PACIFIC JOURNAL OF MATHEMATICS
Vol. 82, No. 2, 1979

ON THE ASYMPTOTIC BEHAVIOR OF LARGE PRIME FACTORS OF INTEGERS

K. ALLADI AND P. ERDÖS

We prove results on the asymptotic behavior of large prime factors of the integers. The basic idea of the paper is that if k is any fixed integer, then the kth largest prime factor of n, denoted by $P_k(n)$ is generally much bigger than $\sum_{j>k} P_j(n)$. We give precise estimates of this phenomenon. This paper is a sequel to an earlier paper by the authors.

1. Notations and definitions. Throughout this paper the letters p and q, with or without subscript will denote primes.

Let $n = \prod_{i=1}^{r} p_i^{\alpha_i}$, $p_1 > p_2 > \cdots > p_r$ be the canonical decomposition of an integer $n > 1$ into primes. We set

$$A(n) = \sum_{i=1}^{r} \alpha_i p_i \, , \qquad A^*(n) = \sum_{i=1}^{r} p_i \tag{1.1}$$

and

$$\Omega(n) = \sum_{i=1}^{r} \alpha_i \, , \qquad \omega(n) = r \, . \tag{1.2}$$

Let $A(1) = A^*(1) = \Omega(1) = \omega(1) = 0$.

We may define the kth largest prime factor in two ways depending on whether we want to count prime factors according to multiplicity or not. To be more precise set

$$\begin{aligned} P_k^*(n) &= p_k \quad \text{for} \quad k \leqq \omega(n) \\ &= 0 \quad \text{for} \quad k > \omega(n) \, . \end{aligned} \tag{1.3}$$

We may also define

$$\begin{aligned} &P_1(n) = p_1 \\ &P_k(n) = P_1\Big(\frac{n}{P_1(n) \cdot P_2(n) \cdots P_{k-1}(n)}\Big) \, , \qquad 1 < k \leqq \Omega(n) \\ &P_k(n) = 0 \quad \text{for} \quad k > \Omega(n) \, . \end{aligned} \tag{1.4}$$

Observe that $P_1(n) = P_1^*(n)$.

The terms "average order" and "normal order" will mean the following: Let f be an arithmetic function and set

$$F(x) = \sum_{1 \leqq n \leqq x} f(n) \, . \tag{1.5}$$

Suppose g is a monotonic function such that

$$\lim_{x\to\infty} \frac{G(x)}{F(x)} = 1 \tag{1.6}$$

where

$$G(x) = \sum_{1\leq n\leq x} g(n) , \tag{1.7}$$

then f has average order g. Next, we say that two functions f and g are "nearly the same almost always" if for each $\varepsilon > 0$

$$\lim_{x\to\infty} \frac{\psi_\varepsilon(x)}{x} = 1 , \tag{1.8}$$

where

$$\psi_\varepsilon(x) = \sum_{\substack{1\leq n\leq x \\ 1-\varepsilon<(f(n)/g(n))<1+\varepsilon}} 1 . \tag{1.9}$$

If in (1.9) and (1.8), the function g is monotonic, we say that f has normal order g.

Consider the sum

$$\psi(x, y) = \sum_{\substack{1\leq n\leq x \\ P_1(n)\leq y}} 1 . \tag{1.10}$$

If $\alpha \geqq 1$ is a real number and $y = x^{1/\alpha}$, it is well known (see [10]) that

$$\rho(\alpha) = \lim_{x\to\infty} \frac{\psi(x, x^{1/\alpha})}{x} \tag{1.11}$$

exists. The limit in (1.11) is also defined if $-\infty < \alpha < 1$ and

$$\rho(\alpha) = \begin{cases} 1 & 0 \leqq \alpha < 1 \\ 0 & -\infty < \alpha < 0 . \end{cases} \tag{1.12}$$

The function $\rho(\alpha)$ is a monotonic decreasing continuous function of α for $\alpha \geqq 1$.

Finally we define the sums

$$\begin{aligned} S_1(x, k) &= \sum_{2\leq n\leq x} \frac{A(n) - P_1(n) - \cdots - P_{k-1}(n)}{P_1(n)} , & k \geqq 1 \\ S_2(x, k) &= \sum_{2\leq n\leq x} \frac{A^*(n) - P_1^*(n) - \cdots - P_{k-1}^*(n)}{P_1(n)} , & k \geqq 1 \\ S_3(x, k) &= \sum_{2\leq n\leq x} \frac{P_k(n)}{P_1(n)} , & k \geqq 1 \\ S_4(x, k) &= \sum_{2\leq n\leq x} \frac{P_k^*(n)}{P_1(n)} , & k \geqq 1 . \end{aligned} \tag{1.13}$$

The aim of this paper is to obtain estimates for these sums $S_i(x, k)$, $i = 1, 2, 3, 4$.

2. General background and main theorem. The results in this paper are in continuation of those in § 2 of [2].

It is a well known theorem of Hardy and Ramanujan [6], [7] that the functions $\Omega(n)$ and $\omega(n)$ both have average and normal order $\log\log n = g(n)$. This means that a number n usually has $\log\log n$ prime factors and most of them occur square free. Thus it is natural to expect the large prime factors to occur with multiplicity one, most of the time. So one should be able to show that the functions A and A^* have the same average order. In an earlier paper [2] we showed this to be true and much more.

Not only do A and A^* have the same average order, but the function $P_1(n)$ dominates the sums in (1.1) to such an extent that A, A^* and P_1 have the same average order. More generally $A(n) - P_1(n) - \cdots - P_{k-1}(n)$ and $P_k(n)$ have the same average order. It was observed in [1] that the functions $P_k^*(n)$ and $A^*(n) - P_1^*(n) - \cdots - P_{k-1}^*(n)$ also have the same average order as $P_k(n)$, since the asymptotic analysis in [2] remains unaffected if the weak inequalities are replaced by strict ones. Thus we restate (without proof) the main theorem in [2] in a more complete form:

THEOREM A. *If k is a fixed positive integer then*

$$\begin{aligned}(2.1)\qquad &\sum_{1\leq n\leq x}\{A(n) - P_1(n) - \cdots - P_{k-1}(n)\} \sim \sum_{1\leq n\leq x} P_k(n) \sim \sum_{1\leq n\leq x} P_k^*(n)\\ &\sim \sum_{1\leq n\leq x}\{A^*(n) - P_1^*(n) - \cdots - P_{k-1}^*(n)\} \sim a_k \frac{x^{1+1/k}}{(\log x)^k}\end{aligned}$$

where a_k is a constant depending only on k, and is a rational multiple of $\zeta(1 + 1/k)$ where ζ is the Riemann zeta function. In addition for each $k \geqq 1$

$$(2.2)\qquad \sum_{1\leq n\leq x}\{A(n) - A^*(n)\} = x\log\log x + O(x) = o\Big(\sum_{1\leq n\leq x} P_k^*(n)\Big)\,.$$

Theorem A says that the average order in (2.1) is $g(n) = a_k^* \cdot n^{1/k}/(\log n)^k$ where $a_k^* = a_k \cdot (1 + 1/k)$. An average is essentially influenced by two things—(i) the abnormally large values of a function, which certainly contribute to (2.1) and (ii) the values a function takes most often.

The question now arises whether A, A^*, and P_1 are nearly the same almost always. The main theorem stated below answers this question in the affirmative.

THEOREM B. *If k is a fixed positive integer then*

$$S_1(x, k) \sim S_2(x, k) \sim S_3(x, k) \sim S_4(x, k) \sim a'_k \frac{x}{(\log x)^{k-1}} \tag{2.3}$$

where $a'_1 = 1$ and a'_k for $k > 1$ is a constant depending only on k, and is a rational multiple of e^γ where γ is Euler's constant. In addition for each $k \geqq 1$

$$\sum_{2 \leqq n \leqq x} \frac{A(n) - A^*(n)}{P_1(n)} = O\left(\frac{x}{e^c\sqrt{\log x \log\log x}}\right) = o\left(\sum_{2 \leqq n \leqq x} \frac{P_k^*(n)}{P_1(n)}\right) \tag{2.4}$$

where c is an absolute constant >0.

3. Consequences and motivation. Statements (2.3) and (2.4) may be looked upon as analogues to (2.1) and (2.2). Theorem A said that A, A^* and P_1 have the same average order, $\pi^2 n/6 \log n$, ($a_1 = \pi^2/12$, see [2]). We can deduce from Theorem B the following.

COROLLARY. *The functions A, A^* and P_1 are all nearly the same almost always. Also all three functions fail to possess a normal order.*

Proof. Consider two arithmetic functions f, g satisfying $f(n) \geqq g(n) > 0$. Suppose that

$$\sum_{1 \leqq n \leqq x} \frac{f(n)}{g(n)} \sim x\ . \tag{3.1}$$

We rewrite (3.1) as

$$\sum_{1 \leqq n \leqq x} \left\{\frac{f(n)}{g(n)} - 1\right\} = o(x)\ . \tag{3.2}$$

Since $f/g \geqq 1$ we infer from (3.2) that

$$\left|\frac{\psi_\varepsilon(x)}{x} - 1\right| < \frac{o(x)}{\varepsilon \cdot x} \longrightarrow 0 \quad \text{as} \quad x \longrightarrow \infty \tag{3.3}$$

for each $\varepsilon > 0$, where $\psi_\varepsilon(x)$ is as in (1.8). So f and g are nearly the same almost always. (We can deduce (3.3) also if $f(n) \leqq g(n)$ for all n).

Setting $k = 1$ in (2.3) we see that (3.1) is true with $f = A(n)$ and $g(n) = P_1(n)$. Therefore A and P_1 are nearly the same almost always. Since $A \geqq A^* \geqq P_1$, the same is true for all three functions.

Now to show that these three functions do not have normal orders it suffices to show that one of them does not. It follows easily from a theorem of Elliott [5] on additive functions

$$f(n) = \sum_{p|n} f(p) \,, \tag{3.4}$$

that A^* does not have a normal order. That proves the corollary.

REMARK. Since $A(n) \geqq \log n$, it follows from (2.2) that

$$\begin{aligned} \left| \sum_{2 \leqq n \leqq x} \left\{ \frac{A^*(n)}{A(n)} - 1 \right\} \right| &\leqq \sum_{2 \leqq n \leqq x} \frac{A(n) - A^*(n)}{\log n} \\ &= 0\left(\frac{x \log \log x}{\log x}\right) . \end{aligned} \tag{3.5}$$

From (3.2), (3.3) and (3.5) we can deduce that A and A^* are nearly the same almost always.

Let us look a little more closely at (2.3) which for $f = A$ or A^* and $g = P_1$ is a more accurate form of (3.1). We may rewrite (2.3) as

$$\begin{aligned} \sum_{2 \leqq n \leqq x} \frac{A^*(n)}{P_1(n)} = \sum_{2 \leqq n \leqq x} 1 &+ \sum_{2 \leqq n \leqq x} \frac{P_2^*(n)}{P_1(n)} + \cdots \\ &+ \sum_{1 \leqq n \leqq x} \frac{P_k^*(n)}{P_1(n)} + \cdots \end{aligned} \tag{3.6}$$

where

$$\sum_{2 \leqq n \leqq x} \frac{P_k^*(n)}{P_1(n)} \sim a_k' x/(\log x)^k \,. \tag{3.7}$$

We show in § 5 that

$$a_k' = \int_1^\infty \rho(s - k) s^{k-2} ds \tag{3.8}$$

where ρ is defined in (1.11). We deduce from (3.8) in § 6 that a_k' is a rational multiple of e^γ for $k > 1$. The integral representation is investigated in § 6 and this leads to pretty connections with some related problems.

The next section is devoted to obtaining upper and lower bounds for $S_i(x, k)$, $i = 1, 2, 3, 4$. This enables us to deduce the first four asymptotic relations in (2.3). It is only § 5 that we prove (3.7) and (3.8). But the upper bound method in § 4 is used in § 5 to take care of the error terms arising out of (3.6) and (2.3). For the reader who does not want to go through the detailed proof, see [1], where some of the ideas of this paper and an earlier paper by the authors [2] are summarized.

We now move on to the proofs of our results.

4. Upper and lower bounds. In what follows, $c_1, c_2, c_3, \cdots$

denote absolute positive constants whose precise values will not be our concern. Also $\exp\{x\} = e^x$. We begin by proving

THEOREM 1. *There exists for each positive integer k a constant b_k and a real number $x_0 = x_0(k)$ such that if $x \geqq x_0$ then $S_i(x, k) > b_k \cdot x/(\log x)^{k-1}$ for $i = 1, 2, 3, 4$.*

To prove this we need

LEMMA 1. *Let s be a positive real number. Then*

$$\sum_{p>x} \frac{1}{p(\log p)^s} = \frac{1}{s(\log x)^s} + O(\exp\{-c_1\sqrt{\log x}\}) .$$

Proof. We use the Prime Number Theorem [4], [9] in the form

$$|\pi(x) - li(x)| = O(x \exp\{-c_2\sqrt{\log x}\}) . \tag{4.1}$$

Now write

$$\begin{aligned}\sum_{p>x} \frac{1}{p(\log p)^s} &= \int_{x+}^{\infty} \frac{d\pi(y)}{y(\log y)^s} = \int_{x}^{\infty} \frac{dy}{y(\log y)^{s+1}} + \int_{x+}^{\infty} \frac{d\{\pi(y) - li(y)\}}{y(\log y)^s} \\ &= \frac{1}{s(\log x)^s} + \frac{\pi(y) - li(y)}{y(\log y)^s}\Bigg|_{x+}^{\infty} \\ &\quad + O\left(\int_{x+}^{\infty} \frac{\{\pi(y) - li(y)\}}{y^2(\log y)^{s+1}}\,dy\right) .\end{aligned} \tag{4.2}$$

Lemma 1 follows from (4.1) and (4.2).

Proof of Theorem 1. It suffices to prove Theorem 1 for the smallest of the four sums $S_4(x, k)$.

Assume first that $k > 1$. For x sufficiently large choose a prime p_1 in the interval

$$k!\,x^{1/k+1} \leqq p_1 \leqq x^{1/k} . \tag{4.3}$$

Now choose primes $p_2, p_3, \cdots, p_k$ satisfying

$$\frac{p_1}{k} < p_k < \frac{p_1}{k-1} < p_{k-1} < \frac{p_1}{k-2} < \cdots < \frac{p_1}{3} < p_3 < \frac{p_1}{2} < p_2 < p_1 . \tag{4.4}$$

Consider any multiple $m \leqq x$ of $p_1 p_2 \cdots p_k$

$$m = n' p_1 p_2 \cdots p_k . \tag{4.5}$$

Because of (4.3) and (4.4) we have

$$p_1 p_2 \cdots p_k \leqq x \tag{4.6}$$

and

$$n' \leqq \frac{m}{p_1 \cdots p_k} \leqq \frac{x}{(k!)^{k-1} x^{k/k+1}} < x^{1/k+1} . \tag{4.7}$$

By (4.7) and (4.4)

$$P_1(n') \leqq n' < x^{1/k+1} \leqq (k-1)!\, x^{1/k+1} < p_k . \tag{4.8}$$

Thus by (4.5) and (4.8) we see that $P_k^*(m) = p_k$. So any multiple $\leqq x$ of $p_1 \cdots p_k$ has p_k as its kth largest prime factor (P_k^*). So

$$S_4(x, k) = \sum_{2 \leqq n \leqq x} \frac{P_k^*(n)}{P_1(n)} \geqq \sum_{\substack{2 \leqq n \leqq x \\ n = n' p_1 \cdots p_k \\ (p_i \text{ satisfying } (4.4))}} \frac{P_k^*(n)}{P_1(n)} . \tag{4.9}$$

We can estimate the second sum in (4.9) by using the well known result [11]

$$\sum_{p \leqq k} \frac{1}{p} = \log\log x + c_3 + O(\exp\{-c_4\sqrt{\log x}\}) . \tag{4.10}$$

Observe that the second sum in (4.9) is

$$\begin{aligned}
&\sum_{k! x^{1/(k+1)} \leqq p_1 \leqq x^{1/k}} \; \sum_{p_1/2 < p_2 < p_1} \; \sum_{p_1/3 < p_3 < p_1/2} \cdots \sum_{p_1/k < p_k < p_1/(k-1)} \frac{x}{p_1 \cdots p_k} \cdot \frac{p_k}{p_1} \\
&= x \sum_{k! x^{1/(k+1)} \leqq p_1 \leqq x^{1/k}} \frac{1}{p_1^2} \left\{ \left(\prod_{j=2}^{k-1} \left(\sum_{p_1/j < p_j < p_1/(j-1)} \frac{1}{p_j} \right) \right) \cdot \left(\sum_{p_1/k < p_k < p_1/(k-1)} 1 \right) \right\} \\
&\geqq b'_k x \sum_{k! x^{1/(k+1)} \leqq p_1 \leqq x^{1/k}} \cdot \frac{1}{p_1 (\log p_1)^{k-1}} > b_k x / (\log x)^{k-1}
\end{aligned} \tag{4.11}$$

by virtue of Lemma 1 and (4.10). Theorem 1 follows from (4.9) and (4.11), for $k > 1$. For $k = 1$, Theorem 1 is trivially true.

Now for an upper bound.

THEOREM 2. *All four sums $S_i(x, k)$, $i = 1, 2, 3, 4$ are $O(x/(\log x)^{k-1})$ where k is an integer $\geqq 1$, and the O-constant depends only on k.*

We need a few preliminary results before proving Theorem 2.

LEMMA 2. *Let k be a nonnegative integer and*

$$S_k^*(x) = \sum_{1 \leqq p \leqq x} (\log\log x - \log\log p)^k .$$

Then

$$S_k^*(x) = \frac{k!\, x}{(\log x)^{k+1}} + O_k\left(\frac{x \cdot \log\log x}{(\log x)^{k+2}} \right) .$$

Proof. If we write $S_k^*(x)$ as a Stieltjes integral, use the fact that

$$d\pi(y) = \frac{dy}{\log y} + d\{\pi(y) - li(y)\} ,$$

integrate the second integral by parts and then use (4.1) we get

$$S_k^*(x) = \int_4^x \frac{(\log\log x - \log\log y)^k}{\log y} dy + O\Big(\frac{x}{(\log x)^{k+2}}\Big) . \tag{4.12}$$

Next

$$\begin{aligned} T_k(x) &= \int_4^x \frac{(\log\log x - \log\log y)^k}{\log y} dy \\ &= (\log\log x - \log\log y)^k li(y)\Big|_4^x \\ &\qquad + k\int_4^x \frac{li(y)(\log\log x - \log\log y)^{k-1}}{y\log y} dy \\ &= O((\log\log x)^k) + k\int_4^x \frac{li(y)(\log\log x - \log\log y)^{k-1}}{y\log y} dy . \end{aligned} \tag{4.13}$$

But

$$li(y) = \frac{y}{\log y} + O\Big(\frac{y}{\log^2 y}\Big) . \tag{4.14}$$

So the integral in (4.13) becomes

$$\begin{aligned} &k\int_4^x \frac{(\log\log x - \log\log y)^{k-1}}{\log^2 y} + O\Big(\int_4^x \frac{(\log\log x - \log\log y)^{k-1}}{\log^3 y} dy\Big) \\ &= I_1 + I_2 . \end{aligned} \tag{4.15}$$

We split I_1 into

$$I_1 = k\int_4^{x/(\log x)^{k+3}} + k\int_{x/(\log x)^{k+3}}^x . \tag{4.16}$$

Clearly in (4.16)

$$\int_4^{x/(\log x)^{k+3}} = O\Big(\frac{x(\log\log x)^{k-1}}{(\log x)^{k+3}}\Big) = O\Big(\frac{x}{(\log x)^{k+2}}\Big) . \tag{4.17}$$

Regarding the second integral in (4.16) we observe that

$$\begin{aligned} k\int_{x/(\log x)^{k+3}}^x &= k\left\{\frac{1}{\log x} + O\Big(\frac{\log\log x}{\log^2 x}\Big)\right\} \\ &\qquad \times \int_{x/(\log x)^{k+3}}^x \frac{(\log\log x - \log\log y)^{k-1}}{\log y} dy . \end{aligned} \tag{4.18}$$

Now the last integral in (4.18) is

$$T_{k-1}(x) + O\Big(\frac{x(\log\log)^{k-1}}{(\log x)^{k+3}}\Big) = T_{k-1}(x) + O\Big(\frac{x}{(\log x)^{k+2}}\Big) . \tag{4.19}$$

From the definition of T_k we have

$$T_0(x) = \frac{x}{\log x} + O\Big(\frac{x}{\log^2 x}\Big) . \tag{4.20}$$

Now make the induction hypothesis that for $k \geqq 1$

$$T_{k-1}(x) = \frac{(k-1)!\,x}{(\log x)^k} + O\Big(\frac{x(\log\log x)}{(\log x)^{k+1}}\Big) . \tag{4.21}$$

Then from equations (4.16) through (4.21) we deduce that

$$I_1 = \frac{k!\,x}{(\log x)^{k+1}} + O\Big(\frac{x\log\log x}{(\log x)^{k+2}}\Big) . \tag{4.22}$$

By analysis very similar to the above one can show that

$$I_2 = O\Big(\frac{x}{(\log x)^{k+2}}\Big) . \tag{4.23}$$

So from (4.22), (4.23), (4.15) and (4.13) we see that (4.21) is true for $T_k(x)$ and so by induction for all $k \geqq 1$. Lemma 2 follows from (4.12) and (4.21).

LEMMA 3. *Let $x,\ y \geqq 4$ be real numbers and $k \geqq 0$ an integer. Then*

$$\sum_{y \leq p \leq x} \frac{(\log\log x - \log\log p)^k}{p} = \frac{(\log\log x - \log\log y)^{k+1}}{k+1} + O_k((\log\log x - \log\log y)^k \exp\{-c_5\sqrt{\log y}\}) .$$

Proof. As in the beginning of the proof of Lemma 2 we convert the above sum into a Stieltjes integral and replace $d\pi(y)$ by $dy/\log y$. Lemma 3 can be easily proved by making the substitution $\log\log x - \log\log y = t$. We do not go through the details.

Proof of Theorem 2. It suffices to prove Theorem 2 for the largest of the four sums $S_1(x, k)$. That is we will show

$$S_1(x, k) = \sum_{2 \leqq n \leqq x} \frac{A(n) - P_1(n) - \cdots - P_{k-1}(n)}{P_1(n)} = O\Big(\frac{x}{(\log x)^{k-1}}\Big) \tag{4.24}$$

for $k \geqq 1$ an integer. We claim that it suffices to prove (4.24), for $k > 1$ because for $k = 1$ we have

$$S_1(x, 1) = \sum_{2 \leq n \leq x} \frac{A(n)}{P_1(n)} = \sum_{2 \leq n \leq x} 1 + \sum_{2 \leq n \leq x} \frac{A(n) - P_1(n)}{P_1(n)} \tag{4.25}$$
$$= x - 1 + S_1(x, 2)$$
$$= O(x)$$

assuming that (4.24) is true for $k = 2$. So from now on we assume that $k > 1$.

We write

$$\frac{A(n) - P_1(n) - P_2(n) - \cdots - P_{k-1}(n)}{P_1(n)} = \frac{P_k(n)}{P_1(n)} + \frac{P_{k+1}(n)}{P_1(n)} + \cdots . \tag{4.27}$$

Let us denote a general nonzero term of (4.27) by p_k/p_1. We would like to know how often this term occurs in $S_1(x, k)$. The term p_k/p_1 occurs as often as we can find integers $n = p_1p_2 \cdots p_{k-1} \cdot p_k m \leqq x$ where the p_i satisfy $p_k \leqq p_{k-1} \leqq \cdots \leqq p_1$ and $P_1(m) \leqq p_{k-1}$. If we fix the primes p_i to satisfy these conditions then the number of such n is given by

$$\psi\left(\frac{x}{p_1p_2 \cdots p_k}, p_{k-1}\right) \tag{4.28}$$

where ψ is defined in (1.10).

Thus we may rewrite (4.27) as

$$S_1(x, k) = \sum_{2 \leq p_1 \leq x} \sum_{p_k \leq p_1} \sum_{p_k \leq p_{k-1} \leq p_1} \sum_{p_{k-1} \leq p_{k-2} \leq p_1} \cdots \tag{4.29}$$
$$\cdots \sum_{p_3 \leq p_2 \leq p_1} \psi\left(\frac{x}{p_1 \cdots p_k}, p_{k-1}\right) \frac{p_k}{p_1} .$$

We first consider a subsum of (4.29) with a restriction on p_1. That is we choose β with $0 < \beta < 1$, whose value will be specified later, and consider p_1 in (4.29) satisfying $x^\beta \leqq p_1 \leqq x$. We shall get an upper bound for this sum.

Observe that the sum in (4.29) with this extra condition on p_1 is

$$\leqq \sum_{x^\beta \leq p_1 \leq x} \sum_{p_k \leq p_1} \sum_{p_k \leq p_{k-1} \leq p_1} \cdots \sum_{p_3 \leq p_2 \leq p_1} \frac{x}{p_1 \cdots p_k} \cdot \frac{p_k}{p_1} \tag{4.30}$$
$$= x \sum_{x^\beta \leq p_1 \leq x} \frac{1}{p_1^2} \sum_{p_k \leq p_1} 1 \sum_{p_k \leq p_{k-1} \leq p_1} \frac{1}{p_{k-1}} \cdots \sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2} .$$

(Note: If $k = 2$ in (4.30) we have only

$$x \sum_{x^\beta \leq p_1 \leq x} \frac{1}{p_1^2} \sum_{p_2 \leq p_1} 1 \tag{4.31}$$

and no other terms. For $k > 2$, there is no confusion in (4.30).) Because of this difference assume for the moment that $k > 2$. Then if we use Lemma 3 we infer

$$\sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2} = O(\log\log p_1 - \log\log p_3) . \tag{4.32}$$

Again by Lemma 3 and (4.32)

$$\sum_{p_4 \leq p_3 \leq p_1} \frac{1}{p_3} \sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2} = O((\log\log p_1 - \log\log p_4)^2) . \tag{4.33}$$

Iterating this process we get in (4.30) for $k > 2$

$$O\Big(x \sum_{x^\beta \leq p_1 \leq x} \frac{1}{p_1^2} \sum_{p_k \leq p_1} (\log\log x - \log\log p_k)^{k-2}\Big) \tag{4.34}$$

by repeated use of Lemma 3. Now observe that because of (4.31) we see that (4.34) is true even for $k = 2$. Thus for $k \geqq 2$, we may replace (4.30) by (4.34). Thus from now on we drop the assumption $k > 2$, but of course still assume $k > 1$.

To estimate (4.34) we use Lemma 2 which gives

$$O\Big(x \sum_{x^\beta \leq p_1 \leq x} \frac{1}{p_1(\log p_1)^{k-1}}\Big) . \tag{4.35}$$

Finally Lemma 1 and (4.35) imply that the sum in (4.35) and hence in (4.30) is

$$O\Big(\frac{x}{\beta^{k-1}(\log x)^{k-1}}\Big) \tag{4.36}$$

where the constant on the O-term in (4.36) depends only on k and not on β.

So (4.36) gives a bound for the sum in (4.29) with the condition $x^\beta \leqq p_1 \leqq x$. For the sum corresponding to $p_1 \leqq x^\beta$ we write

$$\sum_{2 \leq p_1 \leq x^\beta} = \sum_{m=0}^{\infty} \sum_{x^{\beta/2^{m+1}} \leq p_1 \leq x^{\beta/2^m}} . \tag{4.37}$$

To estimate (4.37) we use the following result of de Bruijn [3]; If $y = x^{1/\alpha}$ then

$$\psi(x, y) = O(x \exp\{-c_6\alpha\}) . \tag{4.38}$$

In (4.37) consider the case

$$x^{\beta/2^{m+1}} \leqq p_1 \leqq x^{\beta/2^m} . \tag{4.39}$$

Then in (4.29) with the restriction (4.39) on p_1 we have from (4.38) the following:

$$\alpha = \frac{\log (x/p_1 \cdots p_k)}{\log p_{k-1}} \geqq \left(1 - \frac{k\beta}{2^m}\right) \Big/ \left(\frac{\beta}{2^m}\right) = \frac{2^m - k\beta}{\beta} . \tag{4.40}$$

We choose $\beta = \beta(k)$, depending on k, so small that

$$\alpha \geqq \frac{2^m - k\beta}{\beta} > \frac{2^{m-1}}{\beta} . \tag{4.41}$$

Then by (4.38), (4.39), and (4.41) we will have in (4.29) for the subsum corresponding to (4.39)

$$\psi\left(\frac{x}{p_1 \cdots p_k}, p_{k-1}\right) = O\left(\frac{x}{p_1 \cdots p_k} \exp\{-c_6 2^{m-1}/\beta\}\right) . \tag{4.42}$$

If we substitute (4.42) in (4.29) and analyze this sum just the way we derived (4.36) we get

$$\begin{aligned} &O\left(\frac{x}{\{\log (x^{\beta/2^{m+1}})\}^{k-1}} \exp\{-c_6 2^{m-1}/\beta\}\right) \\ &\quad = O\left(\frac{x(2^{m+1}/\beta)^{k-1}}{(\log x)^{k-1} \exp\{c_6 2^{m-1}/\beta\}}\right) . \end{aligned} \tag{4.42}$$

But then

$$\sum_{m=0}^{\infty} \frac{(2^{m+1}/\beta)^{k-1}}{\exp\{c_6 2^{m-1}/\beta\}} < \infty . \tag{4.43}$$

This means that (4.43), (4.42), and (4.36) imply that in (4.29)

$$S_1(x, k) = O(x/(\log x)^{k-1})$$

for $k > 1$. That completes the proof of Theorem 2.

It is interesting to note that Theorems 1 and 2 actually imply the first four asymptotic relations in Theorem B, as will be shown below. Before establishing this we prove the last part of Theorem B namely

THEOREM 3. *For each positive integer k we have*

$$\begin{aligned} \sum_{2 \leqq n \leqq x} \frac{A(n) - A^*(n)}{P_1(n)} &= O(x \exp\{-c_7 \sqrt{\log x \log \log x}\}) \\ &= o\left(\sum_{2 \leqq n \leqq x} \frac{P_k^*(n)}{P_1(n)}\right) . \end{aligned}$$

Proof. First let $1 \leqq y \leqq x$ and $y = x^{1/\alpha}$. N. G. de Bruijn [3] showed that if $3 < \alpha < 4y^{1/2}/\log y$ then

$$\psi(x, y) = O(x \log^2 y \exp\{-\alpha \log \alpha - \alpha \log\log \alpha + c_8 \alpha\}) . \tag{4.44}$$

Take $y = \exp\{\sqrt{\log x \log\log x}\}$. Then from (4.44) we have

$$\psi(x, y) = O(x \exp\{-c_9 \sqrt{\log x \log\log x}\}) . \tag{4.45}$$

Next observe that

$$\frac{A(n) - A^*(n)}{P_1(n)} \leqq \Omega(n) = O(\log n) . \tag{4.46}$$

We now split

$$\sum_{2 \leqq n \leqq x} \frac{A(n) - A^*(n)}{P_1(n)} = \sum_{\substack{2 \leqq n \leqq x \\ P_1(n) \leqq y}} + \sum_{\substack{2 \leqq n \leqq x \\ P_1(n) > y}} = \sum_1 + \sum_2 . \tag{4.47}$$

Clearly from (4.46) and (4.45)

$$\sum_1 = O(\log x \cdot \psi(x, y)) = O(x \exp\{-c_{10} \sqrt{\log x \log\log x}\}) . \tag{4.48}$$

But then by Theorem A, (2.2), we have

$$\begin{aligned} \sum_2 &\leqq \exp\{-\sqrt{\log x \log\log x}\} \sum_{2 \leqq n \leqq x} (A(n) - A^*(n)) \\ &= O(x \exp\{-c_{11} \sqrt{\log x \log\log x}\}) . \end{aligned} \tag{4.49}$$

The first equation in Theorem 3 follows from (4.47), (4.48) and (4.49). The second equation is a consequence of Theorem 1. That proves Theorem 3.

THEOREM 4. *For every integer $k \geqq 1$ we have*

$$S_1(x, k) \sim S_2(x, k) \sim S_3(x, k) \sim S_4(x, k) .$$

Proof. The smallest of the four sums is $S_4(x, k)$. By Theorem 1

$$S_1(x, k) \geqq S_4(x, k) > b_k x/(\log x)^{k-1} . \tag{4.50}$$

The largest of the four sums is $S_1(x, k)$. Consider the difference

$$\begin{aligned} S_1(x, k) - S_4(x, k) &= \sum_{2 \leqq n \leqq x} \frac{A(n) - P_1(n) - \cdots - P_{k-1}(n) - P_k^*(n)}{P_1(n)} \\ &= \sum_{2 \leqq n \leqq x} \frac{A(n) - P_1(n) - \cdots - P_k(n)}{P_1(n)} \\ &\quad + \sum_{2 \leqq n \leqq x} \frac{P_k(n) - P_k^*(n)}{P_1(n)} \\ &= S_1(x, k+1) + \sum_{2 \leqq n \leqq x} \frac{P_k(n) - P_k^*(n)}{P_1(n)} . \end{aligned} \tag{4.51}$$

By Theorem 2

$$(4.52)\qquad S_1(x, k+1) = O(x/(\log x)^k) .$$

But then

$$(4.53)\quad A(n) - A^*(n) = \sum_{j=1}^{\infty} P_j(n) - \sum_{j=1}^{\infty} P_j^*(n) = \sum_{j=1}^{\infty} \{P_j(n) - P_j^*(n)\} \geqq P_k(n) - P_k^*(n) .$$

So by (4.53) and Theorem 3 we have

$$(4.54)\quad \sum_{2\leqq n\leqq x} \frac{P_k(n) - P_k^*(n)}{P_1(n)} \leqq \sum_{2\leqq n\leqq x} \frac{A(n) - A^*(n)}{P_1(n)} = O\left(x \exp\{-c_7\sqrt{\log x \log\log x}\}\right) .$$

Clearly from (4.51), (4.52) and (4.54)

$$(4.55)\qquad S_1(x, k) - S_4(x, k) = O(x/(\log x)^k) .$$

Thus from (4.55) and (4.50) we deduce

$$(4.56)\qquad S_1(x, k) \sim S_4(x, k) .$$

But since these are the smallest and largest sums, Theorem 4 follows from (4.56).

While proving Theorem 2 we did not use Lemmas 1, 2, and 3 in the forms in which they were stated, but used only the upper bounds they implied. These lemmas will play a role in obtaining asymptotic estimates, which we take up in the next section. We refer to the method of proof of Theorem 2 (namely the choice of β and the convergence of the series (4.43)), as the "upper bound method" and use this method to take care of the error terms arising out of the asymptotic estimates in what follows.

5. **Asymptotic estimates.** Our goal in this section is to prove

THEOREM 5. *Let k be a positive integer. Then all the four sums $S_i(x, k)$, $i = 1, 2, 3, 4$ are asymptotically equal to*

$$a_k' x/(\log x)^{k-1}$$

where

$$a_k' = \int_1^{\infty} \rho(s-k) s^{k-2} ds .$$

We need some lemmas before we go to the proof.

LEMMA 4. *If $\alpha \geqq 1$ and $\varepsilon > 0$ then*

$$|\rho(\alpha+\varepsilon)-\rho(\alpha)| = O\Big(\frac{\varepsilon}{\Gamma(\alpha+1)}\Big).$$

Proof. It is well known (see [10]) that ρ satisfies

$$\rho(\alpha) = 1 - \int_1^\alpha \frac{\rho(t-1)dt}{t}. \tag{5.1}$$

Furthermore (see [10], [3])

$$\rho(\alpha) < \frac{c_{12}}{\Gamma(\alpha+1)}. \tag{5.2}$$

Combining (5.1) and (5.2) we get

$$0 \leqq \rho(\alpha) - \rho(\alpha+\varepsilon) = \int_\alpha^{\alpha+\varepsilon} \frac{\rho(t-1)}{t}dt \leqq \frac{\rho(\alpha-1)\varepsilon}{\alpha} = O\Big(\frac{\varepsilon}{\Gamma(\alpha+1)}\Big)$$

because ρ is monotonic decreasing.

LEMMA 5. *There exists constants* c_{13}, c_{14} *and* c_{15} *such that if* $\alpha \geqq 1$ *and* $y = x^{1/\alpha}$, $x \geqq 1$, *then*

$$|\psi(x, x^{1/\alpha}) - x\rho(\alpha)| \leqq 2\max\Big(\frac{c_{13}x\alpha^2}{\exp\{c_{14}\sqrt{\log y}\}}, \frac{c_{15}x\alpha}{e^{\alpha/4}\cdot\log x}\Big).$$

Proof. Lemma 5 is obtained by combining certain results of de Bruijn [3]. For the function $\Lambda(x, x^{1/\alpha})$ defined by de Bruijn, it is known

$$|\psi(x, x^{1/\alpha}) - \Lambda(x, x^{1/\alpha})| < c_{13}x\alpha^2\exp\{-c_{14}\sqrt{\log y}\} \tag{5.3}$$

and

$$|\Lambda(x, x^{1/\alpha}) - x\rho(\alpha)| < c_{15}x\alpha/(e^{\alpha/4}\cdot\log x). \tag{5.4}$$

Lemma 5 follows from (5.3) and (5.4).

Proof of Theorem 5. Because of Theorem 4 it suffices to prove Theorem 5 for one of sums $S_i(x, k)$. We consider $S_1(x, k)$. So we start with (4.29). (We assume $k > 1$ since Theorem 5 is trivially true for $k = 1$. (See (5.1), (5.2) and Theorem 4.)

In (4.29) we first look at the contribution due to numbers for which

$$\frac{p_k}{p_1} < \frac{1}{(\log p_1)^{k+1}}. \tag{5.5}$$

We will get an upper bound for the contribution due to such numbers.

Let $0 < \beta < 1$ be a real number whose value will be specified later. Then write

$$\sum_{2 \leq p_1 \leq x} = \sum_{x^\beta \leq p_1 \leq x} + \sum_{m=0}^{\infty} \sum_{x^{\beta/2^{m+1}} \leq p_1 \leq x^{\beta/2^m}} . \tag{5.6}$$

In the interval $x^{\beta/2^{m+1}} \leq p_1 < x^{\beta/2^m}$ one has an upper bound for ψ given in (4.42), while for $x^\beta \leq p_1 \leq x$ we use the trivial upper bound

$$\psi\left(\frac{x}{p_1 \cdots p_k}, p_{k-1}\right) \leq \frac{x}{p_1 \cdots p_k} . \tag{5.7}$$

Then for numbers satisfying (5.5) together with $x^\beta \leq p_1 \leq x$, we have the following bound in (4.29)

$$O\left(x \sum_{x^\beta < p_1 \leq x} \frac{1}{p_1 (\log p_1)^{k+1}} \sum_{p_k \leq p_1} \frac{1}{p_k} \sum_{p_k \leq p_{k-1} \leq p_1} \frac{1}{p_{k-1}} \cdots \sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2}\right) . \tag{5.8}$$

Analysis similar to (4.32), (4.33) and (4.34) yields

$$\begin{aligned} &O\left(x \sum_{x^\beta \leq p_1 \leq x} \frac{1}{p_1 (\log p_1)^{k+1}} \sum_{p_k \leq p_1} \frac{1}{p_k} (\log\log p_1 - \log\log p_k)^{k-2}\right) \\ &\quad = O\left(x \sum_{x^\beta \leq p_1 \leq x} \frac{(\log\log p_1)^{k-1}}{p_1 (\log p_1)^{k+1}}\right) = O\left(\frac{x (\log\log x)^{k-1}}{\beta^{k+1} (\log x)^{k+1}}\right) \end{aligned} \tag{5.9}$$

using Lemma 1. To estimate the contribution due to integers satisfying (5.5) for the case $p_1 \leq x^\beta$, we use the decomposition of the last sum of (5.6). Then the upper bound method yields

$$O\left(\frac{x (\log\log x)^{k-1}}{(\log x)^{k+1}}\right) = O\left(\frac{x}{(\log x)^k}\right) \tag{5.10}$$

provided β is suitably chosen. Thus from (5.9) and (5.10) we conclude that the contribution due to terms satisfying (5.5) is given by (5.10), and is smaller than the asymptotic term we are seeking.

Next we observe that the contribution due to terms for which $p_1 = p_1(n)$ is small is negligible. For that purpose set

$$y = (\exp\{(\log x)^{2/3}\}) . \tag{5.11}$$

With y as in (5.11) we have by (4.38)

$$\psi(x, y) = O(x \exp\{-c_{16} (\log x)^{1/3}\}) . \tag{5.12}$$

So, if $p_1 = P_1(n) \leq y$, then

$$\begin{aligned} \sum_{2 \leq n \leq x} \frac{A(n) - P_1(n) - \cdots - P_{k-1}(n)}{P_1(n)} &\leq \sum_{\substack{2 \leq n \leq x \\ P_1(n) \leq y}} \Omega(n) \leq O\left(\sum_{\substack{2 \leq n \leq x \\ P_1(n) \leq y}} \log x\right) \\ &= O(\log x \psi(x, y)) = O(x \exp\{-c_{17} (\log x)^{1/3}\}) . \end{aligned} \tag{5.13}$$

Because of (5.13) and (5.10), we assume from now on that

$$\frac{p_1}{(\log p_1)^{k+1}} \leqq p_k \leqq p_1;\ P_1(n) = p_1 > \exp\{(\log x)^{2/3}\}\ . \tag{5.14}$$

Once we assume (5.14) we can rewrite Lemma 5 as

$$\begin{aligned}\psi\Big(\frac{x}{p_1p_2\cdots p_k}, p_{k-1}\Big) &= x\rho\Big(\frac{\log(x/p_1\cdots p_k)}{\log p_{k-1}}\Big) \\ &\quad + O\Big(\frac{xe^{-\alpha/4}\cdot\alpha}{p_1\cdots p_k\log p_{k-1}}\Big)\end{aligned} \tag{5.15}$$

where $\alpha = \log(x/p_1\cdots p_k)/\log p_{k-1}$.

The idea is to substitute (5.15) in (4.29). It is then easy to take care of the contribution due to the error term in (5.15) in (4.29) by observing that (5.14)

$$\log p_{k-1} \geqq \log p_k \sim \log p_1 > \frac{1}{2}\log p_1\ , \qquad x \geqq x_0\ . \tag{5.16}$$

This means if we substitute the O-term of (5.15) in (4.29), and use the upper bound method we get

$$O(x/(\log x)^k)\ . \tag{5.17}$$

The convergence of a series like (4.43) is ensured this time by the $e^{-\alpha/4}$ term in (5.15). Since (5.17) is smaller than the asymptotic term we are seeking, we may forget the contribution of the O-term in (5.15), in the sum (4.29).

As to the leading term of (5.15) we observe that

$$\rho\Big(\frac{\log(x/p_1\cdots p_k)}{\log p_{k-1}}\Big) = \rho\Big(\frac{\log x - \sum_{i=1}^{k}\log p_i}{\log p_{k-1}}\Big)\ . \tag{5.18}$$

By (5.14) we have

$$\log p_i = \log p_1 + O(\log\log p_1)\ , \qquad 1 \leqq i \leqq k\ . \tag{5.19}$$

Substituting (5.19) in (5.18) we get

$$\rho\Big(\frac{\log(x/p_1\cdots p_k)}{\log p_{k-1}}\Big) = \rho\Big\{\frac{\log x}{\log p_1} - k + O\Big(\frac{\log x\cdot\log\log p_1}{\log^2 p_1}\Big)\Big\}\ . \tag{5.20}$$

Using Lemma 4 to estimate (5.20) we get

$$\rho\Big(\frac{\log x}{\log p_1} - k\Big) + O\Big(\frac{\log x\cdot\log\log p_1}{\log^2 p_1\cdot\Gamma(\alpha)}\Big) \tag{5.21}$$

where α is as in (5.15).

Thus the factor ρ in the leading term of (5.15) is equal to the quantity in (5.21). Recall that our idea is to substitute (5.15) in (4.29) and estimate the sum. The contribution of the O-terms in (5.21) can be obtained by the upper bound method. There is a $\log x$ in the numerator, but a $\log^2 p_1$ in the denominator. This time the presence of $\Gamma(\alpha)$ in the denominator ensures convergence in a series like (4.43). Thus the upper bound method yields

$$O(x \log\log x/(\log x)^k) \tag{5.22}$$

as the contribution due to the O-term of (5.21). Thus we deduce that the main contribution from (4.29) comes by assuming (5.14) and replacing $\psi(x/p_1 \cdots p_k, p_{k-1})$ by

$$\frac{x}{p_1 \cdots p_k}\rho\left(\frac{\log x}{\log p_1} - k\right). \tag{5.23}$$

So we replace (4.29) by

$$\begin{aligned} &x \sum_{\exp\{(\log x)^{2/3}\} \leq p_1 \leq x} \frac{\rho((\log x/\log p_1) - k)}{p_1^2} \\ &\quad \times \sum_{p_1/(\log p_1)^{k+1} \leq p_k \leq p_1} \sum_{p_k \leq p_{k-1} \leq p_1} \frac{1}{p_{k-1}} \cdots \sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2}. \end{aligned} \tag{5.24}$$

To estimate (5.24) we use Lemma 3. First we get

$$\sum_{p_3 \leq p_2 \leq p_1} \frac{1}{p_2} = (\log\log p_1 - \log\log p_3) + O(\exp\{-c_5\sqrt{\log p_3}\}). \tag{5.25}$$

The contribution due to the O-term in (5.25) in (5.24) is taken care of by the upper bound method. This time the presence of ρ in (5.24) ensures convergence of a series like (4.43), because of (5.2). Actually every error term that arises in (5.24) by repeated use of Lemma 3 can be estimated by the upper bound method, yielding

$$O(x \exp\{-c_{18}\sqrt{\log x}\}). \tag{5.26}$$

So we need only look at the leading terms arising out of Lemma 3 in (5.24). After k-2 applications of the lemma we are left with

$$\begin{aligned} &x \sum_{\exp\{(\log x)^{2/3}\} \leq p_1 \leq x} \frac{\rho((\log x)/(\log p_1) - k)}{p_1^2} \\ &\quad \times \sum_{p_1/(\log p_1)^{k+1} \leq p_k \leq p_1} \frac{(\log\log p_1 - \log\log p_k)^{k-2}}{(k-2)!}. \end{aligned} \tag{5.27}$$

In (5.27) we use Lemma 2 to get

$$x \sum_{\exp\{(\log x)^{2/3}\} \leq p_1 \leq x} \frac{\rho((\log x)/(\log p_1)) - k)}{p_1^2} \times \left[\frac{p_1}{(\log p_1)^{k-1}} + O\left(\frac{p_1(\log\log p_1)}{(\log p_1)^k}\right)\right]. \tag{5.28}$$

As before, the O-term in (5.28) contributes

$$O(x \log\log x/(\log x)^k) \tag{5.29}$$

by use of the upper bound method. Finally the leading term in (5.28) is estimated by writing it as a Stieltjes integral. That is

$$\begin{aligned} x & \sum_{\exp\{(\log x)^{2/3}\} \leq p_1 \leq x} \frac{\rho((\log x)/(\log p_1) - k)}{p_1(\log p_1)^{k-1}} \\ &= x \int_{\mathrm{epx}\{(\log x)^{2/3}\}-}^{x+} \frac{\rho((\log x)/(\log y) - k) d\pi(y)}{y(\log y)^{k-1}} \\ &= x \int_{\exp\{(\log x)^{2/3}\}}^{x} \frac{\rho((\log x)/(\log y) - k)}{y(\log y)^k} dy \\ &\quad + x \int_{\exp\{(\log x)^{2/3}\}-}^{x+} \frac{\rho((\log x)/(\log y) - k)}{y(\log y)^{k-1}} \\ &\quad \times d\{\pi(y) - li(y)\} = I_3 + I_4 . \end{aligned} \tag{5.30}$$

We can bound I_4 rather easily. First observe that $|\rho| \leqq 1$. Ignoring o, we integrate by parts, and use (4.1) to deduce

$$I_4 = O((x \exp\{-c_{19}(\log x)^{1/3}\}) . \tag{5.31}$$

To estimate I_3 write $y = x^{1/s}$. Then

$$\begin{aligned} I_3 &= \frac{x}{(\log x)^{k-1}} \int_1^{(\log x)^{1/3}} \rho(s - k) s^{k-2} ds \\ &= \frac{x}{(\log x)^{k-1}} \left[\int_1^{\infty} - \int_{(\log x)^{1\ 3}}^{\infty}\right] \\ &= a_k' \frac{x}{(\log x)^{k-1}} + O\left(\frac{x}{(\log x)^k}\right) \end{aligned} \tag{5.32}$$

because of (5.2). So Theorem 5 follows from (5.32) and the preceding estimates.

REMARKS. Note that we have actually shown that

$$S_1(x, k) = a_k' \frac{x}{(\log x)^{k-1}} + O\left(\frac{x \log\log x}{(\log x)^k}\right) . \tag{5.33}$$

Observe that $S_1(x, k)$ is the largest of the four sums and $S_4(x, k)$ is the smallest. Therefore, because of (4.55), we deduce a stronger form of Theorem 5, namely

$$S_i(x, k) = a'_k \frac{x}{(\log x)^{k-1}} + O\left(\frac{x \log \log x}{(\log x)^k}\right) \tag{5.34}$$

for $i = 1, 2, 3, 4$.

Thus we have proved all the statements of Theorem B, except the relation between a_k and e^γ. We do this in the next section.

6. The constants a'_k. It is obvious from Theorem 4 or (5.1) and (5.2) that $a'_1 = 1$. So we suppose $k \geqq 2$. For $k \geqq 2$ write

$$\begin{aligned} a'_k &= \int_1^\infty \rho(s-k)s^{k-2}ds = \int_k^\infty \rho(s-k)s^{k-2}ds = \int_0^\infty \rho(t)(t+k)^{k-2}dt \\ &= \sum_{j=0}^{k-2} \binom{k-2}{j} k^{k-2-j} \int_0^\infty \rho(t)t^j dt = \sum_{j=0}^{k-2} \binom{k-2}{j} k^{k-2-j} f_j \end{aligned} \tag{6.2}$$

where

$$f_j = \int_0^\infty \rho(t)t^j dt\ . \tag{6.3}$$

In a recent paper, Knuth and Pardo [8], have studied the behavior of

$$\psi_k(x, y) = \sum_{\substack{1 \leqq n \leqq x \\ P_k^*(n) \leqq y}} 1\ . \tag{6.4}$$

In the course of their investigations they show

$$f_j = e^\gamma g_j \tag{6.5}$$

where γ is Euler's constant and the g_j are recursively defined by

$$g_0 = g_1 = 1\ , \qquad g_j = \frac{1}{j} \sum_{1 \leqq i \leqq j} \binom{j}{i} g_{j-i}\ , \qquad j \geqq 0\ . \tag{6.6}$$

Combining (6.2), (6.5), and (6.6) we infer that a'_k is a rational multiple of e^γ for $k \geqq 2$. For instance

$$a'_2 = f_0 = e^\gamma g_0 = e^\gamma\ .$$

That completes the proof of Theorem B.

Acknowledgments. K. Alladi is grateful to Professors R. Miech and E. G. Straus for their advice and encouragement. They went through an early draft of this paper and made helpful suggestions. We also thank the referee for his critical comments and suggestions.

References

1. K. Alladi, *Analogues to the Hardy-Ramanujan theorems*, in Proceedings of the

Conference of Number Theory and Computational Methods, Publications of "Matscience", Madras, India (1977).

2. K. Alladi and P. Erdös, *On an additive arithmetic function*, Pacific J. Math., **71** (2), (1977), 275–294.

3. N. G. de Bruijn, *On the number of positive integers* $\leqq \times$ *and free of prime factors* $> y$, Indag. Math., **13** (1951), 50–60.

4. K. Chandrasekharan, *Arithmetical Functions*, Springer-Verlag, New York, 1970.

5. P. D. T. A. Elliott, *On a problem of Hardy and Ramanujan*, Mathematika **23**, (1976), 10–17.

6. G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number n*, Quarterly J. Math., **58** (1917), 76–92.

7. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, Clarendon, 1960, 4-Ed.

8. D. E. Knuth and L. T. Pardo, *Analysis of a simple factorisation algorithm*, (to appear).

9. W. J. LeVeque, *Topics in Number Theory*, Vol. 2, Addison-Wesley, Reading, Mass., 1956.

10. K. Norton, *Numbers with small prime factors and the least* kth-*power non-residue*, Memoirs of the A. M. S., **106**, Providence, Rhode Island, 1971.

11. J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math., **6** (1962), 64–94.

Received April 3, 1978 and in revised form October 16, 1978. This work was supported in part under National Science Foundation Grants No. MCS 77-01780 and MCS 78-02685.

UNIVERSITY OF CALIFORNIA
DEPARTMENT OF MATHEMATICS
LOS ANGELES, CALIFORNIA 90024
AND
MATHEMATICS INSTITUTE
THE HUNGARIAN ACADEMY OF SCIENCES
BUDAPEST, HUNGARY

Current Address of first author: University of Michigan
Department of Mathematics
Ann Arbor, MI 48109