpha,y)}\ll\frac{y}{u\log^2 y}\sqrt{\frac{u^2\log^3 y}{y}}=\sqrt{y/\log y}\quad\text{for}\quad u>y/\log y.$$

Hence, if

$$M=\frac{x^\alpha\zeta(\alpha,y)}{\alpha\sqrt{2\pi\phi_2(\alpha,y)}}$$

denotes the main term in the estimate of Theorem 1.6, then

$$M\gg\frac{x^\alpha\zeta(\alpha,y)}{\sqrt{y\log y}}.\tag{4.14}$$

We now apply Lemma 3.3 with $\sigma = \alpha$ and $T = \exp\{(\log y)^{5/4}\}$, getting

$$\Psi(x,y) = \frac{1}{2\pi i}\int_{\alpha-iT}^{\alpha+iT} \zeta(s,y)\frac{x^s}{s}\,\mathrm{d}s + R \tag{4.15}$$

with

$$R \ll 1 + x^\alpha\zeta(\alpha,y)\left\{\frac{1}{\sqrt{T}} + \frac{1}{\sqrt{T}}\int_{-\sqrt{T}}^{\sqrt{T}} \frac{|\zeta(\alpha+i\tau,y)|}{\zeta(\alpha,y)}\,\mathrm{d}\tau\right\}. \tag{4.16}$$

Using the trivial bound $|\zeta(\alpha+i\tau,y)| \le \zeta(\alpha,y)$ for $|\tau| \le 1/\log y$ and the estimate of Lemma 4.3 with $\epsilon = 1/4$ for $1/\log y \le |\tau| \le T$, we see that the expression in parentheses is bounded by

$$\ll \frac{1}{\sqrt{T}} + \frac{1}{\sqrt{T}}\int_{1/\log y}^{\sqrt{T}} \frac{|\zeta(\alpha+i\tau,y)|}{\zeta(\alpha,y)}\,\mathrm{d}\tau \ll \frac{1}{\sqrt{T}} + \exp\{-c\log^2 y\} \ll \frac{1}{y^2}.$$

In view of (4.14), this shows that the contribution of the second term on the right of (4.16) is of order $\ll M/y \ll M/\overline{u}$. The same bound holds for the first term since $\zeta(\alpha,y) \ge 1$ and, by Lemma 4.2 and the bound $u \ge (\log y)^4$,

$$\begin{aligned}\log\left(x^\alpha\zeta(\alpha,y)\right) &\ge \log x^\alpha = u\alpha\log y \\ &\gg u(\log y)\min\left(\frac{1}{\log y}, \frac{y}{u\log^2 y}\right) \gg (\log y)^4.\end{aligned}$$

Furthermore, another application of Lemma 4.3 shows that the contribution of the range $1/\log y \le |\tau| \le T$ to the integral in (4.15) is also of order $\ll M/\overline{u}$. Thus we have

$$\Psi(x,y) = \frac{1}{2\pi i}\int_{\alpha-i/\log y}^{\alpha+i/\log y} \zeta(s,y)\frac{x^s}{s}\,\mathrm{d}s + O\left(\frac{M}{\overline{u}}\right), \tag{4.17}$$

and it remains to estimate the last integral. Let $\delta = \overline{u}^{2/3}/(u\log y)$. Applying Lemma 4.3, we see that the contribution of the range $\delta \le |\tau| \le 1/\log y$ to the integral in (4.17) is bounded by

(4.18)

$$\ll x^\alpha\zeta(\alpha,y)\left\{\int_\delta^{1/\log y} \frac{e^{-c\tau^2\phi_2(\alpha,y)}}{|\alpha+i\tau|}\,\mathrm{d}\tau + \int_\delta^{1/\log y}\left(1+\frac{\tau^2}{\alpha^2}\right)^{-c\pi(y)}\frac{\mathrm{d}\tau}{|\alpha+i\tau|}\right\}.$$

The first of the two integrals here is of order

$$\ll \frac{1}{\alpha\sqrt{\phi_2(\alpha,y)}}\int_{\delta\sqrt{\phi_2(\alpha,y)}}^{\infty} e^{-ct^2}\,\mathrm{d}t \ll \frac{e^{-c\delta\sqrt{\phi_2(\alpha,y)}}}{\alpha\sqrt{\phi_2(\alpha,y)}}.$$

Since, by Lemma 4.2,

$$\delta\sqrt{\phi_2(\alpha,y)} \asymp \frac{\overline{u}^{2/3}}{(u\log y)}\sqrt{\frac{u^2\log^2 y}{\overline{u}}} = \overline{u}^{1/6} \tag{4.19}$$

the contribution of this integral to (4.18) is of order $\ll M\exp(-c_1\overline{u}^{1/6})$ for some constant $c_1 > 0$ and hence acceptable as error term. The second integral in (4.18) is at most

$$\int_{\delta/\alpha}^{\infty}\left(1+t^2\right)^{-c\pi(y)}dt \ll \exp\left\{-c_2\min\left(1,\frac{\delta^2}{\alpha^2}\right)\frac{y}{\log y}\right\}$$

with a suitable constant $c_2 > 0$. Since

$$\frac{\delta}{\alpha} = \frac{\overline{u}^{2/3}}{\alpha u\log y} \gg \left(\frac{y}{\log y}\right)^{-1/3}$$

by Lemma 4.1, this integral is of order $\ll e^{-\sqrt{y}}$ and the corresponding term is again negligible.

In the remaining range $|\tau| \le \delta$, we expand the logarithm of the integrand in (4.17) into a Taylor series about $\tau = 0$. Since by Lemma 4.1

$$\alpha \gg \overline{u}/(u\log y) \gg \delta,$$

we have in this range

$$\frac{1}{\alpha+i\tau} = \frac{1}{\alpha}\left\{1 - i\frac{\tau}{\alpha} + O\left(\frac{\tau^2}{\alpha^2}\right)\right\}.$$

Setting $a_k = \phi_k(\alpha,y)$ and using the relation $a_1 = -\log x$ and the estimates

$$|a_k| \asymp (u\log y)^k\overline{u}^{1-k} \asymp a_2^{k/2}\overline{u}^{1-k/2} \asymp \delta^{-k}\overline{u}^{1-k/3} \ll \delta^{-k} \quad (k=3,4),$$

$$\frac{1}{\alpha^2} \ll \left(\frac{u\log y}{\overline{u}}\right)^2 \asymp \frac{a_2}{\overline{u}},$$

from Lemmas 4.1 and 4.2, we can write the integrand in (4.17) for $|\tau| \le \delta$ as

$$\begin{aligned}
&\frac{\zeta(\alpha,y)x^{\alpha+i\tau}}{\alpha+i\tau}\exp\left\{\sum_{k=1}^{3}\frac{i^k a_k\tau^k}{k!} + O\Big(|a_4|\tau^4\Big)\right\}\\
&= \frac{x^\alpha\zeta(\alpha,y)}{\alpha}e^{-a_2\frac{\tau^2}{2}}\left\{1-\frac{i\tau}{\alpha} - \frac{ia_3\tau^3}{3!} + O\left(\frac{\tau^2}{\alpha^2} + |a_3|^2\tau^6 + \left(\frac{|a_3|}{\alpha} + |a_4|\right)\tau^4\right)\right\}\\
&= \frac{x^\alpha\zeta(\alpha,y)}{\alpha}e^{-a_2\frac{\tau^2}{2}}\left\{1 - \frac{i\tau}{\alpha} - \frac{ia_3\tau^3}{3!} + O\left(\frac{1}{\overline{u}}\sum_{k=2}^{4}\left(\sqrt{a_2}\tau\right)^{2k}\right)\right\}.
\end{aligned}$$

Integrating this expression over the interval $(-\delta, \delta)$, we obtain as main term

$$\frac{x^\alpha \zeta(\alpha, y)}{\alpha} \int_{-\delta}^{\delta} e^{-a_2 \frac{t^2}{2}} d\tau = \frac{x^\alpha \zeta(\alpha, y)}{\alpha\sqrt{a_2}} \left\{ \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} \mathrm{d}t + O\Big(\int_{\delta\sqrt{a_2}}^{\infty} e^{-\frac{t^2}{2}} \mathrm{d}t \Big) \right\}$$
$$= \frac{x^\alpha \zeta(\alpha, y)\sqrt{2\pi}}{\alpha\sqrt{a_2}} \left\{ 1 + O\Big(\frac{1}{\overline{u}}\Big) \right\},$$

since $\delta\sqrt{a_2} \asymp \overline{u}^{1/6}$. The integrals involving the linear and cubic terms in τ vanish, and the contribution of the error terms to the integral is by a factor $\ll 1/\overline{u}$ smaller than the main term. Thus we have

$$\frac{1}{2\pi i} \int_{\alpha - i\delta}^{\alpha + i\delta} \zeta(s, y) \frac{x^s}{s} \mathrm{d}s = \frac{x^\alpha \zeta(\alpha, y)}{\alpha\sqrt{2\pi a_2}} \left\{ 1 + O\Big(\frac{1}{\overline{u}}\Big) \right\},$$

which completes the proof of Theorem 1.6.

5. Distribution in short intervals

The results discussed in the preceding sections give a fairly complete picture of the *global* distribution of integers without large prime factors. As one might expect, the *local* distribution of such integers, i.e. the distribution of integers free of prime factors $> y$ in an interval of the form $(x, x+z]$, is much less understood, and non-trivial results are known only for very limited ranges of the parameters x, y, and z. In this section, we shall survey the principal results, outline some of the methods of proof, and give a complete proof of one such result. As in the preceding sections we set, for given x and y, $u = \log x / \log y$.

We first consider the case when the parameter u is small. The Dickman–de Bruijn relation $\Psi(x, y) \sim x\rho(u)$ suggests that in this case a relation of the form $\Psi(x+z, y) - \Psi(x, y) \sim z\rho(u)$ should hold, at least when z is not too small. Such a result is given in the following theorem.

THEOREM 5.1. *For any fixed $\epsilon > 0$, uniformly in the range*

$$y \geq 2, \quad 1 \leq u \leq \exp\left\{(\log y)^{3/5-\epsilon}\right\}, \tag{5.1}$$

and for $xy^{-5/12} \leq z \leq x$, we have

$$\Psi(x+z, y) - \Psi(x, y) = z\rho(u)\Big\{1 + O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\}. \tag{5.2}$$

This result is proved in Hildebrand (1986a), using the same method as for the proof of Theorem 1.1. Note that the range (5.1) coincides with the range (1.10) of Theorem 1.1, which is the largest range in which the asymptotic relation $\Psi(x,y) \sim x\rho(u)$ is known to hold. The exponent $5/12$ in the lower bound $z \geq xy^{-5/12}$ stems from an application of the sharpest known form of the prime number theorem for short intervals. Recent work of Granville (1993b) indicates that, by using "almost all" type estimates for primes and almost-primes in short intervals, this exponent can be increased, possibly to $1-\epsilon$.

Since the global asymptotic relation $\Psi(x,y) \sim x\rho(u)$ is not known to hold beyond the range (5.1), an improvement of this range in Theorem 5.1 does not seem to be possible at present. However, one can obtain non-trivial results for large u if, instead of trying to approximate $\Psi(x+z,y)-\Psi(x,y)$ by a smooth function, one seeks an estimate for this quantity in terms of $\Psi(x,y)$. The following theorem, due to Hildebrand & Tenenbaum (1986), gives such an estimate.

THEOREM 5.2. *For any fixed $\epsilon > 0$ and uniformly for $x \geq y \geq 2$ and $1 \leq z \leq x$ we have*

$$\begin{aligned}\Psi(x+z,y)-\Psi(x,y) =& \frac{z\alpha}{x}\Psi(x,y)\Big(1+O\Big(\frac{z}{x}+\frac{1}{u}+\frac{\log y}{y}\Big)\Big)\\ &+O\Big(\Psi(x,y)R_\epsilon\Big),\end{aligned} \tag{5.3}$$

where $\alpha = \alpha(x,y)$ is defined as in Theorem 1.6, and

$$R_\epsilon = R_\epsilon(x,y) = \exp\big\{-(\log y)^{3/2-\epsilon}\big\} + (\log y)\exp\big\{-cu/(\log(u+2))^2\big\},$$

with a suitable positive constant c.

Using the estimate (1.26) for $\alpha(x,y)$, it is easy to see that the second error term in (5.3) is absorbed by the first error term if

$$(\log y)^{3/2} \leq u \leq \exp\big\{(\log y)^{3/2-\epsilon}\big\} \quad \text{and} \quad z \geq x\exp\big\{-(\log y)^{3/2-\epsilon}\big\}.$$

In this range, (5.3) gives an asymptotic formula for $\Psi(x+z,y)-\Psi(x,y)$, if y and x/z tend to infinity.

The proof of Theorem 5.2 is very similar to that of Theorem 1.6 given in the preceding section. One writes the left-hand side of (5.3) as a complex integral over the function $\zeta(s,y)((x+z)^s-x^s)/s$, where $\zeta(s,y)$ is defined as in Theorem 1.6, and evaluates the integral using the saddle point method in the same way as in the proof of Theorem 1.6. This leads to an asymptotic estimate with main term $(z/x)x^\alpha\zeta(\alpha,y)/\sqrt{2\pi\phi_2(\alpha,y)}$, which by Theorem 1.6 is equal to the main term on the right-hand side of (5.3), apart from a negligible error term.

The above theorems give very precise estimates for the quantity $\Psi(x+z,y)-\Psi(x,y)$, but only when z is relatively close to x. The range for z can be increased, if one is only interested in obtaining upper or lower bounds of the expected order of magnitude. Upper bounds for $\Psi(x+z,y)-\Psi(x,y)$ were given by Wolke (1971) and Hildebrand (1985a). The more difficult problem of obtaining non-trivial lower bounds was considered by Turk (1982), Friedlander (1985), and Friedlander & Lagarias (1987). In the last-mentioned paper the following result is proved.

THEOREM 5.3. *There exists a positive constant c such that, for any fixed $\alpha \in (0,1)$ and $\beta > 1-\alpha-c\alpha(1-\alpha)$ and for all sufficiently large x,*

$$\Psi(x+x^\beta, x^\alpha) - \Psi(x, x^\alpha) \gg_{\alpha,\beta} x^\beta. \tag{5.4}$$

Note that Theorem 5.1 yields (5.4) only for the smaller range $\beta \geq 1-\frac{5}{12}\alpha$. Even with the exponent $\frac{5}{12}$ replaced by $1-\epsilon$ (cf. the remarks following the statement of the theorem), this result would not be sufficient to imply Theorem 5.3.

The proof of Theorem 5.3 is based on the identity

$$\sum_{x<n\leq x+z} \log n = \sum_{p^m \leq x+z} \log p\left(\left[\frac{x+z}{p^m}\right] - \left[\frac{x}{p^m}\right]\right),$$

together with an analysis of the expression on the right-hand side using Fourier techniques and estimates for exponential sums over primes. Assuming that $\Psi(x+z,y)-\Psi(x,y)$ is small compared to z, it can be shown that the contribution of the primes $p \leq y$ to the right-hand side of the identity must be smaller than expected, which under the hypotheses of the theorem leads to a contradiction.

Friedlander & Lagarias also showed, by a different argument, that for any integer $r \geq 2$ and any fixed real number $\beta > 1-(2/r)+(1/r2^{r-1})$, the inequality $\Psi(x+x^\beta, x^{1/r}) - \Psi(x, x^{1/r}) > 0$ holds for all sufficiently large x. Thus, for example, for any $\epsilon > 0$ and $x \geq x_0(\epsilon)$ the interval $(x, x+x^{1/4+\epsilon}]$ contains an integer free of prime factors exceeding $\sqrt{x}$. For intervals of length $\geq x^{1/2+\epsilon}$, Balog (1987) obtained the following stronger result.

THEOREM 5.4. *For any fixed $\epsilon > 0$ and all real $x \geq x_0(\epsilon)$, the interval $(x, x+x^{1/2+\epsilon}]$ contains an integer free of prime factors exceeding x^ϵ.*

To prove this result, Balog considers the weighted sum

$$\sum_{x<\ell mn\le x+x^{1/2+\epsilon}} a_m a_n,$$

where $a_n = 1$ if $N < n \le 2N$ and n has no prime factors $> x^\epsilon$, and $a_n = 0$ otherwise, with $N = x^{1/2-\eta}$ for a sufficiently small $\eta = \eta(\epsilon)$. Using analytic methods, this sum can be asymptotically evaluated and shown to be positive. If η is sufficiently small then a positive contribution from a term ℓmn to this sum implies that ℓmn has no prime factors $> x^\epsilon$, and this yields the desired conclusion. Very recently, Harman (1991) obtained a quantitative refinement of this result, showing that the bound x^ϵ on the size of the prime factors can be replaced by $\exp\{(\log x)^{2/3+\epsilon}\}$.

The results mentioned so far require that either the interval length z or the bound y on the size of the prime factors is at least as large as a fixed power of x. If one only asks for estimates that hold for "almost all" x, then the ranges for y and z can be substantially improved. The first result in this direction is due to Friedlander (1984a) who proved that for any $\epsilon > 0$, any function $z = z(X)$ tending to infinity with X, and all $x \in [1, X]$ with the exception of a set of measure $o(X)$ the interval $(x, x+z]$ contains an integer n with $P(n) \le n^{1/2+\epsilon}$. Friedlander & Lagarias (1987) later proved the following results.

THEOREM 5.5. *For any fixed $\epsilon > 0$, $0 < \beta \le \alpha \le 1$, and for all sufficiently large X, the estimate*

$$\Psi(x + x^\beta, x^\alpha) - \Psi(x, x^\alpha) \ge \tfrac{1}{64}\beta\rho(1/\alpha)x^\beta \tag{5.5}$$

holds for all $x \in [1, X]$ with the exception of a set of measure bounded by $\ll_{\epsilon,\alpha,\beta} X\exp\{-(\log X)^{1/3-\epsilon}\}$.

THEOREM 5.6. *For any fixed $\epsilon > 0$, for all sufficiently large X, and for y and z satisfying*

$$\exp\{(\log X)^{5/6+\epsilon}\} \le y \le X, \qquad y\exp\{(\log X)^{1/6}\} \le z \le X \tag{5.6}$$

the estimate

$$\Psi(x+z, y) - \Psi(x, y) \gg \rho\Big(\frac{\log X}{\log y}\Big) z \tag{5.7}$$

holds for all $x \in [1, X]$ with the exception of a set of measure bounded by $\ll_\epsilon X\exp\{-\tfrac{1}{2}(\log X)^{1/6}\}$.

The range (5.6) is a consequence of Vinogradov's zero-free region for the zeta function and can be extended if one assumes a larger zero-free region. A result of this type has recently been given by Hafner (1993). In particular, Hafner shows that under the Riemann Hypothesis the conclusion of Theorem 5.6 holds for $L(X) \le y \le x$ and $\sqrt{L(X)} \le z \le X$, where $L(x) = \exp\{\sqrt{\log x \log_2 x}\}$.

Results of this type are of interest since the running times of certain factoring algorithms depend on the distribution of integers without large prime factors in short intervals. In particular, the elliptic curve factoring method (ECM) of Lenstra (1987) factors a large integer N in expected time $\ll \exp\{(1+o(1))\sqrt{\log N \log_2 N}\}$ under the assumption that the estimate

$$\Psi(x+z,y) - \Psi(x,y) = zu^{-(1+o(1))u}$$

holds when $z \asymp \sqrt{x}$ and $\log y \asymp \sqrt{\log x \log_2 x}$, as $x \to \infty$. While we are still very far from being able to prove such a result (or even showing that $\Psi(x+z,y) - \Psi(x,y)$ is positive under the same conditions), Pomerance (1987) showed that the above "almost all" type results can be used in Lenstra's analysis to obtain rigorous, though weaker, bounds for the expected running time of ECM. Specifically, Theorem 5.5 leads to an unconditional bound of the form $O\big(\exp\{(\log N)^{5/6+\epsilon}\}\big)$, and Theorem 5.6 gives on the Riemann Hypothesis the bound $O\big(\exp(\sqrt{c \log N \log_2 N})\big)$ for some constant c.

While the results of Friedlander & Lagarias and of Hafner give only lower bounds for $\Psi(x+z,y) - \Psi(x,y)$ of the expected order of magnitude, their method is in fact capable of yielding an asymptotic estimate of the same quality as that of Theorem 5.1. We conclude this section by proving a result of this type which refines Theorem 5.6, except for a slightly weaker bound on the measure of the exceptional set. Theorem 5.5 can also be sharpened to an asymptotic formula, but this would require a somewhat different approach which gives only rather poor estimates for the exceptional set.

THEOREM 5.7. *For any fixed $\epsilon > 0$, for all sufficiently large X, and for y, z satisfying (5.6), the estimate (5.2), i.e.*

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

holds for all $x \in [1,X]$ with the exception of a set of measure bounded by $\ll_\epsilon X \exp\{-(\log X)^{1/6-\epsilon}\}$.

As in Hafner's result, the range (5.6) can be extended if one assumes a larger zero-free region for the zeta function. Moreover, the estimate (5.2) can be sharpened to

$$\Psi(x+z,y)-\Psi(x,y)=\frac{z}{x}\Psi(x,y)\Big\{1+O\Big(\exp\big\{-(\log X)^{\epsilon/2}\big\}\Big)\Big\},$$

if one assumes, in addition, that $0<\epsilon\le\frac12$ and $z\le X\exp\big\{-(\log X)^{\epsilon}\big\}$ and relaxes the bound on the measure of the exceptional set to a quantity $\ll X\exp\big\{-(\log X)^{\epsilon/2}\big\}$.

The main step in the proof of the above theorems is contained in the following lemma, which is a slight generalization of similar results in Friedlander & Lagarias (1987) and Hafner (1993).

LEMMA 5.8. *Let $\epsilon>0$ be fixed. Let $X\ge 2$, $1\le M,V\le X$, and let $a(m)$ $(M<m\le 2M)$ be complex numbers with $|a(m)|\le 1$. Set*

$$w(n)=\sum_{mp=n,M<m\le 2M}a(m)\log p \qquad \text{and} \qquad S(x)=\sum_{x<n\le(1+1/V)x}w(n).$$

Then we have

$$\begin{aligned}\int_X^{2X}&\Big|S(x)-\frac{x}{V}A(1)\Big|^2\mathrm{d}x\\ &\ll X\mathcal{L}^4\Big(\frac{X}{V}\Big)^2\Big\{\Big(\frac{M}{X}\Big)^{\mathcal{L}^{-2/3-\epsilon}}\frac{1}{M}\sum_{M<m\le 2M}|a(m)|^2+\Big(\frac{V}{M}\Big)^2\Big\},\end{aligned}\tag{5.8}$$

where $A(s)=\sum_{M<m\le 2M}a(m)m^{-s}$ and $\mathcal{L}=\log X$.

Proof. We may assume that X is sufficiently large in terms of ϵ, for otherwise the result holds trivially. The generating Dirichlet series for $w(n)$ is given by

$$W(s)=\sum_{n\ge1}\frac{w(n)}{n^s}=\Big(-\frac{\zeta'}{\zeta}(s)-G(s)\Big)A(s),$$

where $A(s)$ is defined as in the lemma and

$$G(s)=\sum_{p,m\ge2}\frac{\log p}{p^{ms}}=\sum_p\frac{\log p}{p^{2s}(1-p^{-s})}.$$

Since $|w(n)|\le\sum_{p|n}\log p\le\log n$, we have, by a standard application of Perron's formula (cf. Lemma 3.19 in Titchmarsh (1986)), for any non-integral value of $x\in[X,2X]$,

$$S(x)=\frac{1}{2\pi i}\int_{\kappa-iM}^{\kappa+iM}W(s)k(s)x^s\mathrm{d}s+O\Big(\frac{X}{M}\mathcal{L}^2\Big),\tag{5.9}$$

where $\kappa = 1 + 1/\mathcal{L}$ and

$$k(s) = \frac{(1+1/V)^s - 1}{s}.$$

We put $\eta = 1 - (\log X)^{-2/3-\epsilon}$ and shift the path of integration to the line $\operatorname{Re} s = \eta$. If X is sufficiently large in terms of ϵ, then Vinogradov's zero-free region for the zeta function implies that the integrand in (5.9) is analytic in the rectangle with vertices $\eta \pm iM$, $\kappa \pm iM$, with the exception of a pole at $s = 1$ with residue $A(1)k(1)x = (x/V)A(1)$. Moreover, on the horizontal segments $[\eta \pm iM, \kappa \pm iM]$ we have the bounds

$$|(\zeta'/\zeta)(s)| \ll \log(M+1) \ll \mathcal{L}, \quad |G(s)| \ll 1, \quad |k(s)| \ll 1/|s| \ll 1/M,$$

and hence

$$\left|W(s)k(s)x^s\right| \ll \frac{\mathcal{L}}{M} \sum_{M<m\le 2M} \frac{|a(m)|}{m^\sigma} x^\sigma \ll \left(\frac{X}{M}\right)^\sigma \mathcal{L} \ll \frac{X}{M}\mathcal{L}.$$

By Cauchy's theorem it follows that, for any non-integral $x \in [X, 2X]$,

$$S(x) - \frac{x}{V}A(1) = \frac{1}{2\pi i} I(x) + O\left(\frac{X}{M}\mathcal{L}^2\right), \tag{5.10}$$

where

$$I(x) = \int_{\eta - iM}^{\eta + iM} W(s)k(s)x^s \mathrm{d}s.$$

From (5.10) we deduce

$$\int_X^{2X} \left|S(x) - \frac{x}{V}A(1)\right|^2 \mathrm{d}x \ll \int_X^{2X} |I(x)|^2 \mathrm{d}x + X\left(\frac{X}{M}\right)^2 \mathcal{L}^4.$$

The last term here is bounded by the right-hand side of (5.8), so it suffices to estimate the first term. We have $I(x) = x^\eta \widehat{f}(-\log x)$ with

$$f(t) = \begin{cases} W(\eta + it)k(\eta + it) & (|t| \le M), \\ 0 & (|t| > M). \end{cases}$$

Hence

$$\begin{aligned} \int_X^{2X} |I(x)|^2 \mathrm{d}x &\le (2X)^{2\eta} \int_{\log X}^{\log(2X)} \left|\widehat{f}(-y)\right|^2 e^y \mathrm{d}y \\ &\ll X^{1+2\eta} \int_{-\infty}^{+\infty} \left|\widehat{f}(y)\right|^2 \mathrm{d}y \ll X^{1+2\eta} \int_{-\infty}^{+\infty} |f(t)|^2 \,\mathrm{d}t \end{aligned}$$

by Plancherel's theorem. Using the elementary estimate $k(\eta+it) \ll 1/V$ together with the bound

$$\left|\frac{\zeta'}{\zeta}(\eta+it)\right| \ll |1-\eta-it|^{-1} + \log(2+|t|) \ll \mathcal{L} \qquad (|t| \le M \le X),$$

which follows from the definition of η and Vinogradov's zero-free region (see, e.g., Titchmarsh (1986), p. 135) and implies in turn

$$W(\eta+it) \ll \mathcal{L}|A(\eta+it)|,$$

we obtain

$$\int_X^{2X} |I(x)|^2 \mathrm{d}x \ll X^{1+2\eta}\mathcal{L}^2 V^{-2} \int_{-M}^{M} |A(\eta+it)|^2 \mathrm{d}t.$$

By a classical mean value theorem (cf. Theorem 6.1 in Montgomery (1971)) the last integral is

$$\ll M \sum_{M<m\le 2M} |a(m)|^2 m^{-2\eta} \ll M^{1-2\eta} \sum_{M<m\le 2M} |a(m)|^2.$$

It follows that

$$\int_X^{2X} |I(x)|^2 \mathrm{d}x \ll X\mathcal{L}^2 \Big(\frac{X}{V}\Big)^2 \Big(\frac{X}{M}\Big)^{2\eta-2} \frac{1}{M} \sum_{M<m\le 2M} |a(m)|^2.$$

This completes the proof of (5.8), since $2\eta - 2 = -2\mathcal{L}^{-2/3-\epsilon}$.

Proof of Theorem 5.7. We begin by reducing the assertion of the theorem to a form that will be more convenient to prove. We may assume that X is sufficiently large in terms of ϵ, for otherwise the assertion holds trivially. Given $\epsilon > 0$ and $X \ge X_0(\epsilon)$, set

$$\mathcal{L} = \log X, \quad \delta = \mathcal{L}^{-3}, \quad \Delta_\alpha = \exp\big\{-\mathcal{L}^{1/6-\alpha}\big\}.$$

By splitting up the interval $[1, X]$ into intervals of the type $[x, (1+\delta)x]$, we see that it suffices to prove that (5.2) holds for all x in $[X, (1+\delta)X]$ with the exception of a set of measure $\ll \Delta_{\epsilon/2}X$. Since, for $x \in [X, (1+\delta)X]$,

$$\Psi(x+z,y) - \Psi(x,y) \begin{cases} \le \Psi\big(x(1+z/X), y\big) - \Psi(x,y), \\ \ge \Psi\Big(x(1+\dfrac{z/X}{1+\delta}), y\Big) - \Psi(x,y), \end{cases}$$

this will follow, if the estimate

$$\Psi\big(x(1+1/V),y\big)-\Psi(x,y)=\frac{x}{V}\rho(u)\Big\{1+O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\} \tag{5.12}$$

holds for $V=X/z$ and $V=(1+\delta)X/z$, for almost all $x\in[X,(1+\delta)X]$, with the same bound on the measure of the exceptional set. Theorem 5.1 shows that this is the case if either $y\geq X^{3/4}$ or $1\leq V\leq 1/\delta$. In view of the conditions (5.6) on z and y, we may therefore restrict ourselves to considering the range

$$\exp\big\{\mathcal{L}^{5/6+\epsilon}\big\}\leq y\leq X^{3/4},\qquad 1/\delta\leq V\leq 2\Delta_0X/y. \tag{5.13}$$

Next, set $U=\mathcal{L}/\log y$. By (5.13) and Corollary 2.3 we have $U\leq\mathcal{L}^{1/6-\epsilon}$ and

$$\rho(U)\gg\exp(-2U\log U)\gg\Delta_{\epsilon/2}. \tag{5.14}$$

Moreover, if $x\in[X,(1+\delta)X]$, then $x=\big(1+O(\delta)\big)X$ and $u=\log x/\log y=\big(1+O(\delta)\big)U=U+O(1/\mathcal{L}^2)$, so that by Corollary 2.4,

$$\rho(u)=\rho(U)\exp\big\{O\big(\log(U+1)/\mathcal{L}^2\big)\big\}=\rho(U)\big(1+O(1/\mathcal{L})\big).$$

Thus, x and u may be replaced by X and U in (5.12) without introducing additional error terms.

Having made these reductions, we now relate the quantity on the left of (5.12) to the sums $S(x)$. Set $r=[\log(\frac{2}{3}y)/\log(1+\delta)]$, define

$$M_i=\frac{X}{y}(1+\delta)^i,\quad P_i=\frac{X}{M_i}\qquad(i=0,\dots,r),$$

and set

$$P_i^+=P_i(1+\delta)^2,\qquad P_i^-=P_i/(1+\delta).$$

If X is sufficiently large, then these definitions together with (5.13) imply

$$X^{1/4}\leq X/y=M_0<M_1<\cdots<M_r\leq\tfrac{2}{3}X<M_r(1+\delta)\leq\tfrac{3}{4}X, \tag{5.15}$$

and

$$\tfrac{5}{4}\leq P_r^\pm<P_{r-1}^\pm<\cdots<P_0^\pm\leq y(1+\delta)^2. \tag{5.16}$$

Now note that any integer $n \geq 2$ can be written uniquely in the form mp with $p \geq 2$ and $P(m) \leq p$. If n is counted in the expression on the left-hand side of (5.12), then we have $2 \leq p \leq y$ in this representation and

$$X < n = mp \leq (1+\delta)(1+1/V)X \leq (1+\delta)^2 X \leq \tfrac{4}{3}X$$

by (5.13) and our assumptions that $V \geq 1/\delta$ and X is sufficiently large. Hence $X/y < m \leq \frac{2}{3}X$, so that by (5.15) each m falls into one of the disjoint intervals $\big(M_i, (1+\delta)M_i\big]$ $(0 \leq i \leq r)$. Moreover, if $m \in \big(M_i, (1+\delta)M_i\big]$ then we have $P_i^- \leq p < P_i^+$, with $P_i^\pm$ defined as above. Replacing for such m the condition $P(m) \leq p$ by the stronger (resp. weaker) condition $P(m) \leq P_i^-$ (resp. $P(m) \leq P_i^+$), we then obtain the inequalities

$$\Psi\big(x(1+1/V), y\big) - \Psi(x,y) \begin{cases} \leq \displaystyle\sum_{i=0}^{r} \sum_{\substack{M_i < m \leq (1+\delta)M_i \\ P(m) \leq P_i^+}} \sum_{x < mp \leq x(1+1/V)} 1, \\ \geq \displaystyle\sum_{i=0}^{r} \sum_{\substack{M_i < m \leq (1+\delta)M_i \\ P(m) \leq P_i^-}} \sum_{x < mp \leq x(1+1/V)} 1. \end{cases}$$

Setting

$$a_i^\pm(m) = \begin{cases} 1 & \text{if} \quad M_i < m \leq (1+\delta)M_i \quad \text{and} \quad P(m) \leq P_i^\pm, \\ 0 & \text{otherwise,} \end{cases}$$

and defining the associated functions $w_i^\pm(n)$, $A_i^\pm(s)$ and $S_i^\pm(x)$ as in Lemma 5.8, we deduce that

$$\Psi\big(x(1+1/V), y\big) - \Psi(x,y) \begin{cases} \leq \displaystyle\sum_{i=0}^{r} \frac{1}{\log P_i^-} S_i^+(x), \\ \geq \displaystyle\sum_{i=0}^{r} \frac{1}{\log P_i^+} S_i^-(x). \end{cases} \tag{5.17}$$

We now apply the lemma with ϵ replaced by $\epsilon/2$ and $M = M_i$ to each of the functions $S_i^\pm(x)$, $0 \leq i \leq r$. Since $P_i^\pm \asymp P_i = X/M_i$ and, by (1.4),

$$\begin{aligned} \sum |a_i^\pm(m)|^2 \leq \Psi\big((1+\delta)M_i, P_i^+\big) &\ll M_i \mathcal{L} \exp\Big\{-\frac{\log(1+\delta)M_i}{2\log P_i^+}\Big\} \\ &\ll M_i \mathcal{L} \exp\Big\{-\frac{\mathcal{L}}{4\log P_i}\Big\}, \end{aligned}$$

the expression on the right of (5.8) is bounded by

$$\ll X\mathcal{L}^4\Big(\frac{X}{V}\Big)^2\Big(R_i+Q_i\Big)$$

with

$$R_i = P_i^{-\mathcal{L}^{-2/3-\epsilon/2}}\mathcal{L}\exp\Big\{-\frac{\mathcal{L}}{4\log P_i}\Big\}, \qquad Q_i=\Big(\frac{V}{M_i}\Big)^2.$$

By (5.13) we have $V/M_i \le Vy/X \ll \Delta_0$, so that $Q_i \ll \Delta_0^2 \ll \Delta_{\epsilon/4}$. Moreover, since

$$\begin{aligned}\log R_i &= -\mathcal{L}^{-2/3-\epsilon/2}\log P_i - \frac{\mathcal{L}}{4\log P_i} + \log\mathcal{L}\\ &\le \max_{t\ge 1}\Big(-\mathcal{L}^{-2/3-\epsilon/2}t-\frac{\mathcal{L}}{4t}\Big)+\log\mathcal{L} = -\mathcal{L}^{1/6-\epsilon/4}+\log\mathcal{L},\end{aligned}$$

we also have $R_i \ll \mathcal{L}\Delta_{\epsilon/4}$. Hence the right-hand side of (5.8) is bounded by

$$\ll \Delta_{\epsilon/4}\mathcal{L}^5X\Big(\frac{X}{V}\Big)^2 \ll \Delta_{\epsilon/2}^3\mathcal{L}^{-20}X\Big(\frac{X}{V}\Big)^2 \ll \Delta_{\epsilon/2}\mathcal{L}^{-20}X\Big(\frac{X}{V}\Big)^2\rho(U)^2,$$

where we have used (5.14) in the last estimate. It follows that for any fixed choice of $(i,\pm)$, the set of those values $x\in[X,(1+\delta)X]$ for which the inequality

$$\Big|S_i^{\pm}(x)-\frac{x}{V}A_i^{\pm}(1)\Big| \le \mathcal{L}^{-5}\frac{X}{V}\rho(U) \tag{5.18}$$

fails, has measure

$$\le \Big(\mathcal{L}^{-5}\frac{X}{V}\rho(U)\Big)^{-2}\int_X^{(1+\delta)X}\Big|S_i^{\pm}(x)-\frac{x}{V}A_i^{\pm}(1)\Big|^2\,dx \ll \mathcal{L}^{-10}\Delta_{\epsilon/2}X.$$

Since $r\ll \mathcal{L}/\delta=\mathcal{L}^4$, the set of such values for which (5.18) fails for *some* choice of $(i,\pm)$, has measure $\ll \mathcal{L}^{-6}\Delta_{\epsilon/2}X \ll \Delta_{\epsilon/2}X$. To complete the proof of the theorem, it now suffices to show that (5.12) holds (with x and u replaced by X and U on the right-hand side) for any $x\in[X,(1+\delta)X]$, for which (5.18) is satisfied for all choices of $(i,\pm)$. By (5.16)–(5.18) and the bound $r\ll\mathcal{L}^4$, this will follow if we can show that

$$\sum_{i=0}^{r}\frac{A_i^{\pm}(1)}{\log P_i} = \rho(U)\Big\{1+O\Big(\frac{\log(U+1)}{\log y}\Big)\Big\}. \tag{5.19}$$

To prove (5.19), we first note that

$$A_i^{\pm}(1) = \sum_{\substack{M_i < m \le (1+\delta)M_i \\ P(m) \le P_i^{\pm}}} \frac{1}{m} = \big(1 + O(\delta)\big) \frac{D(M_i, P_i^{\pm})}{M_i},$$

where

$$D(M, P) = \Psi\big((1+\delta)M, P\big) - \Psi(M, P).$$

To estimate the quantities $D(M_i, P_i^{\pm})$ we define r_1 by

$$\log P_{r_1+1}^{+} \le \sqrt{\mathcal{L}} < \log P_{r_1}^{+},$$

and consider separately the ranges $0 \le i \le r_1$ and $r_1 < i \le r$. In the second range a trivial estimate using (1.4) gives

$$\begin{aligned} D(M_i, P_i^{\pm}) &\le \Psi\big((1+\delta)M_i, \exp\sqrt{\mathcal{L}}\,\big) \\ &\ll M_i \mathcal{L} \exp\Big\{-\frac{\log\big(1+\delta)M_i\big)}{2\sqrt{\mathcal{L}}}\Big\} \ll M_i \Delta_0. \end{aligned}$$

The contribution of these terms to the sum in (5.19) is therefore $\ll r\Delta_0 \ll \rho(U)/\mathcal{L}$ and thus negligible. In the remaining range $0 \le i \le r_1$, we have $\log P_i^{\pm} \gg \sqrt{\mathcal{L}}$, and applying Theorem 5.1 in the case $P_i^{\pm} \le M_i$ and the trivial estimate

$$D(M_i, P_i^{\pm}) = \delta M_i \Big\{1 + O\Big(\frac{1}{\delta M_i}\Big)\Big\}$$

together with the bound $M_i \ge X/y \ge X^{1/4}$ (see (5.15)) otherwise, we obtain in either case

$$D(M_i, P_i^{\pm}) = \Big\{1 + O\Big(\frac{\log(u_i^{\pm} + 1)}{\log P_i^{\pm}}\Big)\Big\}\rho(u_i^{\pm})\delta M_i,$$

where

$$u_i^{\pm} = \frac{\log M_i}{\log P_i^{\pm}} = \frac{\mathcal{L}}{\log P_i^{\pm}} - \frac{\log P_i}{\log P_i^{\pm}}.$$

Since $P_i^{\pm} = \big(1 + O(\delta)\big)P_i$ and $\sqrt{\mathcal{L}} \ll \log P_i^{\pm} \ll \mathcal{L} \ll \log M_i$ in the range under consideration, we have

$$\begin{aligned} \log P_i^{\pm} &= \log P_i + O(\delta) = \log y - i\log(1+\delta) + O(\delta) \\ &= (\log y - i\delta)\big(1 + O(\delta\mathcal{L})\big) = (\log y - i\delta)\big(1 + O\big(1/\mathcal{L}^2\big)\big), \end{aligned} \tag{5.20}$$

and therefore

$$u_i^{\pm} = \big(1 + O(1/\mathcal{L}^2)\big)\big(u(i) - 1\big), \quad \text{where} \quad u(t) = \mathcal{L}/(\log y - t\delta).$$

By Corollary 2.4, it follows that

$$\rho(u_i^{\pm}) = \rho(u(i)-1)\Big(1+O\left(\log(u(i)+2)/\mathcal{L}^2\right)\Big) = \rho\big(u(i)-1\big)\Big(1+O\big(1/\mathcal{L}\big)\Big).$$

The contribution of the terms $i \le r_1$ to the left-hand side of (5.19) therefore becomes

$$\text{(5.21)} \qquad \frac{\delta}{\mathcal{L}} \sum_{i=0}^{r_1} u(i)\rho\big(u(i) - 1\big)\Big\{1 + O\Big(\frac{u(i)\log(u(i)+2)}{\mathcal{L}}\Big)\Big\},$$

and it remains to show that this expression is equal to the right-hand side of (5.19).

We first note that $u(t)$ is monotonically increasing for $0 \le t < (\log y)/\delta$ and that, by the definition of r_1,

$$u(r_1) \asymp \mathcal{L}/\log P_{r_1} \asymp \sqrt{\mathcal{L}}.$$

It follows that

$$u(t) \ll \sqrt{\mathcal{L}} \quad (t \le r_1), \qquad u(t) \gg \sqrt{\mathcal{L}} \quad \big(r_1 < t < (\log y)/\delta\big),$$

and, by Corollary 2.3,

$$\text{(5.22)} \qquad \rho\big(u(t) - 1\big) \ll e^{-\sqrt{\mathcal{L}}} \ll \Delta_0 \qquad \big(r_1 < t \le (\log y)/\delta\big).$$

Moreover, since for $0 \le t \le r_1 + 1$

$$\text{(5.23)} \qquad u'(t) = \delta\mathcal{L}/(\log y - t\delta)^2 = \delta u(t)^2/\mathcal{L} \ll 1/\mathcal{L}^2,$$

we have, by another application of Corollary 2.4,

$$\rho\big(u(t) - 1\big) = \rho\big(u(i) - 1\big)\big(1 + O(1/\mathcal{L})\big) \qquad \big(0 \le i \le r_1,\ i \le t \le i + 1\big),$$

whence

$$u(i)\rho\big(u(i) - 1\big) = \big(1 + O(1/\mathcal{L})\big)\int_i^{i+1} u(t)\rho\big(u(t) - 1\big)\mathrm{d}t \qquad (0 \le i \le r_1).$$

Thus the sum in (5.21) is, apart from a factor $1 + O(1/\mathcal{L})$, equal to the integral

$$\int_0^{r_1+1} u(t)\rho\big(u(t)-1\big)\Big\{1 + O\Big(\frac{u(t)\log(u(t)+2)}{\mathcal{L}}\Big)\Big\}\mathrm{d}t.$$

The change of variables $v = u(t)$, $\mathrm{d}v = u'(t)\mathrm{d}t = (v^2\delta/\mathcal{L})\mathrm{d}t$ (see (5.23)) shows that this integral is equal to

$$\frac{\mathcal{L}}{\delta}\int_U^{U_1} \frac{\rho(v-1)}{v}\Big\{1 + O\Big(\frac{v\log(v+2)}{\mathcal{L}}\Big)\Big\}\mathrm{d}v,$$

where $U_1 = u(r_1+1)$. Since $\rho(v-1)/v = -\rho'(v)$, the main term in the last integral equals

$$-\int_U^{U_1} \rho'(v)\mathrm{d}v = \rho(U) - \rho(U_1) = \rho(U)\big(1 + O(1/\mathcal{L})\big),$$

by (5.22) and (5.14), and the error term is bounded by

$$\ll \frac{1}{\mathcal{L}}\int_U^{U_1} \big(-\rho'(v)\big)v\log(v+2)\mathrm{d}v.$$

Splitting the range in the last integral into the parts $U \le v \le 2U$ and $2U < v \le U_1$, estimating the contribution of the first part by

$$\ll \frac{U\log(U+1)}{\mathcal{L}}\int_U^{U_1} \big(-\rho'(v)\big)\mathrm{d}v \ll \frac{\log(U+1)}{\log y}\rho(U)$$

and that of the second part using the bounds of Corollary 2.4 for $\rho(v-1)$, it is easily seen that the contribution of the error term is by a factor $\log(U+1)/\log y$ smaller than the main term. Combining these estimates gives (5.19) and thus completes the proof of Theorem 5.7.

6. Distribution in arithmetic progressions

A natural problem is to investigate the distribution of integers without large prime factors in various arithmetically interesting sequences. The simplest example for such a sequence is an arithmetic progression. In analogy to the function $\Psi(x,y)$ one defines, for positive integers a and q, $\Psi(x,y;a,q)$ as the number of positive integers $\le x$, free of prime factors $> y$, and satisfying $n \equiv a \bmod q$. Since $\Psi(x,y;a,q) = \Psi(x/d,y;a/d,q/d)$ for any common divisor d of a and q with $P(d) \le y$ and $\Psi(x,y;a,q) = 0$ if $P\big((a,q)\big) > y$, it suffices to consider the case when a and q are coprime. The goal then is to show that numbers without large prime factors are uniformly distributed over the $\phi(q)$ residue classes $a \bmod q$, $(a,q) = 1$, in as large a range for the parameters x, y, and q, as possible.

The first result of this type is due to Buchstab (1949) who showed that for fixed $u = \log x/\log y$ and fixed positive integers q and a, $(a,q) = 1$, one has

$$\Psi(x,y;a,q) = \frac{x\rho(u)}{q}\Big\{1 + O\big((\log x)^{-1/2}\big)\Big\}. \tag{6.1}$$

Ramaswami (1951) gave a uniform estimate which shows that the asymptotic relation in (6.1) remains valid in the range $u \ll (\log_2 x)^{1-\epsilon}$, although his proof was incomplete; for a rigorous proof and discussion of Ramaswami's result see Norton (1971). A substantial improvement over these results is contained in the work of Levin & Fainleib (1967). However, as had been pointed out by Norton (1971) and others, the argument of Levin & Fainleib lacks in clarity and is marred by numerous inaccuracies and misprints. Their principal result, as quoted in Norton (1971), states that for any fixed positive integer k and positive real numbers ϵ and A, and uniformly in the range

$$y \geq y_0(\epsilon), \quad k+1+\epsilon \leq u \leq (\log y)^{3/5-\epsilon}, \quad q \leq (\log x)^A, \quad (a,q) = 1 \tag{6.2}$$

the estimate
(6.3)

$$\Psi(x,y;a,q) = x\Big\{\frac{\rho(u)}{q} + \frac{1}{\phi(q)}\sum_{i=1}^{k} a_i(q)\frac{\rho^{(i)}(u)}{(\log y)^i} + O_{k,\epsilon,A}\Big(\frac{u\rho^{(k)}(u)}{(\log y)^{k+1}}\Big)\Big\}$$

holds, where $a_i(q)$ is the ith Taylor coefficient at the origin of the function $s\zeta(s+1)(s+1)^{-1}\prod_{p|q}\left(1-p^{-s-1}\right)$. Friedlander (1984b) considered the problem of bounding $\Psi(x,y;a,q)$ from below for most residue classes a mod q and a given modulus q. He showed in particular that $\Psi(x,y;a,q)$ is positive for all but $o(q)$ residue classes a mod q, provided $q = o(x)$ and $q \leq y^{2-\epsilon}$.

Significant progress on the problem of estimating $\Psi(x,y;a,q)$ has been made very recently by Fouvry & Tenenbaum (1991) and Granville (1993a, 1993b), and we shall devote most of the remainder of this section to a survey of this important work.

We begin with a result of Fouvry & Tenenbaum (1991) which represents an analogue of Saias' estimate for $\Psi(x,y)$ (Theorem 1.8). Let

$$\Lambda_q(x,y) = \begin{cases} x\displaystyle\int_{-\infty}^{\infty} \rho(u-v)\mathrm{d}R_q(y^v) & (x \notin \mathbb{N}), \\ \Lambda_q(x+0,y) & (x \in \mathbb{N}), \end{cases}$$

where

$$R_q(x) = x^{-1}\#\{n \leq x : (n,q) = 1\} - \phi(q)/q.$$

THEOREM 6.1. *Let A be a fixed positive number. Then, with a suitable constant $c = c(A) > 0$, the estimate*

$$\Psi(x,y;a,q) = \frac{1}{\phi(q)}\Lambda_q(x,y)\Big\{1 + O\Big(e^{-c\sqrt{\log y}}\Big)\Big\} \tag{6.4}$$

holds uniformly in the range

$$x \geq 3, \qquad 1 \leq u \leq e^{c\sqrt{\log y}}, \qquad q \leq (\log x)^A, \qquad (a,q) = 1. \tag{6.5}$$

Fouvry & Tenenbaum have also given an estimate similar to (6.4), but with a weaker error term, for the range $q \leq \exp\big(c\sqrt{\log y}\big)$.

The function $\Lambda_q(x,y)$ generalizes the function $\Lambda(x,y)$ defined in (1.29). It can be shown that, for any fixed integer $k \geq 0$ and real $\epsilon > 0$, and uniformly in $x \geq 2$, $(\log x)^2 \leq y \leq x$, $q \geq 1$ and $u \geq k + \epsilon$ the estimate

$$\Lambda_q(x,y) = x\sum_{i=0}^{k} a_i(q)\frac{\rho^{(i)}(u)}{(\log y)^i} + O_k\Big(x\frac{\rho^{(k+1)}(u)}{(\log y)^{k+1}}(\log_2(qy))^k\Big) \tag{6.6}$$

holds; see Fouvry & Tenenbaum (1991). When combined with (6.6), the estimate (6.4) sharpens the Levin-Fainleib estimate (6.3), while at the same time the range (6.5) is considerably larger than the range (6.2) of Levin & Fainleib.

To prove (6.4), Fouvry & Tenenbaum first establish, in the range (6.5), the estimate

$$\Psi(x,y;a,q) = \frac{1}{\phi(q)}\Psi_q(x,y)\Big\{1 + O\Big(e^{-c\sqrt{\log y}}\Big)\Big\}, \tag{6.7}$$

with

$$\Psi_q(x,y) = \#\{n \leq x : P(n) \leq y, (n,q) = 1\},$$

and then show that in the range

$$x \geq x_0(\epsilon), \quad (\log_2 x)^{5/3+\epsilon} \leq \log y \leq \log x, \quad \log_2(q+2) \leq \Big(\frac{\log y}{\log(u+1)}\Big)^{1-\epsilon}$$

(which contains (6.5) for large x) one has

$$\Psi_q(x,y) = \Lambda_q(x,y)\left\{1 + O\left(\exp\left\{-(\log y)^{3/5-\epsilon}\right\}\right)\right\}. \tag{6.8}$$

The proof of (6.7) depends on bounds for the character sums

$$\sum_{n\in S(x,y)} \chi(n),$$

which are obtained by analytic techniques similar to those used in the proof of the Siegel-Walfisz theorem for primes in arithmetic progressions.

The estimation of the function $\Psi_q(x,y)$ is an interesting and non-trivial problem by itself, especially if one is interested in estimates that are uniform with respect to q. For other results on this subject, see Norton (1971) and Hazlewood (1975b), who obtained uniform estimates for $\Psi_q(x,y)$ in terms of $\rho(u)$, with error terms depending on the number of prime factors of q. Tenenbaum (1993) gave an optimal range of validity and a sharp error term for the approximation of $\Psi_q(x,y)$ by $\big(\phi(q)/q\big)\Psi(x,y)$. The estimate (6.8) is proved using the identity

$$\Psi_q(x,y) = \sum_{\substack{d|q\\ P(d)\le y}} \mu(d)\Psi\Big(\frac{x}{d},y\Big)$$

and estimating the terms on the right-hand side by means of Theorem 1.8 and Corollary 1.7. The same identity was also at the basis of the work of Norton referred to above. An alternative approach to such estimates is to mimic the proofs of the corresponding estimates for $\Psi(x,y)$.

Granville (1993a, 1993b) used a completely different, elementary method to obtain estimates for $\Psi(x,y;a,q)$ which are valid for q as large as a fixed power of y, though in general do not give as good an error term as (6.4). The method is an extension of the argument used in Hildebrand (1986a) to prove Theorem 1.1. It is based on the identities

$$(6.9)\quad \Psi(x,y;a,q)\log x = \int_1^x \frac{\Psi(t,y;a,q)}{t}dt + \sum_{\substack{p^m\le x\\ p\le y}} \Psi\Big(\frac{x}{p^m},y;\frac{a}{p^m},q\Big)\log p,$$

and

$$(6.10)\qquad \Psi_q(x,y)\log x = \int_1^x \frac{\Psi_q(t,y)}{t}dt + \sum_{\substack{p^m\le x\\ p\le y}} \Psi_q\Big(\frac{x}{p^m},y\Big)\log p,$$

which are valid uniformly in $x\ge 1$, $y\ge 2$, and positive integers a and q, and where, in (6.9), a/p^m denotes division in the multiplicative group $(\mathbb{Z}/q\mathbb{Z})^*$. These identities generalize the identity (1.11) for $\Psi(x,y)$ and can be proved in the same way. By using (6.9) and (6.10) iteratively, Granville (1993a) showed that, if the ratio $\Psi(x,y;a,q)\phi(q)/\Psi_q(x,y)$ is close to 1 in an initial range for x of the form $y^A \le x \le y^B$ with some constants A and B, then it remains close to 1 for all $x \ge y$. The proof of an asymptotic estimate valid for all $x \ge y$ is thus reduced to the proof of the same estimate in a small initial range. Starting from a trivial estimate in the initial range, Granville (1993a) proved in this way that, for any fixed positive number A and uniformly for $x \ge y \ge 2$, $q \le \min(x, y^A)$, and $(a,q) = 1$ the estimate

$$\Psi(x,y;a,q) = \frac{1}{\phi(q)}\Psi_q(x,y)\Big\{1 + O_A\Big(\frac{\log q}{\log y}\Big)\Big\} \tag{6.11}$$

holds. By a more delicate argument, which is also based on the identities (6.9) and (6.10), Granville (1993b) gave the following stronger result.

THEOREM 6.2. *For any fixed $\epsilon > 0$ and uniformly for $x \ge y \ge 2$, $1 \le q \le y^{1-\epsilon}$, and $(a,q) = 1$, we have*

$$\Psi(x,y;a,q) = \frac{1}{\phi(q)}\Psi_q(x,y)\Big\{1 + O\Big(\frac{\log q}{u^c \log y} + \frac{1}{\log y}\Big)\Big\}, \tag{6.12}$$

where c is a positive constant.

Granville also shows that the factor u^{-c} in the error term can be replaced by e^{-uc} if there are no exceptional characters modulo q, and he remarks that in the range (6.5) the error term may be further reduced to $O\Big(e^{-c(u+\sqrt{\log y})} + y^{-c}\Big)$.

Theorem 6.2 implies the lower bound

$$\Psi(x,y;a,q) \gg \Psi_q(x,y)/\phi(q)$$

provided q is less than a sufficiently small power of y, while (6.11) shows that the corresponding upper bound holds whenever q does not exceed x and is bounded by a fixed, but arbitrarily large power of y. Previously, Friedlander (1973b) had obtained a non-trivial lower bound for $q \le y^{(\sqrt{e}-1)/\sqrt{e}+\epsilon}$, and in a later paper (1981) gave the upper bound

$$\Psi(x,y;a,q) \ll \frac{x}{q}\rho\Big(u - \frac{\log q}{\log y} - 4\Big)$$

for the range $1 \le u \le (\log y)^{1/4}$, $q \le x^{1/2}y^{-5/2}$. More recently, Balog & Pomerance (1992) have obtained non-trivial bounds which allow the modulus q to be as large as $y^{4/3-\epsilon}$. In particular, they showed that $\Psi(x,y;a,q)$ has order of magnitude x/q when y is a fixed power of x and $q \le \min(y^{4/3-\epsilon}, (x/y)^{4/3-\epsilon})$. For large u, their bounds become weaker, but, by using the iteration method sketched above, Granville (1993a) was able to extend the estimates of Balog & Pomerance to the full range $x \ge y$ without any loss in precision. The resulting estimate is as follows.

THEOREM 6.3. *For any fixed $\epsilon > 0$ and uniformly in the range*

$$y \geq 2, \quad q \leq y^{4/3-\epsilon}, \quad (a,q) = 1, \quad x \geq \max\left(y^{3/2+\epsilon}, yq^{3/4+\epsilon}\right), \tag{6.13}$$

we have

$$\Psi(x,y;a,q) \asymp \frac{1}{\phi(q)} \Psi_q(x,y). \tag{6.14}$$

The admissible range for the modulus q can be further extended, if one asks only for estimates that hold for "almost all" moduli q as in the Bombieri-Vinogradov theorem. Fouvry & Tenenbaum (1991) proved the following result of this type.

THEOREM 6.4. *Let A be a given positive number. There exists a constant $B = B(A)$ such that uniformly for $x \geq y \geq 2$ and $Q = \sqrt{x}(\log x)^{-B}$ we have*

$$\sum_{q \leq Q} \max_{z \leq x} \max_{(a,q)=1} \left| \Psi(z,y;a,q) - \frac{\Psi_q(z,y)}{\phi(q)} \right| \ll \frac{x}{(\log x)^A}. \tag{6.15}$$

The proof of this result is based on an idea of Motohashi (1976) according to which a Bombieri-Vinogradov type theorem holds for any arithmetic function which can be represented as a convolution $\alpha * \beta$, where α and β are arithmetic functions supported on intervals of the form $(M/2, M]$ and $(N/2, N]$ with $MN = x$ and satisfying some mild hypotheses. The theorem is obtained from this via an approximate identity which expresses the characteristic function of integers $\leq x$ free of prime factors $> y$ as a linear combination of such convolutions.

The estimate (6.15) is non-trivial only for relatively small values of u. However, by using this estimate as the initial step of an iteration argument, Granville (1993a) showed that it can be extended to all $x \geq y$ at the cost of decreasing the value of Q, and obtained the following result.

THEOREM 6.5. *Let A be a fixed positive number. Then there exist positive constants $B = B(A)$ and $C = C(A)$ such that uniformly in the range*

$$y \geq 100,\ 1 \leq Q \leq \min\left(\exp\left\{\frac{C \log y \log_2 y}{\log_3 y}\right\}, \sqrt{x}(\log x)^{-B}\right) \tag{6.16}$$

we have

$$\sum_{q \leq Q} \max_{z \leq x} \max_{(a,q)=1} \left| \Psi(z,y;a,q) - \frac{\Psi_q(z,y)}{\phi(q)} \right| \ll_A \frac{\Psi(x,y)}{(\log y)^A}. \tag{6.17}$$

7. Other results

The problems described in the previous sections – the global distribution of integers without large prime factors, the distribution in short intervals, and the distribution in arithmetic progressions – are the most natural problems in this subject and therefore have received the greatest attention in the literature. In this section we survey other results on the distribution of integers without large prime factors, discuss various generalizations of these problems and applications of the results, and we give references to related work in the literature. Norton's memoir (1971) contains a comprehensive survey and bibliography of results obtained before 1970. The bibliography at the end of the paper is intended to complement that of Norton and contains every paper written on the subject since Norton's memoir that we are aware of. For a similarly comprehensive bibliography see Moree's thesis (1993).

Local distribution of integers composed of small prime factors. Let $y \geq 2$ be fixed, and let $1 = n_1 < n_2 < \cdots$ denote the sequence of positive integers free of prime factors $> y$. The asymptotic behavior of $\log n_i$, as $i \to \infty$, can easily be derived from Ennola's theorem (Theorem 1.5). However this result gives no information on the *local* behavior of the sequence $\{n_i\}$, and more specifically the size of the differences between consecutive terms. This problem has been studied by Tijdeman (1973, 1974) who showed, using methods from transcendence theory, that there exist positive constants $c_1 = c_1(y)$ and $c_2 = c_2(y)$ such that for all $i > 1$,

$$\frac{n_i}{(\log n_i)^{c_1}} \leq n_{i+1} - n_i \leq \frac{n_i}{(\log n_i)^{c_2}}.$$

In another direction, a number of authors have obtained upper bounds for the maximal length of a string of consecutive integers $> y$ belonging to the sequence $\{n_i\}$, i.e., the maximal gap $f(y)$ in the sequence of positive integers having a prime factor $> y$. The first result of this type is due to Erdős (1955) who proved that $f(y) \ll y/\log y$. This estimate has subsequently been improved by Ramachandra (1969, 1970, 1971), Tijdeman (1972), Ramachandra & Shorey (1973), Shorey (1973), and Jutila (1974), the strongest result being Shorey's bound $f(y) \ll y \log_3 y/(\log y \log_2 y)$. Some of these results have been obtained using an elementary method of Halberstam & Roth (1951) for bounding gaps between consecutive k-free numbers, while others rely on deep estimates for linear forms in logarithms obtained by Baker's method. The above results may also be phrased in terms of the function $P(n,k) = \max\{P(n+i) : 0 \leq i \leq k-1\}$; for example, Shorey's bound for $f(y)$ is essentially equivalent to the bound

$P(n,k) \gg k \log k \log_2 k / \log_3 k$ for $n \geq k^{3/2}$. A stronger bound for large n was given by Langevin (1975) who showed that for any $\epsilon > 0$, all sufficiently large k and $n \geq n_0(\epsilon, k)$, one has $P(n,k) \geq (1-\epsilon)k \log\log n$. Turk (1980) gave similar results for the greatest prime factor of values of polynomials at consecutive integers. For other results on $P(n,k)$ see, for example, Eggleton & Selfridge (1976).

Distribution in special sequences. The study of the distribution of integers without large prime factors in various "natural" arithmetical sequences is an interesting and difficult problem. The distribution in short intervals and in arithmetic progressions has been discussed in the preceding sections. Other sequences of interest which have been investigated in the literature are polynomial sequences and sequences of "shifted primes".

The case of polynomial sequences has been studied by Hmyrova (1964, 1966), Wolke (1971), and Timofeev (1977), who gave upper bounds for the function $\Psi(f;x,y) = \#\{n \leq x : P(f(n)) \leq y\}$, where f is a polynomial mapping the positive integers into themselves. (As observed by Friedlander, the proof of the main result in Wolke (1971) is defective, and the result as stated therefore remains unproven.) The problem of obtaining lower bounds for $\Psi(f;x,y)$ seems to be much more difficult, and non-trivial results are known only for linear polynomials in which case the problem reduces to that of the distribution in arithmetic progressions.

Another interesting question related to the distribution of integers without large prime factors in polynomial sequences is the estimation of

$$P_x = P\Big(\prod_{n\leq x} F(n)\Big),$$

where F is an irreducible polynomial with integer coefficients and degree $g > 1$. Equivalently, P_x is the smallest y such that $\Psi(F;x,y) = [x]$. The standard conjectures in prime number theory imply that $F(n)$ is prime for at least one $n \in (x/2, x]$, and therefore $P_x \asymp x^g$, if x is sufficiently large. Improving on a result of Nagell (1921), Erdős (1952) showed that

$$P_x > x(\log x)^{c_0(F)\log_3 x} \quad (x > x_0(F))$$

for some positive constant $c_0(F)$. Using estimates for Kloosterman sums, Hooley (1967) and Deshouillers & Iwaniec (1982) improved this bound in the case of a quadratic polynomial F to $P_x > x^{1+\mu}$ with $\mu = \frac{1}{10}$ and $\mu = 0{\cdot}202$, respectively. However their method does not seem to be applicable to higher degree polynomials; see Hooley (1978). For general polynomials, Erdős' bound has been improved only recently, first by Erdős &

Schinzel (1990), and subsequently by Tenenbaum (1990c), who showed that one has for any fixed $\alpha < 2 - \log 4 = 0{\cdot}61370\ldots$

$$P_x > x \exp\{(\log x)^\alpha\} \quad (x \geq x_0(F, \alpha)).$$

The distribution of integers without large prime factors among integers of the form $p - 1$ (and more generally $p + a$) is of interest because of its connection to the distribution of values of the Euler Phi function. An argument of Erdős (1935) shows that if $c > 0$ is such that $P(p-1) < p^c$ holds for a positive proportion of all primes, then for infinitely many n the equation $\phi(m) = n$ has $\gg_\epsilon n^{1-c-\epsilon}$ solutions m. Erdős proved that there exists a constant $c < 1$ for which this holds, and Pomerance (1980) showed that one may take $c = 0{\cdot}44\ldots$. Balog (1984) proved that $P(p-1) < p^{0{\cdot}35}$ holds for infinitely many primes p. The exponent $0{\cdot}35$ was subsequently improved by Fouvry & Grupp (1986) to $0{\cdot}317\ldots$, and later by Friedlander (1989) to $1/(2\sqrt{e}) = 0{\cdot}303\ldots$, which represents the current record. In both cases the corresponding bound was shown to hold for a positive proportion of all primes. Alford, Granville, & Pomerance (1993) have applied the last result in their recent proof of the infinitude of Carmichael numbers and also showed that further improvements in the exponent can be obtained if one assumes certain conjectures on the distribution of prime numbers in arithmetic progressions.

Additive problems involving integers without large prime factors. In analogy to the twin prime and Goldbach problems one can ask whether there are infinitely many pairs (and, more generally, k-tuples) of consecutive integers without large prime factors, and whether every sufficiently large integer is a sum of two such integers. Specifically, we consider positive integers n satisfying $P(n) \leq n^\alpha$ for some α (possibly depending on n), and the goal is to prove such results with α as small as possible.

In contrast to the situation in the case of primes, the first of these problems turns out to be easier than the second. In fact, it follows a general result of Hildebrand (1985b) asserting the existence of pairs of consecutive integers in certain sets that for every $\epsilon > 0$ there exist infinitely many pairs $(n, n+1)$ of consecutive integers having no prime factors $> n^\epsilon$. Balog, Erdős, & Tenenbaum (1990), using an argument of Heath-Brown (1987), gave a quantitative version of this result which shows that the exponent may be taken as $\epsilon(n) = c \log_3 n / \log_2 n$ with a suitable constant c.

Comparable results for k-tuples of integers without large prime factors are not known when $k \geq 3$. However, it follows from simple density considerations together with the relation $\Psi(y^u, y) \sim x\rho(u)$ that for any $k \geq 2$ and any $\epsilon > 0$ there exist infinitely many k-tuples $(n, n+1, \ldots, n+k-1)$ of consecutive integers all of whose prime factors are $\leq n^{\alpha_k+\epsilon}$, where $\alpha_k < 1$ is defined by the equation $\rho(1/\alpha_k) = 1-1/k$. A recent result of Hildebrand (1989) yields a slight improvement over this trivial bound, showing that α_k may be replaced by the smaller quantity α_{k+1}.

Goldbach type problems for integers without large prime factors seem to be more difficult. Balog & Sárközy (1984b) proved that every sufficiently large integer N can be written as a sum $N = n_1 + n_2$ with $P(n_i) < 2N^{2/5}$, and Balog (1989) improved the exponent $\frac{2}{5}$ to $0{\cdot}2695\ldots$. However, it is not known whether such a representation exists with $P(n_i) \leq N^{\epsilon}$, for any fixed $\epsilon > 0$. The corresponding ternary problem is much better understood; Balog and Sárközy (1984a) showed that every sufficiently large integer N has a representation $N = n_1+n_2+n_3$ with $P(n_i) \leq \exp\left\{3\sqrt{\log N \log_2 N}\right\}$.

Representations of integers as sums of powers of numbers without large prime factors have played an important role in recent work on Waring's problem (see, for example, Vaughan (1986) and Wooley (1992)). This is due to the fact that, in certain circumstances, the number of representations of an integer as a sum of powers of integers with restrictions on the size of its prime factors is easier to estimate asymptotically by the Hardy-Littlewood method than the number of unrestricted representations.

Sums over integers without large prime factors. Let $S(x,y)$ denote the set of positive integers $\leq x$ free of prime factors $> y$. A number of authors have studied the asymptotic behavior of sums of the form $\sum_{n\in S(x,y)} f(n)$, which are sometimes referred to as "incomplete sums". The case when $f(n)$ is equal to $\mu^2(n)/\phi(n)$ (or some similar multiplicative function) occurs in sieve theory and has been first investigated by van Lint & Richert (1964). De Bruijn & van Lint (1964) considered other special classes of multiplicative functions. Halberstam & Richert (1971) later gave a much more general result which contained most of previous estimates. Results of this type are also implicit in the papers of Levin & Fainleib (1967), which, however, has been criticized (cf. Norton (1971)) as being inaccurate at many places. More recently, Levin & Chariev (1986) have given rather general estimates for sums of mulitplicative functions over integers with constraints on the size of the prime factors.

Of particular interest is the case $f(n) = z^{\Omega(n)}$, where $\Omega(n)$ denotes the total number of prime factors of n. The behavior of the sums

$$\sum_{n\in S(x,y)} z^{\Omega(n)}$$

is closely related to the distribution of the function $\Omega(n)$ among the elements of $S(x,y)$ and thus can give some insight into the multiplicative structure of these elements. This case has been investigated by De Koninck & Hensley (1978), Alladi (1987), Hensley (1987), and Hildebrand (1987).

Another interesting case is that of $f(n) = \mu(n)$. In view of the relation $\sum_{n\leq x}\mu(n) = o(x)$ one might expect that $\sum_{n\in S(x,y)}\mu(n) = o(\Psi(x,y))$ holds under suitable conditions on x and y. Answering a question of Erdős, Alladi (1982) and Hildebrand (1987) showed that this relation is indeed valid, provided that $y \to \infty$ and $x \geq y$; more precisely, the bound

$$\Big|\sum_{n\in S(x,y)}\mu(n)\Big| \ll \frac{\Psi(x,y)}{\log y}$$

holds uniformly in $x \geq y \geq 2$. Alladi showed that this bound is best-possible as it stands. However, sharper estimates for the sum on the left-hand side can be obtained if $u = \log x/\log y \to \infty$; see Tenenbaum (1990b).

The estimation of the sums $\sum_{n\in S(x,y)} d(n)$, where $d(n)$ is the divisor function, is of some interest since this problem can be regarded as the analog of the classical divisor problem. Xuan (1990, 1991) and Smida (1993) gave estimates for these sums, and for the corresponding sums over the generalized divisor functions $d_k(n)$. Fouvry & Tenenbaum (1990) considered the problem of estimating the sum $\sum_{n\in S(x,y)} d(n-1)$ which is the analog of the "Titchmarsh divisor problem", i.e., the problem of estimating sums of the divisor function over the sequence $\{p-1\}$ of shifted primes. They obtained an asymptotic formula for this sum, which is valid uniformly in the range $\exp\big\{\epsilon\log x\log_3 x/\log_2 x\big\} \leq y \leq x$, for any fixed $\epsilon > 0$.

Exponential sums over integers without large prime factors have been first studied by Vaughan (1989) in connection with Waring's problem. Fouvry & Tenenbaum (1991) have given rather sharp bounds for the exponential sums $E(x,y,\theta) = \sum_{n\in S(x,y)} e^{2\pi i\theta n}$; in particular, they showed that, uniformly for $3 \leq y \leq \sqrt{x}$, and all positive integers a and q with $(a,q) = 1$,

$$E\Big(x,y,\frac{a}{q}\Big) \ll x(\log qx)^3\left\{\frac{\sqrt{y}}{x^{1/4}} + \frac{1}{\sqrt{q}} + \sqrt{\frac{qy}{x}}\right\}.$$

This bound may be regarded as the analog of Vinogradov's bound for exponential sums over primes.

Distribution of arithmetic functions on integers without large prime factors. The problem of estimating sums of arithmetic functions over integers without large prime factors leads naturally to the question of the distribution of values of arithmetic functions among such integers. Perhaps the simplest example is given by the function $\mu^2(n)$ which describes the distribution of squarefree integers. This problem has been investigated by Ivić (1985b), Ivić & Tenenbaum (1986), Naimi (1988), and Granville (1989). Their results show, for example, that if $u = \log x/\log y \geq y^{1/2+o(1)}$ then the number of squarefree integers in $S(x,y)$ is $o(\Psi(x,y))$, as $y \to \infty$. On the other hand, for $u \leq y^{o(1)}$, this number is asymptotic to $(6/\pi^2)\Psi(x,y)$, so that in this case the proportion of squarefree numbers among elements of $S(x,y)$ is equal to the (global) density of squarefree numbers. The corresponding problem for k-free integers has been studied by Hazlewood (1975a).

Several authors have investigated the distribution of the function $\Omega(n)$, the total number of prime factors of n, on the set $S(x,y)$ and obtained various analogs of classical limit theorems; see Alladi (1987), Hensley (1987), and Hildebrand (1987). Roughly speaking, the results state that $\Omega(n)$ is approximately normally distributed on the set $S(x,y)$ with mean M and variance V, where $M = M(x,y)$ and $V = V(x,y)$ are suitable functions satisfying, for example, $M \sim V \sim \log\log x$ if $u = \log x/\log y = o(\log_2 x)$ and $M \asymp u$ and $V \asymp u/\log^u$ if $(\log y)^{20} \leq u \ll y/\log y$; see Alladi (1987) and Hildebrand (1987).

As in the classical case, results of this type can be extended to more general classes of additive functions. For example, Alladi (1982b) obtained a Turán-Kubilius inequality for $S(x,y)$, i.e., a Chebyshev type inequality for additive functions on $S(x,y)$. This inequality can be used to show that, for a fairly general class of additive functions $f(n)$, the values $f(n)$ are for "most" $n \in S(x,y)$ close to the average value of f over the set $S(x,y)$.

Sums involving the largest prime factor of an integer. Various authors have studied sums of the form $\sum_{n\leq x} f\big(P(n)\big)$, where f is some arithmetic function. Such sums can be evaluated using the identity

$$\sum_{n\leq x} f(P(n)) = \sum_{p\leq x} f(p)\Psi\Big(\frac{x}{p},p\Big)$$

together with sufficiently sharp estimates for the function $\Psi(x,y)$.

The case $f(n) = \log n$ was considered by de Bruijn (1951b); his argument, combined with Theorem 1.8, leads to the estimate

$$\sum_{n\leq x} \log P(n) = \lambda x\log x - \lambda(1-\gamma)x + O\Big(x\exp\big\{-(\log x)^{3/8-\epsilon}\big\}\Big),$$

where $\lambda = \int_0^\infty \rho(u)(1+u)^{-2}du = 0{\cdot}624\ldots$ (see Exercise III.5.3 in Tenenbaum (1990a)). The constant λ is known as Golomb's constant, and may be interpreted as the expected proportion between the number of digits of the largest prime factor of an integer and that of the integer itself; see

Knuth & Trabb Pardo (1976). Wheeler (1990) gave a similar formula for the more general sums $\sum_{n\le x}^{*}\left(\log P(n)\right)^{\alpha}$, where the summation is restricted to integers n whose second largest prime factor $P_2(n)$ satisfies $P_2(n)\le P(n)^{1/u}$ for some fixed u. Scourfield (1991) obtained asymptotic formulae for the counting functions $S_\gamma(x)=\#\{n>1: nP(n)^\gamma\le x\}$ and $T_\gamma(x)=\#\{n>1: n^\gamma(\log P(n))^\gamma\le x\}$. De Koninck & Mercier (1989) have given asymptotic formulae for $\sum_{n\le x} f(P(n))$ for strongly additive functions f satisfying $f(p)=p^\alpha L(p)$ where α is any given real number and L is a slowly oscillating function.

Perhaps the most interesting case is that of the function $f(n)=1/n$ which has been studied by Erdős & Ivić (1980), Ivić & Pomerance (1984), Ivić (1981,1987), and Xuan (1988). Erdős, Ivić & Pomerance showed that, as $x\to\infty$,

$$\sum_{n\le x}\frac{1}{P(n)}=x\delta(x)\Big\{1+O\Big(\Big(\frac{\log_2 x}{\log x}\Big)^{1/2}\Big)\Big\},$$

where $\delta(x)=\int_2^x \rho(\log x/\log t)t^{-2}\mathrm{d}t$. Sums such as $\sum_{n\le x}\Omega(n)/P(n)$ can be asymptotically evaluated in a similar way; see Ivić & Pomerance (1984) and Xuan (1989a, 1989b). Brouwer (1974), Alladi & Erdős (1977, 1979) and De Koninck & Ivić (1984a, 1984b) have investigated the asymptotic behavior of the sums $\sum_{n\le x}P_k(n)$ and $\sum_{n\le x}P_k(n)/P(n)$ and similar sums involving the kth largest prime factor $P_k(n)$ of n.

Van de Lune (1974) and van Rongen (1975) considered the asymptotic behavior of sums of the form $\sum_{2\le n\le x} f\big(\log n/\log P(n)\big)$ and showed that for a wide class of functions f, including all continuous bounded functions and all polynomials, the limit

$$\lim_{x\to\infty}\frac{1}{x}\sum_{2\le n\le x} f\Big(\frac{\log n}{\log P(n)}\Big)$$

exists and equals $-\int_1^\infty f(x)\mathrm{d}\rho(x)$.

Bounds for the least kth power non-residue. An important application of estimates for $\Psi(x,y)$, which goes back to Vinogradov (1926), is an upper bound for the least kth power non-residue modulo a prime p satisfying $p\equiv 1 \bmod k$. The connection is based on the observation that if $g_1=g_1(p,k)$ is the least positive kth power non-residue modulo a prime $p\equiv 1 \bmod k$ then, for any $y<g_1$ and $x\ge 1$, every integer counted in $\Psi(x,y)$ is a product of kth power residues, and hence is itself a kth power residue. We therefore have the inequality

$$\Psi(x,y)\le\#\{n\le x: n\equiv a^k \mod p \text{ for some } a\}$$

for any $y < g_1$. It follows from Burgess' character sum estimates that the right-hand side is asymptotic to x/k, as $p \to \infty$, uniformly in $x \geq p^{1/4+\epsilon}$, for any fixed $\epsilon > 0$. On the other hand, the left-hand side is $x\rho(u) + o(x)$ by the Dickman–de Bruijn formula. Taking $x = p^{1/4+\epsilon}$ and $y = g_1 - 1$, it follows that for any given $\epsilon > 0$ and any sufficiently large prime $p \equiv 1 \bmod k$ we have $g_1(p,k) \ll_\epsilon p^{\alpha_k+\epsilon}$ with α_k defined by the equation $\rho(1/4\alpha_k) = 1/k$. The asymptotic formula (1.7) shows that $\alpha_k = (1 + o(1))(\log_2 k)/(4 \log k)$ as $k \to \infty$. This result is due to Y. Wang (1964) and improves on earlier bounds of Vinogradov (1926) and Buchstab (1949). Norton (1968, 1971) generalized these results to bounds for the least kth power non-residue with respect to arbitrary (not necessarily prime) moduli, and to bounds for the least positive representatives of arbitrary cosets of the set of kth power residues.

The bound for $g_1(p,k)$ could be improved if sufficiently uniform upper bounds for integers without large prime factors in arithmetic progressions were available. For example, if the bound $\Psi(x,y;a,q) \leq (1+o(1))\, x\rho(u)/q$ holds whenever y and q are fixed but sufficiently small powers of x, then it would follow that $g_1(p,k) \ll_\epsilon p^\epsilon$ for every sufficiently large prime p with $(p-1,k) > 1$; see Friedlander (1973a). Unfortunately, the known bounds for $\Psi(x,y;a,q)$ are not sufficient to yield any improvements over the above bound $g_1 \ll_\epsilon p^{\alpha_k+\epsilon}$.

Generalizations to algebraic number fields. Given an algebraic number field K, let $\Psi_K(x,y)$ be the number of integral ideals with norm $\leq x$, all of whose prime divisors have norm $\leq y$. The problem of estimating $\Psi_K(x,y)$ has been studied by Jordan (1965), Gillet (1970), Friedlander (1972), Hazlewood (1975a, 1977), Krause (1990), Moree & Stewart (1990), and Moree (1992). The results obtained, and the methods of proof, are to a large extent analogous to those in the classical case of $\Psi(x,y)$. For example, Krause (1990) showed that Theorem 1.1 and Theorem 1.4 both hold with $\Psi_K(x,y)$ in place of $\Psi(x,y)$ with the same uniformity (but with the implied constants depending on the number field K), provided the right-hand side of (1.8) is multiplied by an appropriate constant $\mu = \mu_K$, equal to the residue of the Dedekind zeta function $\zeta_K(s)$ at $s = 1$.

Other generalizations. The function

$$\Theta(x,y,z) = \#\{n \leq x : p|n \Rightarrow z < p \leq y\}$$

has been first studied by Friedlander (1976), who showed, in particular, that for fixed u and v one has $\Theta(x, x^{1/u}, x^{1/v}) \sim x\sigma(u,v)/\log x^{1/v}$ as $x \to \infty$, where $\sigma(u,v)$ is defined by a system of differential-difference equations.

More precise estimates of this kind have recently been given by Saias (1992), Fouvry & Tenenbaum (1991), and Granville (1991a). Goldston & McCurley (1988) and Warlimont (1990) considered the problem of estimating the function $\Psi(x,y;Q) = \#\{n \leq x : p|n, p > y \Rightarrow p \notin Q\}$, where Q is a sufficiently well distributed set of primes of density δ. This function reduces to $\Psi(x,y)$ in the case Q is the set of all primes. However, it turns out that the behavior of $\Psi(x,y,Q)$ is rather different from that of $\Psi(x,y)$ if the set Q has density $\delta < 1$ among the primes; namely, $\Psi(x,y,Q)$ has then order of magnitude x/u^δ, rather than $x/u^{u(1+o(1))}$ as in the case of $\Psi(x,y)$. An analog of the function $\Psi(x,y)$ for polynomials over a finite field is the number $N_q(n,m)$ of monic polynomials of degree n over the field $\mathbb{F}_q$ all of whose irreducible factors have degrees at most m. This function arises in connection with the discrete logarithm problem and was studied by Odlyzko (1985,1993), Car (1987), and Lovorn (1992). Analogs of $\Psi(x,y)$ for general arithmetic semigroups have been considered by Warlimont (1991).

Distribution of the k largest prime factors of an integer. For fixed x, the function $\Psi(x,y)$ describes the distribution of the largest prime factor of an integer selected at random from $\{1,\ldots,[x]\}$. This point of view suggests a natural generalization, namely to consider the joint distribution of $(P_1(n),\ldots,P_k(n))$, where $P_i(n)$ denotes the ith largest prime factor of n. Result of this type can be deduced from general theorems of Levin & Fainleb (1967). A more specific result was given by Billingsley (1972), who showed that for any k-tuple $(\alpha_1,\ldots,\alpha_k) \in (0,1]^k$ the limit

$$F(\alpha_1,\ldots,\alpha_k) = \frac{1}{x}\#\{n \leq x : P_i(n) \leq n^{\alpha_i}\,(i=1,\ldots,k)\}$$

exists and gave an explicit representation for the associated density function. For other results in this direction see Galambos (1976), Knuth & Trabb Pardo (1976), Lloyd (1984), Vershik (1987), and Wheeler (1990).

References

M. Abramowitz & I. Stegun,
(1964) *Handbook of Mathematical Functions*, National Bureau of Standards, Washington 1964; Tenth Printing, Dover, New York, 1972.

W. R. Alford, A. Granville, & C. Pomerance,
(1993) There are infinitely many Carmichael numbers, preprint.

K. Alladi,
(1982a) Asymptotic estimates of sums involving the Moebius function. II, *Trans. Amer. Math. Soc.* **272**, 87–105.

(1982b) The Turán-Kubilius inequality for integers without large prime factors, *J. Reine Angew. Math.* **335**, 180–196.
(1987) An Erdős-Kac theorem for integers without large prime factors, *Acta Arith.* **49**, 81–105.

K. Alladi and P. Erdős,
(1977) On an additive arithmetic function, *Pacific J. Math.* **71**, 275–294.
(1979) On the asymptotic behavior of large prime factors of integers, *Pacific J. Math.* **82**, 295–315.

A. Balog,
(1984) $p + a$ without large prime factors, in: *Séminaire de Théorie des Nombres*, Bordeaux 1983-84, Univ. Bordeaux, Talence, Exp. No. 31, 5 pp.
(1987) On the distribution of integers having no large prime factor, in: *Journées Arithmétiques*, Besançon 1985, *Astérisque* **147/148**, pp. 27–31.
(1989) On additive representations of integers, *Acta Math. Hungar.* **54**, 297–301.

A. Balog, P. Erdős, & G. Tenenbaum,
(1990) On arithmetic functions involving consecutive divisors, in: *Analytic Number Theory* (B. C. Berndt *et al.*, eds.), Proc. Conf. in Honor of Paul T. Bateman, Birkhäuser, *Progress in Math.* **85**, 77–90.

A. Balog & C. Pomerance,
(1992) The distribution of smooth numbers in arithmetic progressions, *Amer. Math. Soc.* **115**, 33–43.

A. Balog & A. Sárközy,
(1984a) On sums of integers having small prime factors. I, *Stud. Sci. Math. Hungar.* **19**, 35–47.
(1984b) On sums of integers having small prime factors. II, *Stud. Sci. Math. Hungar.* **19**, 81–88.

F. Beukers,
(1975) The lattice-points of n-dimensional tedrahedra, *Indag. Math.* **37**, 365–372.

P. Billingsley,
(1972) On the distribution of large prime divisors, *Period. Math. Hungar.* **2**, 283–289.

A. E. Brouwer,
(1974) Two number theoretic sums, *Math. Centrum Amsterdam*, Report ZW 19/74.

N. G. de Bruijn,
(1951a) The asymptotic behaviour of a function occurring in the theory of primes, *J. Indian Math. Soc. (N.S.)* **15**, 25–32.
(1951b) On the number of positive integers $\leq x$ and free of prime factors $> y$, *Nederl. Akad. Wetensch. Proc. Ser. A* **54**, 50–60.
(1966) On the number of positive integers $\leq x$ and free of prime factors $>y$, II, *Nederl. Akad. Wetensch. Proc. Ser. A* **69**, 239–247.

N. G. de Bruijn & J. H. van Lint,
(1964) Incomplete sums of multiplicative functions I, II, *Nederl. Akad. Wetensch. Proc. Ser. A* **67**, 339–347; 348–359.

A. A. Buchstab,

(1949) On those numbers in an arithmetic progression all prime factors of which are small in magnitude (Russian), *Dokl. Akad. Nauk SSSR* **67**, 5–8.

E. R. Canfield,

(1982) On the asymptotic behavior of the Dickman–de Bruijn function, *Congr. Numer.* **38**, 139–148.

E. R. Canfield, P. Erdős, & C. Pomerance,

(1983) On a problem of Oppenheim concerning "Factorisatio Numerorum", *J. Number Theory* **17**, 1–28.

M. Car,

(1987) Théorèmes de densité dans $\mathbb{F}_q[X]$, *Acta Arith.* **48**, 145–165.

S. D. Chowla and T. Vijayaraghavan,

(1947) On the largest prime divisors of numbers, *J. Indian Math. Soc. (N. S.)* **11**, 31–37.

J.-M. De Koninck & D. Hensley,

(1978) Sums taken over $n \le x$ with prime factors $\le y$ of $z^{\Omega(n)}$, and their derivatives with respect to z, *J. Indian Math. Soc. (N. S.)* **42**, 353–365.

J.-M. De Koninck & A. Ivić,

(1984a) Sommes de réciproques de grandes fonctions additives, *Publ. Inst. Math. (Beograd) (N. S.)* **35**, 41–48.

(1984b) The distribution of the average prime divisor of an integer, *Arch. Math.* **43**, 37–43.

J.-M. De Koninck & A. Mercier,

(1989) Les fonctions arithmétiques et le plus grand facteur premier, *Acta Arith.* **52**, 25–48.

J.-M. Deshouillers & H. Iwaniec,

(1982) On the greatest prime factor of n^2+1, *Ann. Inst. Fourier (Grenoble)* **32**, 1–11.

K. Dickman,

(1930) On the frequency of numbers containing prime factors of a certain relative magnitude, *Ark. Mat. Astr. Fys.* **22**, 1–14.

R. Eggleton & J. L. Selfridge,

(1976) Consecutive integers with no large prime factors, *J. Austral. Math. Soc. Ser. A* **22**, 1–11.

V. Ennola,

(1969) On numbers with small prime divisors, *Ann. Acad. Sci. Fenn. Ser. AI* **440**, 16 pp.

P. Erdős,

(1935) On the normal number of prime factors of $p-1$ and some other related problems concerning Euler's Φ-function, *Quart. J. Math. (Oxford)* **6**, 205–213.

(1952) On the greatest prime factor of $\prod f(k)$, *J. London Math. Soc.* **27**, 379–384.

(1955) On consecutive integers, *Nieuw Arch. Wisk.* (3) **3**, 124–128.

P. Erdős & A. Ivić,

(1980) Estimates for sums involving the largest prime factor of an integer and certain related additive functions, *Stud. Sci. Math. Hungar.* **15**, 183–199.

P. Erdős, A. Ivić, and C. Pomerance,

(1986) On sums involving reciprocals of the largest prime factor of an integer, *Glas. Mat.* III. Ser. **21** (41), 283–300.

P. Erdős & A. Schinzel,

(1990) On the greatest prime factor of $\prod_{k=1}^{x} f(k)$, *Acta Arith.* **55**, 191–200.

A. S. Fainleib,

(1967) On the estimation from below of the number of numbers with small prime factors (Russian), *Dokl. Akad. Nauk USSR* **7**, 3–5.

E. Fouvry & F. Grupp,

(1986) On the switching principle in sieve theory, *J. Reine Angew. Math.* **370**, 101–126.

E. Fouvry & G. Tenenbaum,

(1990) Diviseurs de Titchmarsh des entiers sans grand facteur premier, in: *Analytic Number Theory* (E. Fouvry, ed.), Tokyo 1988, Lecture Notes in Math. **1434**, Springer, Berlin, pp. 86–102.

(1991) Entiers sans grand facteur premier en progressions arithmétiques, *London Math. Soc.* (3) **63**, 449–494.

J. B. Friedlander,

(1972) On the number of ideals free of large prime divisors, *J. Reine Angew. Math.* **255**, 1–7.

(1973a) On the least kth power residue in an algebraic number field, *London Math. Soc.* (3) **26**, 19–34.

(1973b) Integers without large prime factors, *Nederl. Akad. Wetensch. Proc. Ser. A* **76**, 443–451.

(1976) Integers free from large and small primes, *Proc. London Math. Soc.* (3) **33**, 565–576.

(1981) Integers without large prime factors. II, *Acta Arith.* **39**, 53–57.

(1984a) Large prime factors in small intervals, in: *Topics in classical number theory,* Colloqu. Math. Soc. Janos Bolyai **34**, North-Holland, Amsterdam, 511–518.

(1984b) Integers without large prime factors. III, *Arch. Math.* **43**, 32–36.

(1985) Notes on Chebyshev's method, *Rocky Mountain J. Math.* **15**, 377–383.

(1989) Shifted primes without large prime factors, in: *Number Theory and Applications* (R. A. Mollin, ed.), Kluver, pp. 393–401.

J. B. Friedlander & J. C. Lagarias,

(1987) On the distribution in short intervals of integers having no large prime factor, *J. Number Theory* **25**, 249–273.

J. Galambos,

(1976) The sequence of prime divisors of integers, *Acta Arith.* **31**, 213–218.

J. R. Gillet,

(1970) On the largest prime divisor of ideals in fields of degree n, *Duke Math. J.* **37**, 589–600.

D. A. Goldston & K. S. McCurley,

(1988) Sieving the positive integers by large primes, *J. Number Theory* **28**, 94–115.

A. Granville,

(1989) On positive integers $\leq x$ with prime factors $\leq t \log x$, in: *Number Theory and Applications* (R. A. Mollin, ed.), Kluver, pp. 403–422.

(1991a) The lattice points of an n-dimensional tedrahedron, *Aequationes Math.* **41**, 234–241.

(1991b) On pairs of coprime integers with no large prime factors, *Expos. Math.* **9**, 335–350.

(1993a) Integers, without large prime factors, in arithmetic progressions. I, *Acta Math.* **170**, 255-273.

(1993b) Integers, without large prime factors, in arithmetic progressions. II, preprint.

J. L. Hafner,

(1993) On smooth numbers in short intervals under the Riemann Hypothesis, preprint.

H. Halberstam & H.-E. Richert,

(1971) Mean value theorems for a class of arithmetic functions, *Acta Arith.* **18**, 243–256; Add. & Corr., *ibid.* **28**, 107–110.

(1974) *Sieve Methods,* Academic Press, London, New York, San Francisco.

H. Halberstam & K. F. Roth,

(1951) On the gaps between consecutive k-free integers, *J. London Math. Soc.* **26**, 268–273.

G. Harman,

(1991) Short intervals containing numbers without large prime factors, *Math. Proc. Cambridge Philos. Soc.* **109**, 1–5.

D. G. Hazlewood,

(1973) On integers all of whose prime factors are small, *Bull. London Math. Soc.* **5**, 159–163.

(1975a) On k-free integers with small prime factors, *Proc. Amer. Math. Soc.* **52**, 40–44.

(1975b) Sums over positive integers with few prime factors, *J. Number Theory* **7**, 189–207.

(1975c) On sums over Gaussian integers, *Trans. Amer. Math. Soc.* **209**, 295–310.

(1977) On ideals having only small prime factors, *Rocky Mountain J. Math.* **7**, 753–768.

D. R. Heath-Brown,

(1987) Consecutive almost-primes, *J. Indian Math. Soc. (N. S.)* **52**, 39–49.

D. Hensley,

(1985) The number of positive integers $\leq x$ and free of prime divisors $> y$, *J. Number Theory* **21**, 286–298.

(1986) A property of the counting function of integers with no large prime factors, *J. Number Theory* **22**, 46–74.
(1987) The distribution of $\Omega(n)$ among numbers with no large prime factors, in: *Analytic Number Theory and Diophantine Problems* (A. Adolphson et al., eds.), Proc. of a Conf. at Oklahoma State University 1984, Birkhäuser, *Progress in Math.* **70**, 247–281.

A. Hildebrand,
(1984a) Integers free of large prime factors and the Riemann Hypothesis, *Mathematika* **31**, 258–271.
(1984b) On a problem of Erdős and Alladi, Monatsh. Math. 97, 119–124.
(1985a) Integers free of large prime factors in short intervals, *Quart. J. Math. (Oxford)* (2) **36**, 57–69.
(1985b) On a conjecture of A. Balog, *Proc. Amer. Math. Soc.* **95**, 517–523.
(1986a) On the number of positive integers $\leq x$ and free of prime factors $> y$, *J. Number Theory* **22**, 289–307.
(1986b) On the local behavior of $\Psi(x, y)$, *Trans. Amer. Math. Soc.* **297**, 729–751.
(1987) On the number of prime factors of integers without large prime divisors, *J. Number Theory* **25**, 81–106.
(1989) Integer sets containing strings of consecutive integers, *Mathematika* **36**, 60–70.

A. Hildebrand & G. Tenenbaum,
(1986) On integers free of large prime factors, *Trans. Amer. Math. Soc.* **296**, 265–290.
(1993) On a class of difference differential equations arising in number theory, *J. d'Analyse Math.* **61**, 145–179.

N. A. Hmyrova,
(1964) On polynomials with small prime divisors (Russian), *Dokl. Akad. Nauk SSSR* **155**, 1268–1271.
(1966) On polynomials with small prime divisors. II (Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* **30**, 1367–1372.

C. Hooley,
(1967) On the greatest prime factor of a quadratic polynomial, *Acta Math.* **117**, 2–16.
(1978) On the greatest prime factor of a cubic polynomial, *J. Reine Angew. Math.* **303/304**, 21–50.

A. Ivić,
(1981) Sum of reciprocals of the largest prime factor of an integer, *Arch. Math.* **36**, 57–61.
(1985a) *The Riemann zeta-function,* Wiley, New York.
(1985b) On squarefree numbers with restricted prime factors, *Stud. Sci. Math. Hungar.* **20**, 183–187.
(1987) On some estimates involving the number of prime divisors of an integer, *Acta Arith.* **49**, 21–33.

B. Hornfeck,
(1959) Zur Verteilung gewisser Primzahlprodukte, *Math. Ann.* **139**, 14–30.

A. Ivić and C. Pomerance,

(1984) Estimates for certain sums involving the largest prime factor of an integer, in: *Topics in classical number theory,* Colloq. Budapest 1981, Colloq. Math. Soc. Janos Bolyai 34, pp. 769–789.

A. Ivić and G. Tenenbaum,

(1986) Local densities over integers free of large prime factors, *Quart. J. Math. (Oxford)* (2) **37**, 401–417.

J. H. Jordan,

(1965) The divisibility of Gaussian integers by large Gaussian primes, *Duke Math. J.* **32**, 503–510.

M. Jutila,

(1974) On numbers with a large prime factor. II, *J. Indian Math. Soc. (N. S.)* **38**, 125–130.

D. E. Knuth & L. Trabb Pardo,

(1976) Analysis of a simple factorization algorithm, *Theoret. Comput. Sci.* **3**, 321–348.

U. Krause,

(1990) Abschätzungen für die Funktion $\Psi_K(x, y)$ in algebraischen Zahlkörpern, *Manuscripta Math.* **69**, 319–331.

M. Langevin,

(1975) Plus grand facteur premier d'entiers voisins, *C. R. Acad. Sci. Paris Sér. A-B* **281**, A491–A493.

D. H. Lehmer & E. Lehmer,

(1941) On the first case of Fermat's last theorem, *Bull. Amer. Math. Soc.* **47**, 139–142.

H. W. Lenstra

(1987) Factoring integers with elliptic curves, *Ann. Math.* **126**, 649–673.

B. V. Levin & U. Chariev,

(1986) Sums of multiplicative functions with respect to numbers with prime divisors from given intervals (Russian), *Dokl. Akad. Nauk. Tadzh. SSSR* **29**, 383–387.

B. V. Levin & A. S. Fainleib,

(1967) Application of some integral equations to problems in number theory, *Russian Math. Surveys* **22**, 119–204.

J. H. van Lint & H.-E. Richert,

(1964) Über die Summe $\sum_{n \le x,\, P(n) \le y} \mu^2(n)/\phi(n)$, *Nederl. Akad. Wetensch. Proc. Ser. A* **67**, 582–587.

S. P. Lloyd,

(1984) Ordered prime divisors of a random integer, *Ann. Probability* **12**, 1205–1212.

R. Lovorn,

(1992) *Rigorous, subexponential algorithms for discrete logarithms over finite fields*, Ph.D. thesis, Univ. of Georgia, 1992.

J. van de Lune,

(1974) The truncated average limit and one of its applications in analytic number theory, *Math. Centrum Amsterdam*, Report ZW 20/74.

H. L. Montgomery,

(1971) *Topics in Multiplicative Number Theory,* Lecture Notes in Math. **227**, Springer, Berlin.

P. Moree,

(1992) An interval result for the number field $\Psi(x, y)$ function, *Manuscripta Math.* **76**, 437–450.

(1993) *Psixyology and Diophantine equations*, Thesis, Leiden University.

P. Moree & C. L. Stewart,

(1990) Some Ramanujan–Nagell equations with many solutions, *Indag. Math. (N. S.)* **1**, 465-472.

Y. Motohashi,

(1976) An induction principle for the generalization of Bombieri's prime number theorem, *Proc. Japan Acad.* **52**, 273–275.

(1979) A note on almost primes in short intervals, *Proc. Japan Acad. Sci. Ser. A Math. Sci.* **55** (1979), 225–226.

M. Naimi,

(1988) Les entiers sans facteur carré $\leq x$ dont les facteurs premiers sont $\leq y$, in: *Groupe de travail en théorie analytique des nombres 1986–87,* Publ. Math. Orsay 88-01, pp. 69–76.

T. Nagell,

(1921) Généralisation d'un théorème de Tchébycheff, *J. Math. Pur. Appl.* (8) **4**, 343–356.

K. K. Norton,

(1968) Upper bounds for kth power coset representatives modulo n, *Acta Arith.* **15**, 161–179.

(1971) *Numbers with small prime factors and the least kth power non residue,* Memoirs of the Amer. Math. Soc. **106**.

A. M. Odlyzko,

(1985) Discrete logarithms in finite fields and their cryptographic significance, in *Advances in Cryptology: Proceedings of Eurocrypt '84,* (T. Beth, N. Cot, I. Ingemarsson, eds.), Lecture Notes in Computer Science **209**, Springer-Verlag.

(1993) Discrete logarithms and smooth polynomials, preprint.

C. Pomerance,

(1980) Popular values of Euler's function, *Mathematika* **27**, 84–89.

(1987) Fast, rigorous factorization and discrete logarithm algorithms, in: *Discrete Algorithms and Complexity (Kyoto, 1986)*, Academic Press, Boston.

K. Ramachandra,

(1969) A note on numbers with a large prime factor. *J. London Math. Soc.* (2) **1**, 303–306.

(1970) A note on numbers with a large prime factor. II, *J. Indian Math. Soc. (N. S.)* **34**, 39–48.

(1971) A note on numbers with a large prime factor. III, *Acta Arith.* **19**, 49–62.

K. Ramachandra & T. N. Shorey,

(1973) On gaps between numbers with a large prime factor, *Acta Arith.* **24**, 99–111.

V. Ramaswami,

(1949) On the number of positive integers less than x and free of prime divisors greater than x^c, *Bull. Amer. Math. Soc.* **55**, 1122–1127.

(1951) Number of integers in an assigned arithmetic progression, $\leq x$ and prime to primes greater than x^c, *Proc. Amer. Math. Soc.* **2**, 318–319.

R. A. Rankin,

(1938) The difference between consecutive prime numbers, *J. London Math. Soc.* **13**, 242–247.

J. B. van Rongen,

(1975) On the largest prime divisor of an integer, *Indag. Math.* **37**, 70–76.

E. Saias,

(1989) Sur le nombre des entiers sans grand facteur premier, *J. Number Theory* **32**, 78–99.

(1992) Entiers sans grand ni petit facteur premier. I, *Acta Arith.* **61**, 347–374.

E. J. Scourfield,

(1991) On some sums involving the largest prime divisor of n, *Acta Arith.* **59**, *no* 4, 339–363.

T. N. Shorey,

(1973) On gaps between numbers with a large prime factor II, *Acta Arith.* **25**, 365–373.

H. Smida,

(1991) Sur les puissances de convolution de la fonction de Dickman, *Acta Arith.* **59**, *no* 2, 123–143.

(1993) Valeur moyenne des fonctions de Piltz sur les entiers sans grand facteur premier, *Acta Arith.* **63**, *no* **1**, 21–50.

W. Specht,

(1949) Zahlenfolgen mit endlich vielen Primteilern, *S.-B. Bayer. Akad. Wiss. Math. Nat. Abt.* 1948, 149–169.

G. Tenenbaum,

(1985) Sur les entiers sans grand facteur premier, in: *Séminaire de Théorie des Nombres*, Bordeaux 1984–85, Univ. Bordeaux I, Talence, 12 pp.

(1988) La méthode du col en théorie analytique des nombres, in: *Séminaire de Théorie des Nombres* (C. Goldstein, ed.), Paris 1985–86, Birkhäuser, *Progress in Math.* **75**, 411–441.

(1990a) *Introduction à la théorie analytique et probabiliste des nombres,* Publ. Inst. Elie Cartan, Vol. 13, Univ. Nancy 1.

(1990b) Sur un problème d'Erdős et Alladi, in: *Séminaire de Théorie des Nombres* (C. Goldstein, ed.), Paris 1988–89, Birkhäuser, *Progress in Math.* **91**, 221–239.

(1990c) Sur une question d'Erdős et Schinzel II, *Inventiones Math.* **99**, 215–224.

(1993) Cribler les entiers sans grand facteur premier, in: R.C.Vaughan (ed.) Theory and applications of numbers without large prime factors, *Phil. Trans. Royal Soc. London, series A*, to appear.

R. Tijdeman,

(1972) On the maximal distance of numbers with a large prime factor, *J. London Math. Soc.* (2) **5**, 313–320.

(1973) On integers with many small prime factors, *Compositio Math.* **26**, 319–330.

(1974) On the maximal distance between integers composed of small primes, *Compositio Math.* **28**, 159–162.

N. M. Timofeev,

(1977) Polynomials with small prime divisors (Russian), *Taškent Gos. Univ. Naučn. Trudy* **548** *Voprosy Mat.*, 87–91.

E. C. Titchmarsh and D. R. Heath-Brown,

(1986) *The theory of the Riemann zeta function,* Second Edition, Oxford Univ. Press, Oxford.

J. Turk,

(1980) Prime divisors of polynomials at consecutive integers, *J. Reine Angew. Math.* **319**, 142–152.

(1982) *Products of integers in short intervals,* Report 8228M, Econometric Institut, Erasmus University, Rotterdam.

R. C. Vaughan,

(1989) A new iterative method in Waring's problem, *Acta Math.* **162**, 1–71.

A. M. Vershik,

(1987) The asymptotic distribution of factorizations of natural numbers into prime divisors, *Soviet Math. Dokl.* **34**, 57–61.

I. M. Vinogradov,

(1926) On a bound for the least nth power non-residue (Russian), *Izv. Akad. Nauk SSSR* **20**, 47–58; English transl.: *Trans. Amer. Math. Soc.* **29** (1927), 218–226.

Y. Wang,

(1964) Estimation and application of a character sum (Chinese), *Shuxue Jinzhan* **7**, 78–83.

R. Warlimont,

(1990) Sieving by large prime factors, *Monatshefte Math.* **109**, 247–256.

(1991) Arithmetical semigroups, II: Sieving by large and small prime elements. Sets of multiples, *Manuscripta Math.* **71**, 197–221.

F. Wheeler,

(1990) Two differential-difference equations arising in number theory, *Trans. Amer. Math. Soc.* **318**, 491–523.

D. Wolke,

(1971) Polynomial values with small prime divisors, *Acta Arith.* **19**, 327–333.

(1973) Über die mittlere Verteilung der Werte zahlentheoretischer Funktionen auf Restklassen. I, *Math. Ann.* **202**, 1–205.

T. Wooley,

(1992) Large improvements in Waring's problem, *Annals Math.*(2) **135**, *no* **1**, 135–164.

T. Z. Xuan,

(1988) On a problem of Erdős and Ivić, *Publ. Inst. Math. (Beograd) (N. S.)* **43** (57), 9–15.

(1989a) On sums involving reciprocals of certain large additive functions. I, *Publ. Inst. Math. (Beograd) (N. S.)* **45** (59), 41–55.

(1989b) On sums involving reciprocals of certain large additive functions. II, *Publ. Inst. Math. (Beograd) (N. S.)* **46** (60), 25–32.

(1990) The average of $d_k(n)$ over integers free of large prime factors, *Acta Arith.* **55**, 249–260.

(1991) The average of the divisor function over integers without large prime factors, *Chinese Ann. of Math. Ser. A* **12**, 28–33.

(1993) On the asymptotic behavior of the Dickman-de Bruijn function, *Math. Ann.* **297**, 519–533.

Adolf Hildebrand
Department of Mathematics
University of Illinois
Urbana, IL 61801
U.S.A.

Gérald Tenenbaum
Département de Mathématiques
Université Nancy 1
BP 239
F-54506 Vandœuvre Cedex
France