ON SUMS INVOLVING RECIPROCALS OF THE LARGEST PRIME FACTOR OF AN INTEGER

P. Erdős, A. Ivić and C. Pomerance*,
Budapest, Beograd and Athens, Ga, USA

Abstract. Sum of reciprocals of $P(n)$, the largest prime factor of $n$, is precisely evaluated asymptotically. Asymptotic formulas for some related sums, involving the function $\Omega(n)$ and $\omega(n)$ (the number of distinct and the total number of prime factors of $n$) are also derived.

§1. Introduction and statement of results

Let as usual $\omega(n)$ and $\Omega(n)$ denote the number of distinct prime factors of $n$ and the total number of prime factors of $n$, respectively. Let $P(n)$ denote the largest prime factor of an integer $n \geq 2$, and let $P(1) = 1$. Several results involving sums of reciprocals of $P(n)$ and some related additive functions were obtained recently in [4], Ch. 6, [6], [7], [9] and [10]. Thus it was shown in [9] that

$$\sum_{n \leq x} \frac{1}{P(n)} = x \exp \left\{ -\left( 2 \log \frac{x}{\log \log x} \right)^{1/2} + O \left( (\log x \log \log \log x)^{1/2} \right) \right\} \quad (1.1)$$

The proof of this result depended on estimates for $\psi(x, y)$, the number of positive integers $n \leq x$ with $P(n) \leq y$. The connection is seen via the easy identity

$$\sum_{n \leq x} \frac{1}{P(n)} = 1 + \sum_{p \leq x} p^{-1} \psi(xp^{-1}, p), \quad (1.2)$$

where $p$ denotes a general prime throughout the paper. By using a better estimate for $\psi(x, y)$ (see [3]), the result (1.1) was slightly sharpened and more general sums were estimated in [10], namely

$$S_r(x) = \sum_{n \leq x} \frac{1}{P^r(n)}, \quad T_r(x) = \sum_{n \leq x, P^r(n) \mid x} \frac{1}{P^r(n)},$$


Key words and phrases: The largest prime factor of an integer, the number of prime factors of an integer, asymptotic formulas for summatory functions, the number of integers not exceeding $x$ all of whose prime factors do not exceed $y$.

* Research supported in part by a grant from the National Science Foundation.
where \( r \geq 0 \) is an arbitrary, fixed real number. It was proved in [10] that
\[
S_r(x) = x \exp \left\{ - (2r)^{1/2} \left( \log x \log x \right)^{1/2} \left( 1 + g_{r-1}(x) + O \left( \log^3 x/\log^2 x \right) \right) \right\}
\]
and
\[
T_r(x) = x \exp \left\{ - (2r+2)^{1/2} \left( \log x \log x \right)^{1/2} \left( 1 + g_r(x) + O \left( \log^3 x/\log^2 x \right) \right) \right\},
\]
where \( \log_k x = \log (\log^{k-1} x) \) is the \( k \)-fold iterated logarithm and
\[
g_r(x) = \frac{\log x + \log (1 + r) - 2 - \log 2}{2 \log x} \left( 1 + \frac{2}{\log x} \right) - \frac{(\log x + \log (1 + r) - \log 2)^2}{8 \log^2 x}.
\]

Recently H. Maier [11] and A. Hildebrand [8] obtained independently much better results concerning \( t(x, y) \), which may be used in connection with our problems. It is now possible to obtain asymptotic formulas for the sums \( S_r(x) \) and \( T_r(x) \). We shall work out the details only for the sum in (1.1), namely \( S_1(x) \). The other sums can be handled by the same method. We prove

**Theorem 1.**

\[
\sum_{n \leq x} 1/P(n) = x \delta(x) \left( 1 + O \left( \left( \frac{\log \log x}{\log x} \right)^{1/2} \right) \right),
\]

where
\[
\delta(x) = \int_2^x \rho \left( \frac{\log t}{\log x} \right) (\log t)^{-2} dt,
\]

and \( \rho(u) \) is the continuous solution to the differential delay equation
\[
up' (u) = - \rho (u - 1)
\]
with the initial condition \( \rho(u) = 1 \) for \( 0 \leq u \leq 1 \).

Here \( \rho(u) \) is the so-called Dickman-de Bruijn function, for which the latter [1] obtained the estimate
\[
\rho(u) = \exp \left\{ - u \left( \log u + \log_2 u - 1 + \frac{\log_2 u}{\log u} - \frac{1}{\log u} + O \left( \frac{\log^2 u}{\log^2 u} \right) \right) \right\},
\]

(1.8)

By comparing (1.3) and (1.5) (or by direct evaluation with the aid of (1.8)) we find that
\[
\delta(x) = \exp \left\{ - (2 \log x \log x)^{1/2} \left( 1 + g_0(x) + O \left( \log^3 x/\log^2 x \right) \right) \right\}.
\]
Although $\delta(x)$ is a fairly complicated function, (1.9) gives for most purposes a sufficiently sharp approximation. Moreover $\delta(x)$ is slowly oscillating, i.e. for any constant $C > 0$ we have

$$\lim_{x \to \infty} \delta(Cx)/\delta(x) = 1,$$

(1.10)

which is obtained in §3. As a corollary of (1.5) and (1.10) we have, for example,

$$\sum_{n \leq x} 1/P(n) \sim \sum_{x < n \leq 2x} 1/P(n),$$

which seems to be difficult to obtain without using finer information about $\psi(x, y)$. The notation $f(x) \sim g(x)$ means as usual that

$$\lim_{x \to \infty} f(x)/g(x) = 1,$$

and other notation used throughout the paper is also standard. For example, $f(x) = O(g(x))$ and $f(x) \ll g(x)$ both mean that $|f(x)| \leq Cg(x)$ for some absolute $C > 0$ and $x \geq x_0$, while $f(x) = o(g(x))$ means that

$$\lim_{x \to \infty} f(x)/g(x) = 0.$$

It seems interesting to investigate how much the sum in (1.1) changes when $1/P(n)$ is replaced by $\omega(n)/P(n)$ or $\Omega(n)/P(n)$. This problem has already been investigated in [7], where it was shown that

$$\left(\frac{\log x}{\log \log x}\right)^{1/2} \sum_{n \leq x} 1/P(n) \ll \sum_{n \leq x} \Omega(n)/P(n) \ll (\log x \log \log x)^{1/2} \sum_{n \leq x} 1/P(n),$$

and the method would yield the same result if $\Omega(n)$ is replaced by $\omega(n)$. By using Theorem 1, we prove

**THEOREM 2.** There is a positive constant $c$ such that

$$\sum_{n \leq x} (\Omega(n) - \omega(n))/P(n) \sim cx \delta(x),$$

(1.11)

$$\sum_{n \leq x} \Omega(n)/P(n) \sim \sum_{n \leq x} \omega(n)/P(n) \sim (2 \log x / \log \log x)^{1/2} x \delta(x),$$

(1.12)

where $\delta(x)$ is defined by (1.6).

The results (1.5) and (1.12) imply that in a certain sense the main contribution to the sums in (1.12) comes from those $n \leq x$ with about $(2 \log x / \log \log x)^{1/2}$ prime factors. This fact is interesting in view of the classical result of G. H. Hardy and S. Ramanujan (see [12]) that both the normal and average order of $\omega(n)$ and $\Omega(n)$ is $\log \log n$.

Theorem 2 shows that sums of $\omega(n)/P(n)$ and $\Omega(n)/P(n)$ behave similarly. On the other hand, quite a different situation arises when one estimates sums of $P(n) - \omega(n)$ and $P(n) - \Omega(n)$. The corresponding summatory functions turn out to be completely different, as shown by
THEOREM 3.

\[ \sum_{n \leq x} P(n)^{-\omega(n)} = \exp \left\{ (4 + o(1)) \frac{(\log x)^{1/2}}{(\log \log x)} \right\}, \]  
(1.13)

\[ \sum_{n \leq x} P(n)^{-\Theta(n)} = \log \log x + D + O(1/\log x). \]  
(1.14)

Here \( D > 0 \) denotes an absolute constant that is effectively computable.

It seems interesting to compare (1.13) with the estimate

\[ \sum_{n \leq x} 1/\omega(n) = \exp \left\{ (2 \sqrt{2} + o(1)) \frac{(\log x/\log \log x)^{1/2}}{} \right\}, \]  
(1.15)

where \( \omega(n) = \prod_{p \mid n} p \) is the largest square-free divisor of \( n \). This result is due to N. G. de Bruijn [2] and was sharpened by W. Schwarz [13].

Since evidently

\[ \omega(n) \leq P(n)^{\omega(n)}, \]

the sum in (1.13) is majorized by the sum in (1.15), but it turns out that even the logarithms of these sums are of a different order of magnitude. The sum in (1.13) is more difficult to estimate than the sum in (1.14), where the main contribution comes from primes.

For our last result, let \( Q(n) \) denote the largest prime power which divides \( n \geq 2 \), and let \( Q(1) = 1 \). One naturally expects sums of \( 1/P(n) \) and \( 1/Q(n) \) to behave similarly, and this is precisely what is established in

THEOREM 4. There is a constant \( C > 0 \) such that

\[ \sum_{n \leq x} 1/Q(n) = \left\{ 1 + O(\exp(-C(\log x \log \log x)^{1/2})) \right\} \sum_{n \leq x} 1/P(n). \]  
(1.16)

§2. Proof of Theorem 1

We begin by establishing the identity (1.2). We have

\[ \sum_{n \leq x} 1/P(n) = 1 + \sum_{p \leq x} 1/p \sum_{n \leq x, P(n) = p} 1 = 1 + \sum_{p \leq x} 1/p \sum_{m \leq x, p | m, P(m) \leq p} 1 = \]

\[ = 1 + \sum_{p \leq x} p^{-1} \psi(x/p, p), \]

so that (1.2) holds. To facilitate notation, let from now on

\[ L = L(x) = \exp \{(\log x \log \log x)^{1/2}\}. \]

Using (1.2) and following the proof of (1.1) given in [9] we see that the contribution from the primes \( p \) with \( p < L^{1/2} \) or \( p > L \) is at most

\[ xL^{-3/2 + o(1)} \quad (x \to \infty). \]
Thus in view of (1.1) it suffices to consider only those values of $p$ with $L^{1/2} \leq p \leq L$.

We now state the new result of A. Hildebrand [8] (H. Maier's theorem [11] is slightly weaker) mentioned in the introduction.

THEOREM (A. Hildebrand). The estimate

$$
\psi (x, x^{1/u}) = x \rho (u) \left( 1 + O \left( \frac{u \log (u+1)}{\log x} \right) \right)
$$

holds uniformly in the range

$$
x \geq 3, \ 1 \leq u \leq \log x / (\log \log x)^{5/3 + \varepsilon},
$$

where $\varepsilon$ is any fixed positive number.

The importance of this result lies in the wide range for $u$. By N. G. de Bruijn [1] (2.1) was known to hold for $1 \leq u \leq (\log x)^{3/5} - \varepsilon$, but this range is not sufficient for our purposes.

Applying (2.1) to the $\psi (x/p, p)$ for $L^{1/2} \leq p \leq L$ we obtain the uniform estimate

$$
\psi (x/p, p) = \frac{x}{p} \rho \left( \frac{\log x}{\log p} - 1 \right) \left( 1 + O \left( \left( \frac{\log \log x}{\log x} \right)^{1/2} \right) \right).
$$

Thus from the above comments we have

$$
\sum_{n \leq x} \frac{1}{P(n)} = \left( 1 + O \left( \left( \frac{\log \log x}{\log x} \right)^{1/2} \right) \right) \sum_{L^{1/2} \leq p \leq L} \frac{x p^{-2} \rho \left( \frac{\log x}{\log p} - 1 \right)}
$$

$$
= \left( 1 + O \left( \left( \frac{\log \log x}{\log x} \right)^{1/2} \right) \right) \int_{L^{1/2}}^{L} \frac{x t^{-2} \rho \left( \frac{\log x}{\log t} - 1 \right)}{\log t} \, d\pi (t) \tag{2.2}
$$

$$
= \left( 1 + O \left( \left( \frac{\log \log x}{\log x} \right)^{1/2} \right) \right) \int_{L^{1/2}}^{L} \frac{x}{t^2 \log t} \rho \left( \frac{\log x}{\log t} - 1 \right) \, dt.
$$

Using (1.7) the last integral becomes

$$
\int_{L^{1/2}}^{L} \frac{-x \log x}{t^2 \log^2 t} \rho' \left( \frac{\log x}{\log t} \right) \, dt = \int_{L^{1/2}}^{L} \frac{x}{t} \, d\rho \left( \frac{\log x}{\log t} \right)
$$

$$
= \frac{x}{L} \rho \left( \frac{\log x}{\log L} \right) - \frac{x}{L^{1/2}} \rho \left( \frac{\log x}{\log L^{1/2}} \right) + \int_{L^{1/2}}^{L} \frac{x}{t^2} \rho \left( \frac{\log x}{\log t} \right) \, dt. \tag{2.3}
$$
Using (1.8) we see that the right-hand side of (2.3) is equal to

\[ x \delta (x) + O \left( xL^{-3/2 + \varepsilon} \right) \]

for any \( \varepsilon > 0 \), hence Theorem 1 follows then from (2.2) and (2.3).

§3. Proof of Theorem 2

A positive integer \( m \) is called square-full if \( p^2 \mid m \) for every prime \( p \mid m \). Let \( s(n) \) denote the largest square-full divisor of \( n \). Then we have

\[ \Omega(n) - \omega(n) = \Omega(s(n)) - \omega(s(n)). \]  

To estimate the sum in (1.11) we first show that those \( n \) with \( s(n) > \log^3 x \) or with \( p^2(n) \mid n \) contribute only \( o(x \delta(x)) \) to the sum. To do this first note that, for \( n \leq x \),

\[ \Omega(n) - \omega(n) \leq \frac{\log x}{\log 2} - 1, \]

hence

\[ \sum_{n \leq x, p^2(n) \mid n} (\Omega(n) - \omega(n))/P(n) \ll T_1(x) \log x = xL^{-2 + o(1)}, \]  

where we used (1.4) with \( r = 1 \).

Next note that there are \( \sim Cx^{1/2} \) square-full integers not exceeding \( x \). Therefore using partial summation we obtain

\[ \sum_{n \leq x, s(n) > L^2} (\Omega(n) - \omega(n))/P(n) \ll \log x \sum_{n \leq x, s(n) > L^3} 1 \]

\[ = \log x \sum_{s > L^3} \sum_{n \leq x, s(n) = s} 1 \leq \log x \sum_{s > L^3} 1/s \]

\[ = xL^{-3/2 + o(1)} = o(x \delta(x)), \]

where \( \sum' \) denotes a sum over square-full integers.

To estimate the sum in (1.11) for \( s(n) < L^3 \), we need the following

LEMMA. Uniformly for \( x \geq 3 \) and \( 1 \leq s \leq L^3 \), we have

\[ \delta(x/s) \leq (1 + o(1))s^{2(\log \log x/\log x)^{0.5}} \delta(x). \]

Proof. We use the following result, which is Lemma 1–(v) in Hildebrand [8]:

\[ -\rho'(u)/\rho(u) \leq \log (u \log^2 u) \quad (u \geq e^4). \]

Integrating this inequality over \([u - \lambda, u] \) we obtain, for \( 0 \leq \lambda \leq u - e^4 \),

\[ \rho(u - \lambda)/\rho(u) \leq (u \log^2 u)^{\lambda}. \]
Thus for \(1 \leq s \leq L^3\) and for large \(x\) we have
\[
\delta(x/s) = \int_{2}^{x/s} t^{-2} \rho\left(\frac{\log x}{\log t} - \frac{\log s}{\log t}\right) dt
\]
\[
= (1 + o(1)) \int_{L^{1/2}}^{x/s} t^{-2} \rho\left(\frac{\log x}{\log t} - \frac{\log s}{\log t}\right) \rho\left(\frac{\log x}{\log t}\right) dt
\]
\[
\leq (1 + o(1)) \int_{L^{1/2}}^{x/s} t^{-2} \rho\left(\frac{\log x}{\log t}\right) \left(\frac{\log x}{\log t}\right)\log x \log t \log \frac{x}{\log L^{1/2}} dt
\]
\[
\leq (1 + o(1)) \left(\frac{\log x}{\log L^{1/2}}\right)^2 \left(\frac{\log x}{\log L^{1/2}}\right)^2 \delta(x)
\]
\[
\leq (1 + o(1)) (\log x)^{\log x / \log L^{1/2}} \delta(x) = (1 + o(1)) s^2 (\log \log x / \log x)^{1/2} \delta(x),
\]
which establishes the lemma.

By Theorem 1 and the Lemma, we have
\[
\sum_{n < x, \log^2 x < s(n) < L^3} (\Omega(n) - \omega(n)) \ll \log x \sum_{n < x, \log^2 x < s(n) < L^3} 1/P(n)
\]
\[
\leq \log x \sum_{\log^2 x < s < L^3} \sum_{n < x, s(n) < L} 1/P(n)
\]
\[
\leq \log x \sum_{\log^2 x < s < L^3} \sum_{m < x/s} 1/P(m) \ll \log x \sum_{\log^2 x < s < L^3} (x/s) \delta(x/s)
\]
\[
\ll x \delta(x) \log x \sum_{s \geq \log^2 x} s^{-1 + 2(\log \log x / \log x)^{1/2}} \ll x \delta(x) \log^{-1/2} x.
\]

Therefore to show (1.11) it will be sufficient to restrict the sum to those \(n \leq x\) for which \(P^2(n)/p\) and \(s(n) \leq \log^3 x\). Also, as in Section 2, we may assume \(L^{1/2} \leq P(n) \leq L\). If \(\sum\) denotes a sum over integers \(k\) with the restrictions \(P^2(k)/k\) and \(P(k) \geq L^{1/2}\), then by (3.1) we have
\[
\sum_{n \leq x, s(n) \leq \log^3 x} (\Omega(n) - \omega(n))/P(n) = \sum_{s \leq \log^3 x} \sum_{n \leq x, s(n) < s} (\Omega(n) - \omega(n))/P(n)
\]
\[
= \sum_{s \leq \log^3 x} (\Omega(s) - \omega(s)) \sum_{m \leq x, s(m) - s} \mu^2(m)/P(m)
\]
\[
= \sum_{s \leq \log^3 x} (\Omega(s) - \omega(s)) \sum_{m \leq x, s(m) - s} 1/P(m) \sum_{d \mid m} \mu(d)
\]
\[
= \sum_{s \leq \log^3 x} (\Omega(s) - \omega(s)) \sum_{d \mid m \leq x, s(m) - s} \mu(d) \sum_{k \leq x(d, s) \leq \log^2 x} 1/P(k).
\]
Here \( z(s) = \prod_{p \mid s} p \) and the last equality follows from the fact that \( d^2 \mid m \chi(s) \) if and only if \( m \) is of the form \( kd^2/(d, \chi(s)) \) (and so \( P(m) = P(k) \) since \( P^2(m) m, P(m) \geq L^{1/2} \) and \( s \leq \log^3 x \)).

The last and innermost sum in (3.5) is majorized by \( x/d^2 \), so that the contribution to (3.5) by those \( d > L^{3/2}/\log^3 x \) is \( O(xL^{-3/2+\alpha(1)}) \) and is thus negligible. For \( \log^3 x \leq d \leq L^{3/2}/\log^3 x \) and \( s \leq \log^3 x \) we have

\[
\log^6 x \leq \frac{sd^2/(d, \chi(s))}{L^3},
\]

so that from the proof of (3.4), the last sum in (3.5) for such a value of \( d \) is

\[
O\left(x\delta(x)\frac{d^{-2}}{(\log \log x)^{1/2}}\right).
\]

Thus the contribution to (3.5) from these values of \( d \) is \( \ll x\delta(x)\log^{-1/2} x \), which is also negligible.

Hence we may restrict attention in (3.5) to those \( d \) with \( d < \log^3 x \). For such \( d \)'s we have

\[
\frac{sd^2/(d, \chi(s))}{\log^3 x},
\]

and so by Theorem 1 and (3.2) the last sum in (3.5) for these \( d \)'s is uniformly

\[
(1+o(1)) \frac{x(d, \chi(s))}{sd^2} \delta \left( \frac{x(d, \chi(s))}{sd^2} \right).
\]

(3.6)

Note that from the definition of \( \delta(x) \) it is possible to show that \( \delta(x) \) is decreasing for \( x \geq x_0 \), so that using the Lemma we have \( \delta(x/t) = (1+o(1)) \delta(x) \) uniformly for \( 1 \leq t \leq \log^3 x \). In particular, this remark implies that (1.10) holds. Thus (3.6) is

\[
(1+o(1)) \frac{x(d, \chi(s))}{sd^2} \delta(x).
\]

Putting this estimate in (3.5) we have

\[
\sum_{n \leq x, s(n) \leq \log^3 x} \frac{(\Omega(n) - \omega(n))/P(n)}{P(n)} =
\]

\[
= (1+o(1)) x\delta(x) \sum_{s \leq \log^3 x} \frac{\Omega(s) - \omega(s)}{s} \sum_{d < \log^3 x} \frac{(d, a(s)) \mu(d)}{sd^2}
\]

\[
= (1+o(1)) \frac{6}{\pi^2} x\delta(x) \sum_{s \leq \log^3 x} \frac{\Omega(s) - \omega(s)}{s} \prod_{p \mid s} \frac{p}{p+1}
\]

\[
= (1+o(1)) \frac{6}{\pi^2} x\delta(x) \sum_{s=1}^{\infty} \frac{\Omega(s) - \omega(s)}{s} \prod_{p \mid s} \frac{p}{p+1},
\]
On sums involving reciprocals of the largest prime factor of an integer

since by multiplicativity
\[
\sum_{d=1}^{\infty} \mu(d) (d, \chi(s)) d^{-2} = \prod_p (1 + \mu(p) (p, \chi(s)) p^{-2})
\]
\[
= \prod_{p \not| s} \left(1 - \frac{1}{p^2}\right) \prod_{p | s} \left(1 - \frac{p}{p^2}\right) = \frac{1}{\zeta(2)} \prod_{p \not| s} \left(1 - \frac{1}{p}\right) \left(1 - \frac{1}{p^2}\right)^{-1} = \frac{6}{\pi^2} \prod_{p \not| s} \frac{p}{p+1}.
\]

Thus from this calculation and (3.2), (3.3) and (3.4) we obtain (1.11) with the constant
\[
c = \frac{6}{\pi^2} \sum_{s=1}^{\infty} \frac{\Omega(s) - \omega(s)}{s} \prod_{p \not| s} \frac{p}{p+1} > 0.
\]

A more careful analysis shows that in (1.11) can be replaced by
\[
1 + O (\frac{1}{\log \log x}^{3/2} \log^{-1/2} x),
\]
since for \(1 < t < \log^2 x\) we obtain by following the proof of the Lemma
\[
\delta(x/t) = (1 + O (\frac{1}{\log \log x}^{3/2} \log^{-1/2} x)) \delta(x).
\]

We now turn our attention to the proof of (1.12). From the proof of (1.1) or from the proof of Theorem 1 it may be seen that only the values of \(n \leq x\) for which
\[
L^{\sqrt{2}/2 - \varepsilon} < P(n) < L^{\sqrt{2}/2 + \varepsilon}
\]
for any fixed \(\varepsilon > 0\) make a non-negligible contribution to the sums in (1.12).

We next note that if \(n \leq x/L^2\), then
\[
\sum_{n \leq x/L^2} \omega(n)/P(n) \leq \sum_{n \leq x/L^2} \Omega(n)/P(n) \ll xL^{-2} \log x = o(x\delta(x)),
\]
so we may assume that
\[
x/L^2 < n < x.
\]

Thus combining (3.7), (3.8) and using the trivial inequality \(n \leq (P(n))^{\Omega(n)}\), we obtain
\[
\Omega(n) \geq (\sqrt{2} - 2\varepsilon) (\log x/\log \log x)^{1/2},
\]
and so for \(x \geq x_0 (\varepsilon)\) it follows that
\[
\sum_{n \leq x} \Omega(n)/P(n) \geq (\sqrt{2} - 3\varepsilon) (\log x/\log \log x)^{1/2} \sum_{n \leq x} 1/P(n).
\]

To estimate the sum of \(\omega(n)/P(n)\) from above we use the classical elementary inequality of G. H. Hardy and S. Ramanujan ([12], p. 265):
\[
\sum_{n \leq x, \omega(n) = k} 1 \leq \frac{Ex (\log \log x + F)^k}{k! \log x},
\]
where $E, F > 0$ are absolute constants. Using (3.10) we obtain that the number of $n \leq x$ with

$$\omega(n) > (\sqrt{2} + 5\varepsilon) (\log x / \log \log x)^{1/2} \quad (3.11)$$

is at most $x L - \sqrt{2} - 2\varepsilon$. Thus the sum of $\omega(n)/P(n)$ for $n \leq x$ and such that (3.7) and (3.11) both hold is at most

$$x \log x L - \sqrt{2} = o(x \delta(x)).$$

On the other hand, the sum of $\omega(n)/P(n)$ for $n \leq x$ and (3.11) failing is clearly at most

$$(\sqrt{2} + 5\varepsilon) (\log x / \log \log x)^{1/2} \sum_{n \leq x} 1/P(n). \quad (3.12)$$

Combining (3.9) and (3.12) with Theorem 1 and (1.11), we have (1.12), completing the proof of Theorem 2. We finally remark that we can prove (1.12) without using the Hildebrand-Maier result on $\psi(x, y)$. However this result seems to be essential for the proof of (1.11).

§4. Proof of Theorem 3

We denote by $\sum(x)$ the sum in (1.13) and proceed first to derive the lower bound of the correct order of magnitude. In what follows $p$ will always denote primes and $p_r$ will denote the $r$-th prime. Let $A$ be a large positive integer, and consider integers $m \leq x$ such that $\omega(m) = k$ and $P(m) \leq p_{(A+1)k}$, where $k = k(x)$ is an integer which will be suitably determined later. If

$$m = p_{i_1}^{a_{i_1}} \cdots p_{i_k}^{a_{i_k}} \quad (4.1)$$

is the canonical decomposition of $m$, then there are $\binom{(A+1)k}{k}$ ways we can choose $p_{i_1} \ldots p_{i_k} = \alpha(m)$. Once $\alpha(m)$ is fixed, we can choose the exponents $a_{i_1}, \ldots, a_{i_k}$ in (4.1) by considering positive integer solutions of

$$a_1 \log p_{i_1} + \ldots + a_k \log p_{i_k} \leq \log x.$$

Note that the number of positive, integer solutions of the above inequality is certainly not less than the corresponding number of solutions of

$$a_1 \log p_{(A+1)k} + \ldots + a_k \log p_{(A+1)k} \leq \log x,$$

which is $\binom{v}{k}$, where $v = \lceil \log x / \log p_{(A+1)k} \rceil$. Therefore

$$\sum(x) \geq \sum_{m \leq x} (P(m))^{-k} \geq \binom{v}{k} \binom{Ak+k}{k} P_{(A+1)k}^{Ak+k}. \quad (4.2)$$
To evaluate binomial coefficients which appear in the last inequality we shall use Stirling's formula in the form

\[ \log(n!) = \exp\left(\log n \cdot \log n - n + \log \sqrt{2\pi n} + o(1)\right). \quad (n \to \infty) \quad (4.3) \]

When \( k, v \to \infty \) and \( k = o(v) \), (4.3) gives

\[
\log\binom{Ak + k}{k} = k (A + 1) \log k (A + 1) - k (A + 1) - k A \log (k A) + k A - k \log k + k + O(\log k),
\]

\[
= k (A + 1) \log (A + 1) - k A \log A + O(\log \log x),
\]

\[
\log\binom{v}{k} = v \log v - k \log k - (v - k) \log (v - k) + O(\log v)
\]

\[
= k \log v - k \log k + k + o(k) + O(\log \log x).
\]

From the prime number theorem we have

\[
p_r = r \left( \log r + O(\log \log r) \right), \quad \log p_r = \log r + \log \log r + O(\log \log r / \log r),
\]

\[
\log \log p_r = \log \log r + O(\log \log r / \log r).
\]

Hence from (4.2) we obtain

\[
\sum (x) \geq \exp\left( O(\log_2 x) \right) \exp \left[ k \log_2 x - k \log k + k + o(k) - k \log p_{(4A+1)k} + k ((A + 1) \log (A + 1) - A \log A) - k \log p_{(4A+1)k} \right] = \exp \left[ k \log_2 x - 2k \log k - 2k \log_2 k + 2k + o(k) + O(k/A) + O(k \log A / \log k) + O(\log_2 x) \right],
\]

where the \( o \) and \( O \)-notation is uniform in \( A \).

Suppose now that \( \varepsilon > 0 \) is given. Then (4.4) implies, for \( k \geq k_0(\varepsilon) \),

\[
\sum (x) \geq \exp\left( O(\log_2 x) \right) \exp \left\{ k \log \log x - 2 \log k - 2 \log \log k + 2k + R(k, A) \right\},
\]

where for some \( B > 0 \)

\[
|R(k, A)| < \varepsilon / 3 + B/A + B \log A / \log k.
\]

At this point we choose \( A = \left[ 3B \varepsilon^{-1} + 1 \right] \). Then for \( k \geq k_1(\varepsilon) \) we have \( B \log A / \log k < \varepsilon / 3 \), hence for \( k \geq \max(k_0, k_1) \)

\[
\sum (x) \geq \exp\left( O(\log_2 x) \right) \exp (f(k)),
\]

where

\[
f(k) = k \log_2 x - 2k \log k - 2k \log_2 k + (2 - \varepsilon) k,
\]

so that

\[
f'(k) = \log_2 x - 2 \log k - 2 \log_2 k - 2 / \log k - \varepsilon.
\]
and \( f''(k) < 0 \) for \( k \) sufficiently large. This means that \( f(k) \) attains its maximum for \( k = k_2 \), where \( k_2 = k_2(x) \) is the solution of \( f''(k) = 0 \), which yields

\[
\log x = k_2^2 \log^2 k_2 \exp (\varepsilon + 2/\log k_2), \quad \log x = \left( \frac{1}{4} + o(1) \right) k_2^2 (\log \log x)^2,
\]

hence

\[
k_2 = k_2(x) = (2 + o(1)) \log^{1/2} x (\log \log x)^{-1}, \quad (x \to \infty).
\]

Taking \( k = k_3 = \lceil k_2(x) \rceil \) we see that \( k = o(n) \) holds (this is needed in the evaluation of the integrals and therefore from (4.5) and (4.6) we obtain

\[
\sum (x) \geq \exp (O(\log_2 x)) \exp ((2 + o(1)) k_3) \geq \exp ((4 - \varepsilon_1) \log^{1/2} x (\log \log x)^{-1}),
\]

for \( x \geq x_1(\varepsilon_1), \varepsilon_1 = \varepsilon_1(\varepsilon) \) and \( \lim_{\varepsilon \to 0} \varepsilon_1 = 0 \). Therefore we have proved the lower bound for \( \sum (x) \).

In proving the upper bound for \( \sum (x) \) we shall make use of

\[
\psi(x, p_t) = \exp (t \log_2 x - t \log t - t \log_2 t + t + o(t)),
\]

(4.8)

Actually we need only the upper bound implied by (4.8), but the lower bound follows from a simple combinatorial argument (e.g. see [1] or [5]) which gives

\[
\psi(x, p_t) \geq \left( t + \left\lfloor \log x / \log t \right\rfloor \right),
\]

(4.9)

and then evaluating the binomial coefficient by (4.3) we arrive at the lower bound implied by (4.8). The upper bound could be also obtained from known results on \( \psi(x, y) \), but it seems more appropriate to proceed directly. Note that \( \psi(x, p_t) \) represents the number of lattice points \( (a_1, \ldots, a_t) \in (N \cup \{0\})^t \) such that

\[
a_1 \log p_1 + \ldots + a_t \log p_t \leq \log x.
\]

Each such lattice point lies in a "lower left corner" of a unit hypercube. If \( (w_1, \ldots, w_t) \in (Re^+)^t \) is in one of these hypercubes, then

\[
\sum_{i \leq t} w_i \log p_i \leq \log x + \sum_{i \leq t} \log p_i \leq \log x + (1 + \varepsilon) p_t,
\]

by the prime number theorem. Thus \( \psi(x, p_t) \) does not exceed the \( t \)-dimensional volume of

\[
\{(w_1, \ldots, w_t) \in (Re^+)^t : \sum_{i \leq t} w_i \log p_i \leq \log x + (1 + \varepsilon) p_t,\}
\]
and consequently

\[
\psi(x, p) \leq \prod_{i \leq t} \frac{\log x + (1 + \epsilon) p_i}{t! \log p_i} = \\
\exp \left( -t \log t - o(t) + \log \log \log t + \right.

\[
\left. + (1 + \epsilon) p_i - \sum_{i < t} \log \log p_i \right)
\]

\[
= \exp \left( t \log \log x - t \log t \log \log \log x - t \log t + \\
+ t + o(t) + O(\log t / \log x) \right).
\]

By hypothesis \( p_t \leq \log^{1-\epsilon} x \), hence \( O(tp_t / \log x) = o(t) \) and (4.8) follows.

Having (4.8) at our disposal we may obtain the upper bound for \( \sum (x) \) as follows. Let \( \omega(n) = t, P(n) = p \) for \( n \) counted by \( \sum (x) \). Then obviously \( p \geq p_t \), and moreover for a fixed \( p \) there are \( \binom{\pi(p) - 1}{t - 1} \) choices for the remaining \( t - 1 \) prime factors of \( n \). Once the \( t \) prime factors of \( n \) are known, there are at most \( \psi(x, p_t) \) numbers with those prime factors counted by \( \sum (x) \), giving

\[
\sum (x) \leq \sum_{t \leq \log \log x} \psi(x, p_t) \sum_{p_t \leq p \leq x} p^{-t} \left( \frac{\pi(p) - 1}{t - 1} \right), \tag{4.10}
\]

since \( t = \omega(n) \leq 2 \log n / \log x, n \leq 2 \log x / \log x \). To estimate the inner sum in (4.10) we use the prime number theorem to obtain

\[
\sum_{p_t \leq p \leq x} p^{-t} \left( \frac{\pi(p) - 1}{t - 1} \right) \leq \sum_{p_t \leq p \leq x} p^{-t} \left( \frac{1 + \epsilon}{p / \log p} \right)^{-1} / (t - 1)! \tag{4.11}
\]

\[
\leq \frac{(1 + \epsilon)^{-1}}{(t - 1)!} \sum_{p_t \leq n \leq x} n^{-1} \log^{1-t} n \ll (1 + \epsilon)^{t!} (t! \log^{-2} t)^{-1}.
\]

To be able to use (4.8) we restrict \( t \) in (4.10) to the range \( t_0 \leq t \leq \log^{1-\epsilon} x \). The contribution of \( t^i \)’s for which \( t < t_0 \) is seen to be negligible by using the trivial bound

\[
\psi(x, y) < \left( \frac{\log x}{\log 2 + 1} \right)^{\pi(y)},
\]

while for the range \( t > \log^{1-\epsilon} x \) we may use the estimates of N. G. de Bruijn [1] or the elementary estimate

\[
\psi(x, y) < \left( \frac{\pi(y) + u}{u} \right)^{1 + \epsilon}, \quad u = \left[ \log x / \log y \right]
\]
of [5] and the trivial inequality $\binom{n}{k} \leq n^k / k!$. For the range $t_0 \leq t \leq \log^{1-\varepsilon} x$ in (4.10) we use (4.8) to obtain a contribution which is

$$
\ll \sum_{t_0 < t < \log^{1-\varepsilon} x} \psi(x, p_t) \sum_{p_t < p \leq x} p^{-t} \left( \frac{\pi(p) - 1}{t-1} \right)
$$

$$
\ll \log x \max_{t_0 < t < \log^{1-\varepsilon} x} \exp \left( t \log_2 x - t \log t - 2 t \log_2 t + t + o(t) - \log t! - \log x \max \exp (g(t)) \right),
$$

(4.12)

where

$$
g(t) = t \log \log x - 2 t \log t - 2 t \log \log t + (2 + \varepsilon) t.
$$

(4.13)

The function $g(t)$ differs from $f(t)$ (as defined by (4.6)) only that it has $\varepsilon$ in place of $-\varepsilon$, and its maximal value is determined analogously by solving the equation $g'(t) = 0$. This gives the value

$$
t = t_1(x) = (2 + o(1)) \log^{1/2} x (\log \log x)^{-1}, \quad (x \to \infty)
$$

thus completing the proof of (1.13), since with the above value (4.12) gives

$$
\sum (x) \ll \exp \left( (4 + \varepsilon) \log^{1/2} x (\log \log x)^{-1} \right).
$$

The proof of (1.14) is considerably simpler than the proof of (1.13). It is sufficient to prove

$$
\sum' 1/(P(n))^{\omega(n)} \ll 1/\log x,
$$

where $\sum'$ denotes summation over composite $n$, since

$$
\sum 1/p^{\omega(p)} = \sum_{p < x} 1/p = \log \log x + B + O(1/\log x). \quad (B = 0.26419 \ldots)(4.14)
$$

Write

$$
S = \sum' 1/(P(n))^{\omega(n)} = S_1 + S_2,
$$

say, where in $S_1$ we have $P(n) \leq y$ and in $S_2$, $P(n) > y$, and $y = y(x)$ will be suitably determined in a moment. Using the trivial

$$(P(n))^{\omega(n)} \geq n$$

and partial summation, we obtain

$$
S_1 \leq \int_{x}^{x} t^{-1} d\psi(t, y) \ll \log (x)^{-1} \psi(x, y) + \int_{x}^{x} \psi(t, y) t^{-2} dt.
$$

(4.15)

From [1] one has, for $y \leq x$ and some absolute $C > 0$,

$$
\psi(x, y) < x \exp \left( -C \log x / \log y \right).
$$

(4.16)
so that (4.15) gives

\[ S_1 \ll \exp(-C \log x / \log y) = 1/ \log x, \]

if we choose

\[ y = y(x) = \exp(C \log x / \log \log x). \]

To estimate \( S_2 \) note that the number of \( n \) with \( \Omega(n) = k \) and \( P(n) = p \) is at most \( \pi^{k-1}(p) \). If \( \Omega(n) = 2 \) and \( n > x \), then \( P(n) > x^{1/2} \).

From these elementary observations, we have

\[
S_2 \leq \sum_{n > x, \Omega(n) = 2} (P(n))^{-\Omega(n)} + \sum_{n > x, \Omega(n) > 2, P(n) > y} (P(n))^{-\Omega(n)}
\]

\[
\leq \sum_{p > x^{1/2}} p^{-2} \pi(p) + \sum_{k = 3}^{\infty} \sum_{p > y} p^{-k} \pi^{k-1}(p)
\]

\[
= \sum_{p > x^{1/2}} p^{-2} \pi(p) + \sum_{p > y} \pi^2(p) / (p^3 - p^2 \pi(p))
\]

\[ \ll \sum_{p > x^{1/2}} 1 / (p \log p) + \sum_{p > y} 1 / (p \log^2 p)
\]

\[ \ll 1 / \log x + 1 / \log^2 y \ll 1 / \log x, \]

using elementary estimates on the distribution of primes. Therefore

\[ S = S_1 + S_2 \ll 1 / \log x, \]

which proves (1.14). We finally remark that the error term \( O(1 / \log x) \) is best possible since

\[ \sum_{n > x} (P(n))^{-\Omega(n)} \gg \sum_{p > x} p^{-2} \pi(p) \ll 1 / \log x. \]

With more work we could replace the term \( O(1 / \log x) \) in (1.14) with \( (C + o(1)) / \log x \) for some positive constant \( C \), but we do not undertake this here.

§5. Proof of Theorem 4

We write the summatory function of \( 1 / Q(n) \) as

\[
\sum_{n < x} 1 / Q(n) = \sum_{n < x, Q(n) = P(n)} 1 / Q(n) + \sum_{n < x, Q(n) = P^2(n), k > 1} 1 / Q(n) + \sum_{n < x, Q(n) = P^k(n), k > 1} 1 / Q(n) = S_1 + S_2 + S_3,
\]

say. By (1.4) with \( r = 1 \) we have first

\[ S_2 \ll x \exp(-(2 + o(1)) (\log x \log_2 x)^{1/2}) \ll L^{-1/2} \sum_{n < x} 1 / P(n). \quad (5.1) \]
Next by writing
\[ \sum_{n \leq x, Q(n) = p} 1/Q(n) = \sum_{n \leq x} 1/P(n) - \sum_{n \leq x, Q(n) > P(n)} 1/P(n), \]
we see that (1.16) follows from (5.1) and
\[ \sum_{n \leq x, Q(n) = p} 1/P(n) \ll L^{-C} \sum_{n \leq x} 1/P(n), \tag{5.2} \]
where \( C > 0 \) is some absolute constant. As in the proof of (1.12) we may consider only those \( n \) for which (3.7) holds. Since \( Q(n) \) is the largest prime power dividing \( n \), we may assume that \( q^a = Q(n) > p = P(n) \) for some \( a \geq 2 \). Observe that in estimating the left-hand side of (5.2) we may also assume that \( q^a \ll L^4 \), since
\[ \sum_{q^a > L^4, \alpha > 2} q^{-\alpha} \ll L^{-2}. \]

If \( 0 < C_1 < C_2 \) are any two fixed constants, then from (1.8) and (2.1) or from earlier estimates on \( \psi(x, y) \) (see [1], [3]) we have, for \( L^{C_1} \ll y \ll L^{C_2} \) and \( u = \log x / \log y \),
\[ \psi(x, y) = x \exp \left( - (1 + o(1)) u \log u \right), \quad (x \to \infty). \tag{5.3} \]

In view of (1.2) and the remarks above it follows that we are left with \( O(\log x) \) sums of the form
\[ \sum_{a > 2} \sum' \sum_{p < q^a \ll L^4} p^{-1} \psi(xp^{-1}q^{-a}, p), \]
where \( a \geq 2 \) is a fixed integer and \( \sum' \) denotes summation over primes \( p \) which satisfy (3.7). Using (5.3) to estimate \( \psi(xp^{-1}q^{-a}, p) \), we obtain
\[ \sum_{a} \ll \sum' \sum_{p} x \exp \left\{ - (2^{-1/2} - 2\epsilon)(\log x \log_2 x)^{1/2} \right\} p^{-2} \sum_{q^a > p} q^{-a} \ll xL^{-2^{1/2} - 2\epsilon} \sum' \sum_{p} p^{-5/2} < xL^{-5(2^{1/2} - 5\epsilon}). \]

Since \( 5 \cdot 2^{-3/2} = 2^{1/2} + \frac{1}{4}2^{1/2} \), we obtain in view of (1.1)
\[ \sum_{a > 2} \sum_{a} \ll L^{-C(\epsilon)} \sum_{n \leq x} 1/P(n) \]
for some \( C(\epsilon) > 0 \), if \( \epsilon \) in (3.7) is sufficiently small. This proves (5.2) and completes the proof of (1.16). Presumably by methods similar to those used to prove Theorem 1 we could improve (1.16) and obtain an asymptotic formula for the sum
\[ \sum_{n \leq x} (1/P(n) - 1/Q(n)). \]
Finally we remark that by methods similar to those used in the proof of Theorem 4 we may obtain

\[ \sum_{2 \leq n \leq x} 1/f(n) = \{1 + \exp(-C(\log x \log \log x)^{1/2})\} \sum_{2 \leq n \leq x} 1/P(n). \] (5.4)

Here \( C > 0 \) is an absolute constant and \( f(n) \) denotes either \( \beta(n) = \sum_{p | n} p \) or \( B(n) = \sum_{p^r | n} \alpha p \) (see [6], Ch. 6 and [7] for some results concerning these functions). Thus (5.4) and Theorem 1 provide an asymptotic formula for sums of reciprocals of \( \beta(n) \) and \( B(n) \).

REFERENCES:

[2] N. G. de Bruijn, On the number of integers \( \leq x \) whose prime factors divide \( n \), Illinois J. Math. 6 (1962), 137–141.
[5] P Erdős and J. H. van Lint, On the number of positive integers \( \leq x \) and free of prime factors \( \geq y \), Simon Stevin 40 (1966), 73–76.

(Received September 14, 1984)