On the average ratio of the smallest and largest prime divisor of n

by P. Erdös and J.H. van Lint

Mathematical Institute, Hungarian Academy of Sciences, Budapest, Hungary
Department of Mathematics, Eindhoven University of Technology, P.O. Box 513, Eindhoven, the Netherlands

Communicated at the meeting of September 28, 1981

1. THE MAIN THEOREM

Let $p(n)$ denote the smallest prime factor of n and let $P(n)$ be the largest prime factor of n; (throughout this paper n stands for an integer >1). In [2] J. van de Lune considered the asymptotic behavior of some sums in which the terms are elementary functions of $p(n)$ and $P(n)$. For the sum $\sum_{n\le x} p(n)/P(n)$ he did not go beyond showing that it is $o(x)$, $(x\to\infty)$. The purpose of this note is to provide a more accurate estimate. In fact we shall prove the following theorem.

THEOREM 1.

$$\sum_{n\le x} p(n)/P(n) = \pi(x)(1+o(1)),\ (x\to\infty).$$

We first give an elementary proof of this theorem. In Section 2 we estimate the remainder term.

The proof of Theorem 1 is in a number of steps. Denote the sum by S.

(i) Clearly the primes contribute $\pi(x)$ to S.

(ii) Since there are $O(x^{1/2})$ prime powers $\le x$ their contribution to S is $O(x^{1/2}) = o(\pi(x))$.

(iii) We now consider the integers $n\le x$ for which $P(n)\ge p(n)(\log n)^2$. Their contribution to S is at most

$$\sum_{1<n\le x} (\log n)^{-2} = O(x(\log x)^{-2}) = o(\pi(x)).$$

In the following steps we assume $P(n)<p(n)(\log n)^2$.

(iv) We consider integers n with two prime factors, i.e. $n=p^\alpha q^\beta$, where $p=P(n)$, $q=p(n)$ and $p(\log x)^{-2}<q<p$. If we require that $p<x^{2/5}$ then there are $O((\pi(x^{2/5})\log x)^2)=O(x^{4/5})$ such numbers n and again their contribution to S is $o(\pi(x))$.

If, on the other hand, $p\geq x^{2/5}$ then $q\geq x^{2/5}(\log x)^{-2}$ and then we may assume that $\alpha=\beta=1$ (with a finite number of exceptions). Ignoring the restriction on p we simply consider all numbers $n=pq$. We split these into two classes: those with $p\leq x^{1/2}(\log x)^{1/3}$ resp. $p>x^{1/2}(\log x)^{1/3}$. For the first class the number of choices for p is $\pi(x^{1/2}(\log x)^{1/3})=O(x^{1/2}(\log x)^{-2/3})$. Therefore the number of integers in the first class is $o(\pi(x))$. For the second class $q<x^{1/2}(\log x)^{-1/3}$ and therefore the corresponding terms q/p in S are less than $(\log x)^{-2/3}$. It is known (cf. [1], Theorem 437) that the number of integers $n=pq\leq x$ is

$$O\left(\frac{x\log\log x}{\log x}\right).$$

Therefore the second class also contributes $o(\pi(x))$ to S.

In Section 2 we shall show that with a little more work one can prove that the contribution of the numbers $n=pq$ to S has order $x(\log x)^{-2}$. At this point the assertion of the theorem is already quite plausible since we do not expect a more significant contribution from the integers with more than two prime factors although it will take some work to show this.

(v) In the remaining part of the proof we consider integers with at least three distinct prime factors subject to the restriction $P(n)<p(n)(\log n)^2$. We introduce two parameters

$$A:=\exp\left(\frac{\log x}{10\log\log x}\right),\quad m:=5\log\log x.$$

If $p_1,\dots,p_k$ are the prime factors of $n\leq x$, say $n=p_1^{\alpha_1}\dots p_k^{\alpha_k}$, then

$$\sum_{i=1}^{k}\alpha_i\log p_i\leq\log x.$$

Therefore there are at most $(\log x)^k$ choices for the exponents. It follows that the integers n with less than m distinct prime factors and $P(n)\leq A$ contribute at most

$$\sum_{3\leq k\leq m}\binom{\pi(A)}{k}(\log x)^k\leq(\pi(A)\log x)^m=O(x^{1/2+\varepsilon})=o(\pi(x))$$

to S.

Next, consider the cases where n has at least m distinct prime factors. Let $d(n)$ denote the number of divisors of n. It is known that $\sum_{n\leq x}d(n)=O(x\log x)$, (cf. [3]).

It follows that there are at most

$$O\left(\frac{x\log x}{2^m}\right)=o\left(\frac{x}{(\log x)^2}\right)=o(\pi(x))$$

integers n with at least m distinct prime factors and their contribution to S is therefore $o(\pi(x))$.

(vi) It remains to consider the case $P(n)>A$. Given primes p, q, r with

$$p(\log x)^{-2}\leq q<r<p,\ p>A$$

there are at most $x/(pqr)$ terms in S such that pqr divides n. Furthermore (cf. (2.7) below).

$$\sum_{p(\log x)^{-2}\leq q<r\leq p}(qr)^{-1}\leq\{\sum_{p(\log x)^{-2}\leq q\leq p}q^{-1}\}^2$$
$$\leq\{\log\log p-\log\log[p(\log x)^{-2}]+O((\log p)^{-1})\}^2$$
$$\leq c\left(\frac{\log\log x}{\log p}\right)^2\leq c\left(\frac{\log\log x}{\log A}\right)^2\leq c'\frac{(\log\log x)^4}{(\log x)^2}.$$

It follows that the remaining contribution to S is at most

$$\frac{c'x(\log\log x)^4}{(\log x)^2}\sum p^{-1}=O\left(\frac{x(\log\log x)^5}{(\log x)^2}\right)=o(\pi(x)).$$

This completes the proof.

2. THE REMAINDER TERM

The method of Section 1 was fairly elementary. In order to estimate the remainder term we shall now first exclude the integers n with $P(n)\leq x^{1/\log\log x}$. We use a method due to R.A. Rankin (cf. [4]).

Let $y:=x^{1/\log\log x}$ and as usual let $\Psi(x,y)$ denote the number of integers n with $n\leq x$, $P(n)\leq y$.

Then, for $\eta>0$ we have

$$\Psi(x,y)\leq\sum_{n\leq x,P(n)\leq y}\left(\frac{x}{n}\right)^\eta\leq x^\eta\prod_{p\leq y}(1-p^{-\eta})^{-1}.$$

Taking $\eta=1-\varepsilon$ we find

$$\log\Psi(x,y)\leq\log x-\varepsilon\log x-\int_2^y\log(1-t^{-1+\varepsilon})d\pi(t).$$

A suitable choice for ε is $\varepsilon=5(\log\log x)/\log x$. We have $t^\varepsilon\leq y^\varepsilon=O(1)$. Hence

$$\log\Psi(x,y)\leq\log x-5\log\log x+O(\int_2^y t^{-1}d\pi(t))$$

and it then follows from (2.7) below that

(2.1) $\quad\Psi(x,y)=o(x/(\log x)^3)$.

In fact we can replace the exponent 3 by any $k>0$.

In the remainder of this section we shall restrict ourselves to

(2.2) $\quad P(n)>x^{1/\log\log x}$.

As in Section 1 it is obvious that the contribution to S of the integers n with $p(n) \le P(n)/(\log x)^3$ is $O(x/(\log x)^3)$. From now on we consider only n with

(2.3) $\quad P(n) < p(n) \cdot (\log x)^3.$

In the following we shall use a number of results which are direct consequences of the prime number theorem. We illustrate one of these. We have

$$\sum_{p \le x} p = \int_{2-\varepsilon}^{x} t d\pi(t) = x\pi(x) - \int_{2-\varepsilon}^{x} \pi(t) dt$$

and then the prime number theorem yields

$$\sum_{p \le x} p = \tfrac{1}{2} \frac{x^2}{\log x} (1 + o(1)) \; (x \to \infty). \tag{2.4}$$

In the same way one finds

$$\sum_{p \le x} \frac{p}{\log p} = \tfrac{1}{2} \frac{x^2}{(\log x)^2} (1 + o(1)) \; (x \to \infty). \tag{2.5}$$

$$\sum_{p > x} \frac{1}{p^3} = \tfrac{1}{2} \frac{1}{x^2 \log x} (1 + o(1)) \; (x \to \infty). \tag{2.6}$$

We shall often use the well known result (cf. [3])

$$\sum_{p \le x} \frac{1}{p} = \log \log x + A + O\left(\frac{1}{\log x}\right) \; (x \to \infty). \tag{2.7}$$

We now state the result of this section.

THEOREM 2.

$$\sum_{n \le x} \frac{p(n)}{P(n)} = \frac{x}{\log x} + \frac{3x}{(\log x)^2} + o\left(\frac{x}{(\log x)^2}\right).$$

PROOF.

(i) We saw in Section 1 (i) and (ii) that the primes and prime powers contribute $\pi(x) + O(x^{1/2})$ to S. By the prime number theorem

$$\pi(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + O\left(\frac{x}{(\log x)^3}\right).$$

(ii) Consider integers n of the form $a^2 b$, where $a > t$. The number of such integers $n \le x$ is less than $\sum_{a > t} x/a^2 < x/t$. From (2.2) and (2.3) we have

$$p(n) > x^{1/\log \log x} (\log x)^{-3}.$$

From these two observations it follows that from now on we may assume that n is *squarefree* since the contribution to S of the integers n which are not squarefree is less than

$$x^{1 - 1/\log \log x} (\log x)^3 = O(x/(\log x)^3).$$

(iii) We consider integers n with two prime factors, say $n=pq$ with $q<p$, satisfying (2.2) and (2.3). To find the contribution to S we first assume $p\le x^{1/2}$. Using (2.4) and (2.5) we find that the contribution of these n is

$$\sum_{q<p\le x^{1/2}} \frac{q}{p} = \sum_{p\le x^{1/2}} \frac{1}{p} \sum_{q<p} q =$$
$$= \tfrac{1}{2} \sum_{p\le x^{1/2}} \frac{p}{\log p}(1+o(1)) = \frac{x}{(\log x)^2}(1+o(1)).$$

Next, consider $p>x^{1/2}$ and $q<x/p$. By (2.3) we have $p<x^{1/2}(\log x)^3$. From (2.4) it then follows that

$$\sum_{q<x/p} q = \frac{(x/p)^2}{\log x}(1+o(1)).$$

Therefore the contribution to S is (applying (2.6) twice)

$$\frac{x^2}{\log x} \sum_{x^{1/2}<p\le x^{1/2}(\log x)^3} \frac{1}{p^3}(1+o(1)) = \frac{x}{(\log x)^2}(1+o(1)).$$

Therefore the integers $n=pq$ contribute

$$\frac{2x}{(\log x)^2}(1+o(1))$$

to S.

(iv) We now consider integers n with exactly three prime factors, say $n=pqr$ with $q<r<p$ satisfying (2.2) and (2.3), and treat these in the same way as we did in (iii) for two prime factors. First, assume $p\le x^{1/3}$. Given p and q there are less than $\pi(p)$ choices for r. Hence the contribution to S of these integers n is less than

$$\sum_{p\le x^{1/3}} \frac{\pi(p)}{p} \sum_{q<p} q < \sum_{p\le x^{1/3}} \left(\frac{p}{\log p}\right)^2 = O\left(\frac{x}{(\log x)^3}\right).$$

If $p>x^{1/3}$ then by (2.3) $p<x^{1/3}(\log x)^3$. Again there are at most $\pi(p)$ choices for r. We have $q^2<x/p$. From (2.4) we find that the corresponding sum

$$\sum q \text{ is } \le \frac{cx}{p\log x}.$$

The contribution to S is at most

$$\frac{cx}{\log x}\cdot \sum_{x^{1/3}<p\le x^{1/3}(\log x)^3} \frac{1}{p\log p} \le \frac{c'x}{(\log x)^2}\cdot \sum_{x^{1/3}<p\le x^{1/3}\log x} \frac{1}{p}$$
$$= O\left(\frac{x\log\log x}{(\log x)^3}\right) = o\left(\frac{x}{(\log x)^2}\right),$$

by (2.7).

(v) It remains to treat integers with at least four prime factors. The proof is complete if we can show that for these integers the value of

$$\sum n^{-1} \text{ is } o((\log x)^{-2}).$$

Now

$$\sum n^{-1} = \sum_1 q^{-1} \sum_2 t^{-1},$$

where in $\sum_1$ we have, by (2.2) and (2.3),

$$x^{1/2 \log \log x} < q < x$$

and in $\sum_2$ every t has at least three prime factors, and all its prime factors in the interval $(q, q(\log x)^3)$. Hence

$$\sum_2 t^{-1} < \frac{1}{3!} \Big(\sum_{q<p<q(\log x)^3} p^{-1} \Big)^3 + \frac{1}{4!} \Big(\sum_{q<p<q(\log x)^3} p^{-1} \Big)^4 + \ldots$$
$$< c \, (\log \log x)^6 / (\log x)^3,$$

since

$$\sum_{q<p<q(\log x)^3} p^{-1} = O((\log \log x)^2 / \log x)$$

by (2.7).

Since $\sum_1 q^{-1} = O(\log \log x)$ by (2.7), the proof is complete.

It is possible to generalize the results of this paper in the following way. Let $p_1(n) < p_2(n) < \ldots < p_k(n)$ be the distinct prime factors of n. So, k depends on n and $p(n) = p_1(n)$, $P(n) = p_k(n)$. In the following, if $k=1$ then $p_{k-1}(n)$ should be read as $p_k(n)$ and if $k=2$ then $p_{k-2}(n)$ should be read as $p_{k-1}(n)$. Then

$$\sum_{n \le x} \frac{p_{k-1}(n)}{p_k(n)} = (1 + c + o(1)) \frac{x}{\log x}, \quad (x \to \infty)$$

where $c > 0$, and

$$\sum_{n \le x} \frac{p_{k-2}(n)}{p_k(n)} = (1 + o(1)) \frac{x}{\log x}, \quad (x \to \infty).$$

The proofs are similar to those given above.

ACKNOWLEDGEMENT

The authors thank A. Ivić and N.G. de Bruijn for valuable comments and suggestions.

REFERENCES

1. Hardy, G.H. and E.M. Wright – The Theory of Numbers, Oxford Univ. Press Oxford (1954).
2. Lune, J. van de – Some sums involving the largest and smallest prime divisor of a natural number, Report ZW 25/74, Mathematical Centre, Amsterdam (1974).
3. Prachar, K. – Primzahlverteilung, Springer Verlag (1957).
4. Rankin, R.A. – The difference between consecutive prime numbers, J. London Math. Soc. **13**, 242–247 (1938).