

THE LOGARITHM OF THE LEAST PRIME FACTOR II

RAFAEL JAKIMCZUK¹ *

ABSTRACT. Let $p(n)$ be the least prime factor of the positive integer n . In a previous note we prove the equation $\sum_{i=2}^n \log p(i) = o(n \log n)$. In this note we obtain more precise results.

1. PRELIMINARY RESULTS

In this note p_n denotes the n -th prime number and p denotes a positive prime. Then $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, p_5 = 11, \dots$

Let $P(n)$ be the greatest prime factor of the positive integer n . N. G. de Bruijn [1] proved the following asymptotic formula

$$\sum_{i=2}^n \log P(i) = an \log n + O(n),$$

where a is a positive constant defined in [1].

Let $p(n)$ be the least prime factor of the positive integer n . Clearly we have

$$\sum_{i=2}^n \log p(i) \leq \sum_{i=2}^n \log P(i).$$

In a previous note [3] we prove the following theorem

Theorem 1.1. *We have*

$$\sum_{i=2}^n \log p(i) = ng(n),$$

where $g(n) \rightarrow \infty$ and $g(n) = o(\log n)$. Therefore

$$\sum_{i=2}^n \log p(i) = o(n \log n).$$

Proof. See [3]. □

In this note we obtain more precise results.

We need the following theorems.

Date: Received: Nov 10, 2013; Accepted: Apr 1, 2014.

* Corresponding author.

2010 *Mathematics Subject Classification.* Primary 11A99; Secondary 11B99.

Key words and phrases. Least prime factor of a positive integer, logarithm.

The following theorem is sometimes called either the principle of inclusion-exclusion or the principle of cross-classification. We now enunciate the principle.

Theorem 1.2. *Let S be a set of N distinct elements, and let $S_1, \dots, S_r$ be arbitrary subsets of S containing $N_1, \dots, N_r$ elements, respectively. For $1 \leq i < j < \dots < l \leq r$, let $S_{ij\dots l}$ be the intersection of $S_i, S_j, \dots, S_l$ and let $N_{ij\dots l}$ be the number of elements of $S_{ij\dots l}$. Then the number K of elements of S not in any of $S_1, \dots, S_r$ is*

$$K = N - \sum_{1 \leq i \leq r} N_i + \sum_{1 \leq i < j \leq r} N_{ij} - \sum_{1 \leq i < j < k \leq r} N_{ijk} + \dots + (-1)^r N_{12\dots r}.$$

Proof. See, for example, either [2, page 233] or [4, page 84]. $\square$

Theorem 1.3. *Let $\sum_{i=1}^{\infty} a_i$ and $\sum_{i=1}^{\infty} b_i$ be two series of positive terms such that $\lim_{i \rightarrow \infty} \frac{a_i}{b_i} = 1$. If $\sum_{i=1}^{\infty} b_i$ is divergent then we have the following limit*

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i} = 1.$$

Therefore $\sum_{i=1}^{\infty} a_i$ is also divergent.

Proof. See [5, page 332]. $\square$

Theorem 1.4. (Mertens) *The following asymptotic formula holds*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log x},$$

where γ is Euler's constant.

Proof. See [2, chapter XXII]. $\square$

Theorem 1.5. *The following asymptotic formula holds*

$$\sum_{p \leq x} \frac{1}{p} \sim \log \log x.$$

Proof. See [2, chapter XXII]. $\square$

Theorem 1.6. (Prime Number Theorem) *Let $\pi(x)$ be the prime counting function. We have*

$$\pi(x) \sim \frac{x}{\log x}, \quad p_n \sim n \log n, \quad \sum_{p \leq x} \log p \sim x.$$

We also need the following well-known formulae,

$$(a + b)^k = \sum_{i=0}^k \binom{k}{i} a^{k-i} b^i, \tag{1.1}$$

$$2^0 + 2^1 + 2^2 + \dots + 2^n = 2^{n+1} - 1. \tag{1.2}$$

In this note (as usual) $\lfloor \cdot \rfloor$ denotes the integer-part function. Note that

$$0 \leq x - \lfloor x \rfloor < 1. \tag{1.3}$$

2. MAIN RESULTS

Let $\beta_{p_h}(n)$ be the number of numbers that do not exceed n such that their least prime factor is p_h . Therefore

$$\sum_{p_h \leq n} \beta_{p_h}(n) = n - 1. \quad (2.1)$$

We have the following theorem.

Theorem 2.1. *The following asymptotic formula holds*

$$\beta_{p_h}(n) = \left(\prod_{i=1}^{h-1} \left(1 - \frac{1}{p_i} \right) \right) \frac{1}{p_h} n + O_{p_h}(1) \quad (n \geq 1),$$

where $|O_{p_h}(1)| \leq 2^{h-1}$ ($n \geq 1$). Consequently these numbers have positive density

$$D_{p_h} = \left(\prod_{i=1}^{h-1} \left(1 - \frac{1}{p_i} \right) \right) \frac{1}{p_h}.$$

Proof. We have (see Theorem 1.2)

$$\begin{aligned} \beta_{p_h}(n) &= \left\lfloor \frac{n}{p_h} \right\rfloor - \sum_{1 \leq i \leq h-1} \left\lfloor \frac{n}{p_h p_i} \right\rfloor + \sum_{1 \leq i < j \leq h-1} \left\lfloor \frac{n}{p_h p_i p_j} \right\rfloor - \sum_{1 \leq i < j < k \leq h-1} \left\lfloor \frac{n}{p_h p_i p_j p_k} \right\rfloor \\ &+ \cdots + (-1)^{h-1} \left\lfloor \frac{n}{p_h p_1 \cdots p_{h-1}} \right\rfloor = \frac{n}{p_h} - \sum_{1 \leq i \leq h-1} \frac{n}{p_h p_i} + \sum_{1 \leq i < j \leq h-1} \frac{n}{p_h p_i p_j} \\ &- \sum_{1 \leq i < j < k \leq h-1} \frac{n}{p_h p_i p_j p_k} + \cdots + (-1)^{h-1} \frac{n}{p_h p_1 \cdots p_{h-1}} + O_{p_h}(1) \\ &= \frac{n}{p_h} \left(\prod_{i=1}^{h-1} \left(1 - \frac{1}{p_i} \right) \right) + O_{p_h}(1), \end{aligned} \quad (2.2)$$

where (see (2.2), (1.3) and (1.1))

$$|O_{p_h}(1)| \leq \sum_{i=0}^{h-1} \binom{h-1}{i} = (1+1)^{h-1} = 2^{h-1}.$$

□

Theorem 2.2. *The following asymptotic formula holds*

$$\sum_{p_h \leq \log n} \log p_h \beta_{p_h}(n) \sim e^{-\gamma} n \log \log \log n. \quad (2.3)$$

Proof. We have (Theorem 2.1, Theorem 1.6 and Theorem 1.4)

$$D_{p_h} \sim \frac{e^{-\gamma}}{p_h \log p_h}.$$

That is

$$\log p_h D_{p_h} \sim \frac{e^{-\gamma}}{p_h}.$$

Hence (Theorem 1.3)

$$\sum_{h=1}^k \log p_h D_{p_h} \sim \sum_{h=1}^k \frac{e^{-\gamma}}{p_h},$$

and therefore (Theorem 1.5)

$$\sum_{p_h \leq n} \log p_h D_{p_h} \sim \sum_{p_h \leq n} \frac{e^{-\gamma}}{p_h} \sim e^{-\gamma} \log \log n. \quad (2.4)$$

We have (Theorem 2.1)

$$\begin{aligned} \sum_{p_h \leq \log n} \log p_h \beta_{p_h}(n) &= \sum_{p_h \leq \log n} (\log p_h D_{p_h} n + \log p_h O_{p_h}(1)) \\ &= \left(\sum_{p_h \leq \log n} \log p_h D_{p_h} \right) n + \sum_{p_h \leq \log n} \log p_h O_{p_h}(1). \end{aligned} \quad (2.5)$$

Now (see (2.4))

$$\left(\sum_{p_h \leq \log n} \log p_h D_{p_h} \right) n \sim e^{-\gamma} n \log \log \log n, \quad (2.6)$$

and, on the other hand, we have (Theorem 2.1, Theorem 1.6 and (1.2))

$$\begin{aligned} \left| \sum_{p_h \leq \log n} \log p_h O_{p_h}(1) \right| &\leq \sum_{p_h \leq \log n} \log p_h |O_{p_h}(1)| \leq \log \log n \sum_{p_h \leq \log n} 2^{h-1} \\ &= \log \log n (2^{1-1} + 2^{2-1} + \dots + 2^{\pi(\log n)-1}) = \log \log n (2^{\pi(\log n)} - 1) \\ &\leq \log \log n 2^{\pi(\log n)} \leq \log \log n e^{c \log 2 \frac{\log n}{\log \log n}} = n^{\frac{c \log 2}{\log \log n}} \log \log n \\ &\leq \sqrt{n} \log \log n, \end{aligned} \quad (2.7)$$

where $c > 1$. Equations (2.5), (2.6) and (2.7) give (2.3). $\square$

Note that the number of numbers not exceeding n such that their least prime factor do not exceed $\log n$ considered in Theorem 2.2, namely

$$\sum_{p_h \leq \log n} \beta_{p_h}(n)$$

is not a small number, since we have (see [3] and equation (2.1))

$$\sum_{p_h \leq \log n} \beta_{p_h}(n) \sim n.$$

In the following theorem we obtain a more precise result.

Theorem 2.3. *The following asymptotic formulae hold*

$$\sum_{p_h \leq \log n} \beta_{p_h}(n) = n - e^{-\gamma} \frac{n}{\log \log n} + o\left(\frac{n}{\log \log n}\right), \quad (2.8)$$

$$\sum_{\log n < p_h \leq n} \beta_{p_h}(n) \sim e^{-\gamma} \frac{n}{\log \log n}. \quad (2.9)$$

Proof. The following equality can be proved without difficulty by mathematical induction (see Theorem 2.1)

$$\sum_{h=1}^k D_{p_h} = \sum_{h=1}^k \left(\prod_{i=1}^{h-1} \left(1 - \frac{1}{p_i} \right) \right) \frac{1}{p_h} = 1 - \prod_{i=1}^k \left(1 - \frac{1}{p_i} \right) \quad (k \geq 1).$$

Therefore (Theorem 1.4)

$$\sum_{p_h \leq n} D_{p_h} = 1 - \prod_{p_h \leq n} \left(1 - \frac{1}{p_h} \right) = 1 - \frac{e^{-\gamma}}{\log n} + o\left(\frac{1}{\log n}\right). \quad (2.10)$$

We have (Theorem 2.1)

$$\sum_{p_h \leq \log n} \beta_{p_h}(n) = n \sum_{p_h \leq \log n} D_{p_h} + \sum_{p_h \leq \log n} O_{p_h}(1). \quad (2.11)$$

Now (Theorem 2.1, Theorem 1.6 and (1.2))

$$\begin{aligned} & \left| \sum_{p_h \leq \log n} O_{p_h}(1) \right| \leq \sum_{p_h \leq \log n} |O_{p_h}(1)| \leq \sum_{p_h \leq \log n} 2^{h-1} \\ &= 2^{1-1} + 2^{2-1} + \dots + 2^{\pi(\log n)-1} = 2^{\pi(\log n)} - 1 \\ &\leq 2^{\pi(\log n)} \leq e^{c \log 2 \frac{\log n}{\log \log n}} = n^{\frac{c \log 2}{\log \log n}} \\ &\leq \sqrt{n}. \end{aligned} \quad (2.12)$$

Equations (2.11), (2.10) and (2.12) give (2.8). Finally (2.9) is an immediate consequence of (2.8). $\square$

Now, we can obtain a better theorem than Theorem 2.2.

Theorem 2.4. *Let k be a positive integer. The following asymptotic formula holds*

$$\sum_{p_h \leq \log^k n} \log p_h \beta_{p_h}(n) \sim e^{-\gamma} n \log \log \log n. \quad (2.13)$$

Proof. We have

$$\sum_{p_h \leq \log^k n} \log p_h \beta_{p_h}(n) = \sum_{p_h \leq \log n} \log p_h \beta_{p_h}(n) + \sum_{\log n < p_h \leq \log^k n} \log p_h \beta_{p_h}(n). \quad (2.14)$$

On the other hand (see (2.9)) we have

$$\begin{aligned} & \sum_{\log n < p_h \leq \log^k n} \log p_h \beta_{p_h}(n) \leq \log(\log^k n) \sum_{\log n < p_h \leq n} \beta_{p_h}(n) \\ &\sim k \log \log n e^{-\gamma} \frac{n}{\log \log n} \sim k e^{-\gamma} n. \end{aligned} \quad (2.15)$$

Equations (2.14), (2.3) and (2.15) give (2.13). $\square$

Now, we can prove a better theorem than Theorem 1.1.

Theorem 2.5. *The following inequality holds*

$$\sum_{i=2}^n \log p(i) = \sum_{p_h \leq n} \log p_h \beta_{p_h}(n) \leq c \frac{1}{\log \log n} n \log n, \quad (2.16)$$

where c is a positive constant.

Proof. We have

$$\sum_{p_h \leq n} \log p_h \beta_{p_h}(n) = \sum_{p_h \leq \log n} \log p_h \beta_{p_h}(n) + \sum_{\log n < p_h \leq n} \log p_h \beta_{p_h}(n). \quad (2.17)$$

On the other hand (see (2.9)) we have

$$\sum_{\log n < p_h \leq n} \log p_h \beta_{p_h}(n) \leq \log n \sum_{\log n < p_h \leq n} \beta_{p_h}(n) \sim \log n e^{-\gamma} \frac{n}{\log \log n}. \quad (2.18)$$

Equations (2.18), (2.3) and (2.17) give (2.16). $\square$

We establish the following conjecture.

$$\sum_{i=2}^n \log p(i) \sim e^{-\gamma} n \log \log n.$$

To finish, we have the following theorem.

Theorem 2.6. *Let $k \geq 2$ be an arbitrary but fixed positive integer. We have the following asymptotic formulae*

$$\sum_{\frac{n}{k} < p_h \leq n} \log p_h \beta_{p_h}(n) \sim \left(1 - \frac{1}{k}\right) n,$$

$$\sum_{\frac{n}{k} < p_h \leq n} \beta_{p_h}(n) \sim \left(1 - \frac{1}{k}\right) \frac{n}{\log n},$$

$$\sum_{\sqrt{n} < p_h \leq n} \log p_h \beta_{p_h}(n) \sim n,$$

$$\sum_{\sqrt{n} < p_h \leq n} \beta_{p_h}(n) \sim \frac{n}{\log n}.$$

Proof. Note that if $\sqrt{n} < p_h \leq n$ we have $p_h^2 > n$. Hence there exists a unique multiple of p_h not exceeding n such that p_h is its least prime factor, namely p_h . Therefore in the four formulae of the theorem we have

$$\beta_{p_h}(n) = 1. \quad (2.19)$$

Now, the theorem is an immediate consequence of (2.19) and Theorem 1.6. $\square$

Acknowledgement. The author is very grateful to Universidad Nacional de Luján.

REFERENCES

1. N. G. de Bruijn, *On the number of positive integers $\leq x$ and free of prime factors $> y$* , Kon. Nederl. Akad. Wetensch. Proceedings A 54(= Indagationes Math. 13) (1951), 50 - 60.
2. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford, 1960.
3. R. Jakimczuk, *The logarithm of the least prime factor*, Gulf J. Math. **1**(2013), 1-4.
4. W. J. LeVeque, *Topics in Number Theory*, Addison-Wesley, 1958.
5. J. Rey Pastor, P. Pi Calleja y C. A. Trejo, *Análisis Matemático (Volumen 1)*, Kapelusz, 1969.

¹ DIVISIÓN MATEMÁTICA, UNIVERSIDAD NACIONAL DE LUJÁN, BUENOS AIRES, ARGENTINA.

E-mail address: jakimczu@mail.unlu.edu.ar