COLLOQUIA MATHEMATICA SOCIETATIS JÁNOS BOLYAI
34. TOPICS IN CLASSICAL NUMBER THEORY
BUDAPEST (HUNGARY), 1981.

ESTIMATES FOR CERTAIN SUMS INVOLVING THE LARGEST PRIME FACTOR OF AN INTEGER

A. IVIĆ and C. POMERANCE

1. INTRODUCTION AND STATEMENT OF RESULTS

Let $P(n)$ for $n \geq 2$ denote the largest prime factor of an integer n, and let $\beta(n)$ and $B(n)$ denote the additive functions

$$\beta(n) = \sum_{p \mid n} p, \qquad B(n) = \sum_{p^a \| n} ap,$$

where as usual p denotes primes and $p^a \| n$ means that p^a divides n but p^{a+1} does not. Several results concerning sums with the functions $P(n)$, $\beta(n)$ have been recently obtained. Thus for instance it was proved in [8] that

(1.1) $$\sum_{2\le n\le x} 1/P(n) = x\exp\{-(2\log x\cdot \log\log x)^{1/2} + o((\log x\cdot \log\log\log x)^{1/2})\},$$

and the same formula holds also for sums of reciprocals of $\beta(n)$ and $B(n)$. Moreover [5] contains a proof of

(1.2) $$\sum_{2\le n\le x} B(n)/\beta(n) = x + x\exp(-c(\log x\cdot \log\log x)^{1/2}),\quad c>0,$$

and for additional results concerning other related sums the reader is referred to [1], [2] and [6]. Our aim here is to prove a result which will yield considerable sharpenings of (1.1) and (1.2) and will have other applications as well. In what follows we suppose that $r\ge 0$ is arbitrary but fixed, and we let

(1.3) $$g_r(x) = \frac{\log_3 x + \log(1+r) - 2 - \log 2}{2\log_2 x}\left(1+\frac{2}{\log_2 x}\right) - \frac{(\log_3 x + \log(1+r) - \log 2)^2}{8\log_2^2 x},$$

(1.4) $$L = L(x) = (\log x\cdot \log_2 x)^{1/2},$$

where $\log_k x$ is the k-fold iterated logarithm. Further we define

(1.5) $$S_r(x) = \sum_{2\le n\le x} 1/P^r(n),\quad T_r(x) = \sum_{\substack{2\le n\le x\\ P^2(n)\mid n}} 1/P^r(n).$$

Thus for instance $T_0(x)$ represents the number of $n\le x$ of the form $n=p_1^{a_1}\cdots p_k^{a_k}$ where $p_1>\cdots>p_k$ and $a_1\ge 2$. Our result is the following

THEOREM.

(1.6) $$S_r(x) = x\exp\{-(2r)^{1/2}L(x)(1+g_{r-1}(x) + O(\log_3^3 x/\log_2^3 x))\},$$

(1.7) $T_r(x) =$

$$= x \exp\{-(2r+2)^{1/2}L(x)(1+g_r(x) + O(\log_3^3 x/\log_2^3 x))\},$$

From (1.6) and (1.7) we shall deduce the following

COROLLARY.

(1.8) $\sum_{2\le n\le x} 1/\beta(n) =$

$$= x \exp\{-2^{1/2}L(x)(1+g_0(x) + O(\log_3^3 x/\log_2^3 x))\} ,$$

(1.9) $\sum_{2\le n\le x} 1/B(n) =$

$$= x \exp\{-2^{1/2}L(x)(1+g_0(x) + O(\log_3^3 x/\log_2^3 x))\} ,$$

(1.10) $\sum_{2\le n\le x} B(n)/\beta(n) =$

$$= x + x \exp\{-2^{1/2}L(x)(1+g_0(x) + O(\log_3^3 x/\log_2^3 x))\} ,$$

(1.11) $\sum_{2\le n\le x} 1/\beta(n) - \sum_{2\le n\le x} 1/B(n) =$

$$= x \exp\{-2L(x)(1+g_1(x) + O(\log_3^3 x/\log_2^3 x))\} .$$

The method of proof of our theorem may be used for the estimation of other related arithmetic sums. If $\omega(n)$ and $\Omega(n)$ denote as usual the number of distinct prime factors of n and the number of total prime factors of n respectively, then following the proof of (1.10) and (1.11) we obtain

(1.12) $\sum_{2\le n\le x} \Omega(n)/P(n) - \sum_{2\le n\le x} \omega(n)/P(n) =$

$$= x \exp\{-2^{1/2}L(x)(1+g_0(x) + O(\log_3^3 x/\log_2^3 x))\} ,$$

$$(1.13) \qquad \sum_{2\le n\le x} 1/(\beta(n)\omega(n)) - \sum_{2\le n\le x} 1/(B(n)\Omega(n)) =$$

$$= x \exp\{-2^{1/2}L(x)(1+g_0(x) + O(\log_3^3 x/\log_2^3 x))\}.$$

Also if $U_r(x)$ denotes the number of $n\le x$ of the form $n=p_1^{a_1}\cdots p_k^{a_k}$, where $p_1 >\dots> p_k$ and $a_1\ge r$, where $r\ge 2$ fixed, then we have

$$(1.14) \qquad U_r(x) =$$

$$= x \exp\{-(2r-2)^{1/2}L(x)(1+g_{r-2}(x) +$$

$$+ O(\log_3^3 x/\log_2^3 x))\}.$$

2. ESTIMATES FOR $\psi(x,y)$

The proof of (1.1) which was given in [9] depended on estimates for

$$\psi(x,y) = \sum_{n\le x, P(n)\le y} 1,$$

the number of n not exceeding x all of whose prime factors do not exceed y. Our proof of (1.6) and (1.7) will also depend on estimates for $\psi(x,y)$, and the formula for $S_1(x)$ furnished by (1.6) is sharper than (1.1). This is due to a modified method of proof and to the fact that we are now able to use the following recent result proved in [4]:

$$(2.1) \qquad \psi(x,y) =$$

$$= x \exp\{-u(\log u + \log_2 u - 1 + \frac{\log_2 u-1}{\log u} +$$

$$+ O((\log_2 u/\log u)^2))\},$$

where $e^e \le u = \log x/\log y \le (1-\varepsilon)\log x/\log_2 x$, and the O-constant depends only on ε. The upper bound implicit in (2.1) can already be found in N.G. DE BRUIJN [3].

However the lower bound improves on that of H. HALBERSTAM [7], the estimate used in [8]. In particular [7] gives

$$(2.2)\qquad \psi(x,y) \gg x\exp(-u(\log u + \log_2 u + R(u))),$$

where $u\geq 3$, $y\geq y_1(u)$, and $R(u)$ is explicitly given and is $o(1)$. Though [7] gives no explicit evaluation of $y_1(u)$, it is seen by following Halberstam's proof that one may take $y_1(u) = \exp(\log^{5/3+\varepsilon} u)$ for $u\geq u_0(\varepsilon)$; the main observation is that on p.106 of [7] instead of $O(1/\log x)$ for the error term in the formula for $\sum_{p\leq x} 1/p$ one can use the sharper error term $O(\exp(-\log^{3/5-\varepsilon} x))$ which comes from the strongest version of the prime number theorem. Thus certainly one can apply (2.2) in the range needed in the proof of (1.1), but nevertheless (2.2) is now superseded by (2.1). Though (2.1) is not an asymptotic formula in the strict sense, it is particularly well-suited for our purposes, since it yields an equality for $\psi(x,y)$ for a large range of $u=\log x/\log y$. For ranges of validity of a true asymptotic formula for $\psi(x,y)$ of the form

$$\psi(x,y) = (1 + o(1))x\rho(u),$$

the reader should consult the papers of DE BRUIJN [3].

3. PROOF OF THE THEOREM

We begin by noting that

$$(3.1)\qquad T_r(x) = \sum_{\substack{p^2 m\leq x\\ P(m)\leq p}} p^{-r} = \sum_{p\leq x^{1/2}} p^{-r}\psi(xp^{-2},p),$$

$$(3.2)\qquad S_r(x) = \sum_{\substack{pm\leq x\\ P(m)\leq p}} p^{-r} = \sum_{p\leq x} p^{-r}\psi(xp^{-1},p),$$

so that the expressions for $T_r(x)$ and $S_r(x)$ are similar, and so will be the proofs. We give therefore all the details only for (1.7). First from (3.1) we have

$$(3.3)\qquad T_r(x) = \sum_{p<\exp(L/(4r+4))} + \sum_{\exp(L/(4r+4))\leq p\leq \exp((4r+4)L)} + \sum_{p>\exp((4r+4)L)} = \Sigma_1 + \Sigma_2 + \Sigma_3 ,$$

where $L=L(x)$ is given by (1.4). Note that from (2.1) one has

$$(3.4)\qquad \phi(x,\exp(L/(4r+4))) = x\exp((-1+o(1))(2r+2)L),$$

which gives

$$(3.5)\qquad \Sigma_1 = \sum_{p<\exp(L/(4r+4))} p^{-r}\phi(xp^{-2},p) \le$$

$$\le \exp(L/(4r+4))\cdot\phi(x,\exp(L/(4r+4))) =$$

$$= x\exp(((4r+4)^{-1} - (2r+2) + o(1))L).$$

Note that $(4r+4)^{-1} - (2r+2) < -(2r+2)^{1/2}$ for all $r\ge 0$, so that Σ_1 is negligible.

Next, trivially we have

$$(3.6)\qquad \Sigma_3 = \sum_{p>\exp((4r+4)L)} p^{-r}\phi(xp^{-2},p) \le$$

$$\le x\sum_{p>\exp((4r+4)L)} p^{-2} \ll x\exp(-(4r+4)L),$$

so that Σ_3 is negligible as well.

For the remaining sum

$$\Sigma_2 = \sum_{\exp(L/(4r+4))\le p\le\exp((4r+4)L)} p^{-r}\phi(xp^{-2},p),$$

we further restrict the interval of summation. Let

$$(3.7)\qquad T_{r,c}(x) = \sum_{\exp(cL-1)\le p\le\exp(cL)} p^{-r}\phi(xp^{-2},p)$$

and let c_0 denote that value of $c\in[(4r+4)^{-1}, 4r+4]$ for which $T_{r,c}(x)$ is maximal. Then clearly

$$(3.8)\qquad T_{r,c_0}(x) \le \Sigma_2 \le (4r+4)LT_{r,c_0}(x).$$

Thus from (3.3), (3.5), (3.6) and (3.8) it will suffice to show that $T_{r,c_0}(x)$ is equal to the right-hand side of (1.7). For $T_{r,c}(x)$ we have

$$u = \log xp^{-2}/\log p = \log x/\log p - 2 =$$

$$= \log x/cL + O(1),$$

$$\log u = \tfrac{1}{2}\log_2 x - \tfrac{1}{2}\log_3 x - \log c + O((\log_2 x/\log x)^{1/2}),$$

$$\log_2 u = \log_3 x - \log 2 - (\log_3 x + 2\log c)/\log_2 x +$$

$$+ o(\log_3^2 x/\log_3^2 x)$$

and therefore (2.1) gives

$$(3.9) \qquad T_{r,c}(x) = \sum_{\exp(cL-1)\le p\le \exp(cL)} xp^{-r-2} \times$$

$$\times \exp\{-u(\log u + \log_2 u - 1 + \frac{\log_2 u-1}{\log u} +$$

$$+ o((\frac{\log_2 u}{\log u})^2))\} =$$

$$= x \exp\{-L(x)((1+r)c + (2c)^{-1}(1+\varepsilon(x,c)))\},$$

where

$$(3.10) \qquad \varepsilon(x,c) = \frac{\log_3 x - 2\log c - 2\log 2 - 2}{\log_2 x} +$$

$$+ \frac{2\log_3 x - 4\log c - 4\log 2 - 4}{\log_2^2 x} + o((\log_3^2 x/\log_2^3 x)).$$

To obtain the maximal value we set

$$(3.11) \qquad c = ((1+d)/(2+2r))^{1/2},$$

where $d=d(x,r)$ is to be suitably determined. Before we proceed to determine d precisely, we may remark that using weaker estimates for $\phi(x,y)$ than (2.1) we may obtain a weaker result than our theorem. Specifically, if we use [3] for the upper bound for $\phi(x,y)$ and [7] for the lower bound for $\phi(x,y)$, then we have (3.9) with $\varepsilon(x,c) = (\log_3 x + O(1))/\log_2 x$ only. Then obviously one has to choose in (3.11) $d = (\log_3 x + O(1))/\log_2 x$ if $T_{r,c}(x)$ is to be maximal, and this leads to

$$T_r(x) =$$

$$= x \exp\{-(2r+2)^{1/2} L(x)(1+\log_3 x/(2\log_2 x) +$$

$$+ O(1/\log_2 x))\},$$

and to a corresponding result for $S_r(x)$ with $(2r)^{1/2}$ instead of $(2r+2)^{1/2}$. This is weaker than (1.6) and (1.7), but still improves (1.1). For a more precise determination of d in (3.11) with $\varepsilon(x,c)$ given by (3.10) we use

$$2\log c = \log(1+d) - \log(2+2r) =$$

$$= d + O(d^2) - \log 2 - \log(1+r),$$

$$(1+d)^{1/2} = 1 + \frac{1}{2} d - \frac{1}{8} d^2 + O(d^3),$$

$$(1+d)^{-1/2} = 1 - \frac{1}{2} d + \frac{3}{8} d^2 + O(d^3).$$

Then we have

$$(1+r)c + (2c)^{-1}(1+\varepsilon(x,c)) =$$

$$= \left(\frac{1+r}{2}\right)^{1/2}\left(2 + \frac{1}{4} d^2 + O(d^3) + \varepsilon(x,c)\left(1 - \frac{d}{2}\right)\right) =$$

$$= \left(\frac{1+r}{2}\right)^{1/2}\left(F_r(d) + O(\log_3^3 x/\log_2^3 x)\right),$$

where

$$F_r(d) = 2 + \frac{1}{4} d^2 + \left(1 - \frac{d}{2}\right) \frac{\log_3 x + \log(1+r) - \log 2}{\log_2 x} \times$$

$$\times \left(1 + \frac{2}{\log_2 x}\right) - \frac{2}{\log_2 x}\left(1 + \frac{2}{\log_2 x}\right).$$

Now $F_r(d)$ is a quadratic function in d whose

minimal value $F(d_0)$ is attained for

$$d = d_0 = \frac{\log_3 x + \log(1+r) - \log 2}{\log_2 x}\left(1 + \frac{2}{\log_2 x}\right),$$

and equals

$$F_r(d_0) = 2 + \frac{\log_3 x + \log(1+r) - 2 - \log 2}{\log_2 x} \times$$

$$\times \left(1 + \frac{2}{\log_2 x}\right) - \frac{(\log_3 x + \log(1+r) - \log 2)^2}{4 \log_2^2 x} +$$

$$+ O(\log_3^3 x/\log_2^3 x) = 2 + 2g_r(x) + O(\log_3^3 x/\log_2^3 x),$$

where $g_r(x)$ is given by (1.3). Therefore we obtain (1.7), and the proof of (1.6) is analogous.

4. PROOF OF THE COROLLARY

From (1.6) with $r=1$ and

$$P(n) \le \beta(n) \le B(n) \le \Omega(n)P(n) \le (P(n)\log n)/\log 2$$

we obtain at once (1.8) and (1.9). For (1.10) first note that

$$(4.1)\qquad \sum_{2\le n\le x} B(n)/\beta(n) - x =$$

$$= \sum_{2\le n\le x} (B(n)-\beta(n))/\beta(n) + O(1) \ge$$

$$\ge \log 2(\log x)^{-1} \sum_{\substack{2\le n\le x\\ P^2(n)\mid n}} (B(n)-\beta(n))/P(n) + O(1) \ge$$

$$\ge T_0(x)/(2\log x) + O(1) =$$

$$= x\exp\{-2^{1/2}L(x)(1+g_0(x) +$$

$$+ O(\log_3^3 x/\log_2^3 x))\},$$

where we used (1.7) with $r=0$ and the fact that $B(n) - \beta(n) \ge P(n)$ when $P^2(n)\mid n$. For the upper bound we have

$$\sum_{2\le n\le x} (B(n)-\beta(n))/\beta(n) \ll$$

$$\ll \sum_{\substack{2\le n\le x\\ P(n)\,\|\, n}} (B(n)-\beta(n))/P(n) + T_0(x)\log x,$$

and the proof of (1.10) will be finished when we show that

$$(4.2)\qquad S = \sum_{\substack{2\le n\le x\\ P(n)\,\|\, n}} (B(n)-\beta(n))/P(n) \le$$

$$\le x\exp\{-2^{1/2}L(x)(1+g_0(x) +$$

$$+ O(\log_3^3 x/\log_2^3 x))\}.$$

To prove (4.2) observe that every n may be written uniquely as $n = q(n)s(n)$, $(q(n),s(n)) = 1$, where $q(n)$ is square-free and $s(n)$ is square-full (n is square-full if $p^2\mid n$ whenever $p\mid n$). Since there are $O(x^{1/2})$ square-full numbers not exceeding x we have

$$\sum_{\substack{n\le x\\ s(n)\ge\exp(5L)}} (B(n)-\beta(n))/P(n) \ll$$

$$\ll \log x \sum_{\substack{n\le x\\ s(n)\ge\exp(5L)}} 1 \le$$

$$\le \log x \sum_{\substack{s \text{ square-full}\\ s\ge\exp(5L)}} xs^{-1} \ll x\exp(-2L).$$

Denoting square-free numbers by q and square-full numbers by s we have then

$$S = \sum_{\substack{n\le x\\ s(n)<\exp(5L),\ P(n)\| n}} (B(n)-\beta(n))/P(n) + O(x\exp(-2L))$$

$$= \sum_{s<\exp(5L)} (B(s)-\beta(s)) \sum_{\substack{q\le x/s,(q,s)=1,\\ P(q)>P(s)}} 1/P(q)+O(x\exp(-2L$$

$$\ll \sum_{s<\exp(5L)} s^{1/2}\log s \sum_{n\le x/s} 1/P(n) +$$

$$+ O(x\exp(-2L)) = x \sum_{s<\exp(5L)} s^{-1/2}\log s \times$$

$$\times \exp\{-2^{1/2}L(x/s)(1+g_0(x/s)+$$

$$+ O(\log_3^3 x/\log_2^3 x))\} + O(x\exp(-2L)) =$$

$$= x\exp\{-2^{1/2}L(x)(1+g_0(x)+$$

$$+ O(\log_3^3 x/\log_2^3 x))\}.$$

Here we used (1.6) with $r=1$. Also for $(q,s)=1$, q squarefree, s square-full we have

$$B(sq) - \beta(sq) = B(s) - \beta(s),$$

$$B(s) - \beta(s) \le \Omega(s)\log s \ll s^{1/2}\log s,$$

$$\sum_{s<\exp(5L)} s^{-1/2}\log s \ll L^2,$$

and for $s \le \exp(5L)$

$$L(x)(1+g_0(x)) - L(x/s)(1+g_0(x/s)) \ll$$

$$\ll \log_3^3 x/\log_2^3 x.$$

Finally to see that (1.11) holds note that by (1.7) with $r=1$ we have

$$\sum_{2\le n\le x} (1/\beta(n) - 1/B(n)) \gg$$

$$\gg \log^{-2}x \sum_{\substack{2\le n\le x\\ P^2(n)|n}} (B(n)-\beta(n))/P^2(n) \gg$$

$$\gg \log^{-2}x \sum_{\substack{2\le n\le x\\ P^2(n)|n}} 1/P(n) =$$

$$= T_1(x)\log^{-2}x =$$

$$= x \exp\{-2L(x)(1 + g_1(x) + o(\log_3^3 x/\log_2^3 x))\}.$$

For the upper bound we have

$$\sum_{2\le n\le x} (1/\beta(n) - 1/B(n)) \ll$$

$$\ll T_1(x)\log x + \sum_{\substack{2\le n\le x\\ P(n)\| n}} (B(n) - \beta(n))/P^2(n).$$

Here again by (1.7) $T_1(x)\log x$ is of the right order of magnitude, while the remaining sum is estimated in the same way as was the sum S in (4.2).

REFERENCES

[1] Alladi, K. and Erdős P., On an additive arithmetic function, *Pacific J. Math.* 71 (1977), 275-294.

[2] Alladi, K. and Erdős P., On the asymptotic behaviour of large prime factors of an integer, *Pacific J. Math.* 82 (1979), 295-315.

[3] De Bruijn, N.G., On the number of positive integers $\le x$ and free of prime factors $>y$, *Nederl. Akad. Wet. Proc.* 54 (1951), 50-60, and *ibid.* 28 (1966), 239-247.

[4] Canfield, E.R., Erdős P. and Pomerance, C., On a problem of Oppenheim concerning "Factorisatio Numerorum", to appear in *Journal of Number Theory.*

[5] De Koninck, J.-M. and Ivić, A., *Topics in arithmetical functions*, Notas de Matemática 72, Amsterdam, 1980.

[6] Erdős P. and Ivić, A., Estimates for sums involving the largest prime factor of an integer and certain related additive functions, to appear in *Studia Sci. Math. Hung.*

[7] Halberstam, H., On integers all of whose prime factors are small, *Proc. London Math. Soc.* (3) 21 (1970), 102-107.

[8] Ivić, A., Sum of reciprocals of the largest prime factor of an integer, *Archiv Math.* 36 (1980), 57-61.

A. Ivić
Katedra Matematike RGF-a
Universiteta u Beogradu
Djušina 7, 11000 Beograd
Yugoslavia

C. Pomerance
Department of Mathematics
University of Georgia
Athens, Georgia 30602
USA