

Sum of Prime Factors in the Prime Factorization of an Integer

Rafael Jakimczuk

División Matemática, Universidad Nacional de Luján
Buenos Aires, Argentina
jakimczu@mail.unlu.edu.ar

In memory of my sister Fedra Marina Jakimczuk (1970-2010)

Abstract

In this article we obtain some results on the sequence $c(n)$, where $c(n)$ is the sum of the prime factors in the prime factorization of n .

Mathematics Subject Classification: 11A99, 11B99

Keywords: Prime factorization, sum of prime factors

1 Introduction

In this note we obtain some results on the sequence $c(n)$, where $c(n)$ is the sum of the prime factors in the prime factorization of $n \geq 2$. For example if $n = 12$ then $c(n) = c(12) = c(2.2.3) = 2 + 2 + 3 = 7$ and if n is prime then $c(n) = n$.

The first few terms of the integer sequence $c(n)$ are

2, 3, 4, 5, 5, 7, 6, 6, 7, 11, 7, 13, 9, 8, 8, 17, 8, 19, 9, 10, 13, 23, 9, 10, 15, 9, 11, 29,

10, 31, 10, 14, 19, 12, 10, 37...

Therefore $c(2) = 2$, $c(3) = 3$, $c(4) = 4$, $c(5) = 5$, $c(6) = 5$, ...

We see that the sequence $c(n)$ is very irregular. On the other hand, we see that there are numbers repeated.

2 Main Results

First, we establish two general lemmas.

Lemma 2.1 *If a_i ($i = 1, \dots, n$) are positive integers such that $a_i \geq 2$ ($i = 1, \dots, n$) then we have the following inequality*

$$a_1 + a_2 + \dots + a_n \leq a_1 a_2 \dots a_n \quad (n \geq 1). \quad (1)$$

Proof. We apply mathematical induction. If $n = 1$ we obtain $a_1 \leq a_1$. Consequently the lemma is true for $n = 1$. If $n = 2$ we can suppose that $a_1 \geq a_2$. Therefore

$$a_1 \cdot a_2 \geq a_1 \cdot 2 = a_1 + a_1 \geq a_1 + a_2.$$

That is

$$a_1 \cdot a_2 \geq a_1 + a_2. \quad (2)$$

Consequently the lemma is true for $n = 2$.

Suppose that the lemma is true for $n \geq 2$, that is (inductive hypothesis)

$$a_1 + a_2 + \dots + a_n \leq a_1 a_2 \dots a_n. \quad (3)$$

We shall prove that the lemma is also true for $n + 1$.

Equation (2) and equation (3) give

$$a_1 \dots a_n a_{n+1} = (a_1 \dots a_n) a_{n+1} \geq (a_1 \dots a_n) + a_{n+1} \geq a_1 + \dots + a_n + a_{n+1}.$$

The lemma is proved.

Lemma 2.2 *If a_i ($i = 1, \dots, n$) are positive integers such that $a_i \geq 2$ ($i = 1, \dots, n$) then we have the following inequality*

$$a_1 + a_2 + \dots + a_n \geq \frac{3}{\log 3} \log (a_1 a_2 \dots a_n). \quad (4)$$

Proof. If we study the function $f(x) = \frac{x}{\log x}$ then we obtain the following inequality

$$\frac{k}{\log k} \geq \frac{3}{\log 3},$$

where k is a positive integer such that $k \geq 2$. Consequently we have ($n \geq 1$)

$$\frac{a_i}{\log a_i} \geq \frac{3}{\log 3} \quad (i = 1, \dots, n).$$

That is

$$a_i \geq \frac{3}{\log 3} \log a_i \quad (i = 1, \dots, n).$$

Therefore

$$a_1 + a_2 + \cdots + a_n \geq \frac{3}{\log 3} (\log a_1 + \log a_2 + \cdots + \log a_n) = \frac{3}{\log 3} \log (a_1 a_2 \cdots a_n)$$

The lemma is proved.

We have the following theorems.

Theorem 2.3 *We have the following inequalities*

$$c(n) \leq n \quad (n \geq 2), \quad (5)$$

$$c(n) \geq \frac{3}{\log 3} \log n \quad (n \geq 2). \quad (6)$$

Proof. If we consider the prime factorization of n and apply (1) and (4) then we obtain (5) and (6) respectively. The theorem is proved.

Remark 2.4 *Note that in theorem 2.3 n is the least upper bound since if n is prime then $c(n) = n$. Besides, $\frac{3}{\log 3} \log n$ is the greatest lower bound since if $n = 3^k$ then $c(3^k) = \frac{3}{\log 3} \log 3^k = 3k$.*

Corollary 2.5 *The following limit holds.*

$$\lim_{n \rightarrow \infty} c(n) = \infty.$$

Proof. It is an immediate consequence of equation (6).

Let n be a positive integer greater than 1. $\vartheta(n)$ denotes the number of partitions of n into positive prime numbers.

Lemma 2.6 *If $n = 2, 3, 4$ then $\vartheta(n) = 1$. If $n \geq 5$ then $\vartheta(n) \geq 2$.*

Proof. If $n \geq 6$ is even we have $n = 2 + \cdots + 2 = 2 + \cdots + 2 + 3 + 3$. If $n \geq 9$ is odd we have $n = 2 + \cdots + 2 + 3 = 2 + \cdots + 2 + 3 + 3 + 3$. If $n = 5$ we have $5 = 5$ and $5 = 2 + 3$. If $n = 7$ we have $7 = 7$, $7 = 2 + 5$ and $7 = 2 + 2 + 3$. The lemma is proved.

Theorem 2.7 *The equation $c(i) = n$ where $n \geq 2$ has $\vartheta(n)$ solutions. Consequently if $n = 2, 3, 4$ then the equation has one solution. On the other hand if $n \geq 5$ the equation has at least two solutions.*

Proof. It is an immediate consequence of the definition of $c(n)$ and of lemma 2.6. The theorem is proved.

Theorem 2.8 *The sequence $c(n)$ is not decreasing. The sequence $c(n)$ is not increasing.*

Proof. Clearly the sequence $c(n)$ is not decreasing (see corollary 2.5). On the other hand, the sequence $c(n)$ is not increasing since if $p \geq 7$ is prime we have $p < 2^{\frac{p-1}{2}}$ and $c(p) = p > c\left(2^{\frac{p-1}{2}}\right) = p - 1$. The theorem is proved.

Theorem 2.9 *We have the following asymptotic formula*

$$\sum_{i=2}^n c(i) \sim \frac{\pi^2}{12} \frac{n^2}{\log n}. \quad (7)$$

Proof. Let $S_1(n)$ be the sum of the prime factors in the prime factorization of $n!$. In a previous article [1] we prove the asymptotic formula $S_1(n) \sim \frac{\pi^2}{12} \frac{n^2}{\log n}$. On the other hand, clearly $S_1(n) = \sum_{i=2}^n c(i)$. The theorem is proved.

Theorem 2.10 *Let us consider the set $\{c(2), c(3), \dots, c(n)\}$. Let n_0 be the number of numbers in this set such that $c(i) \geq \frac{i}{\log^{1-\epsilon} i}$ where $0 < \epsilon < 1$ is fixed. We have $\lim_{n \rightarrow \infty} \frac{n_0}{n} = 0$.*

Proof. Note that there exists a positive integer $q > 2$ such that the function $f(x) = \frac{x}{\log^{1-\epsilon} x}$ is strictly increasing on the interval $[q, \infty]$.

Let us consider the set $\{c(q), c(q+1), \dots, c(n)\}$. Let m_0 be the number of numbers in this set such that $c(i) \geq \frac{i}{\log^{1-\epsilon} i}$. Suppose that the limit $\lim_{n \rightarrow \infty} \frac{m_0}{n} = 0$ is not fulfilled. Therefore there exists $\alpha > 0$ such that for infinite values of n we have $\frac{m_0}{n} \geq \alpha$.

Note that (L'Hospital's rule)

$$\lim_{x \rightarrow \infty} \frac{\int_q^x \frac{t}{\log^{1-\epsilon} t} dt}{\frac{x^2}{2 \log^{1-\epsilon} x}} = 1.$$

That is

$$\int_q^x \frac{t}{\log^{1-\epsilon} t} dt = h_1(x) \frac{x^2}{2 \log^{1-\epsilon} x},$$

where $h_1(x) \rightarrow 1$.

Therefore we have

$$\begin{aligned} \sum_{i=q}^n c(i) &\geq \sum_{\substack{c(i) \geq \frac{i}{\log^{1-\epsilon} i}, \\ i \geq q}} c(i) \geq \sum_{i=q}^{q+m_0-1} \frac{i}{\log^{1-\epsilon} i} \geq \int_q^{q+m_0-1} \frac{x}{\log^{1-\epsilon} x} dx \\ &\geq \int_q^{q-1+\alpha n} \frac{x}{\log^{1-\epsilon} x} dx = h_1(q-1+\alpha n) \frac{(q-1+\alpha n)^2}{2 \log^{1-\epsilon} (q-1+\alpha n)} \\ &= h_2(n) \frac{\alpha^2}{2} \frac{n^2}{\log^{1-\epsilon} n} \geq C \frac{n^2}{\log n} \log^\epsilon n, \end{aligned}$$

where $h_2(n) \rightarrow 1$ and $0 < C < \frac{\alpha^2}{2}$. That is

$$\sum_{i=q}^n c(i) \geq C \frac{n^2}{\log n} \log^\epsilon n. \quad (8)$$

Now, equation (8) and equation (7) are an evident contradiction. Therefore $\lim_{n \rightarrow \infty} \frac{m_0}{n} = 0$ and consequently $\lim_{n \rightarrow \infty} \frac{n_0}{n} = 0$. The theorem is proved.

Corollary 2.11 *Let us consider the set $\{c(2), c(3), \dots, c(n)\}$. Let n_1 be the number of numbers in this set such that $c(i) < \frac{i}{\log^{1-\epsilon} i}$ where $0 < \epsilon < 1$ is fixed. We have $\lim_{n \rightarrow \infty} \frac{n_1}{n} = 1$.*

Theorem 2.12 *There exists n_0 such that if $n \geq n_0$ we have the following inequality*

$$\frac{1}{2} \log n < \sum_{i=2}^n \frac{1}{c(i)} < 2 \frac{n}{\log n}. \quad (9)$$

Proof. Theorem 2.3 gives

$$\sum_{i=2}^n \frac{1}{i} \leq \sum_{i=2}^n \frac{1}{c(i)} \leq \sum_{i=2}^n \frac{1}{\log i}. \quad (10)$$

Now, we have

$$\sum_{i=2}^n \frac{1}{i} = \int_2^n \frac{1}{x} dx + O(1) = h(n) \log n > \frac{1}{2} \log n, \quad (11)$$

where $h(n) \rightarrow 1$.

On the other hand, we have

$$\sum_{i=2}^n \frac{1}{\log i} = \int_2^n \frac{1}{\log x} dx + O(1) = g(n) \frac{n}{\log n} < 2 \frac{n}{\log n}, \quad (12)$$

where $g(n) \rightarrow 1$. Since (L'Hospital rule)

$$\lim_{x \rightarrow \infty} \frac{\int_2^x \frac{1}{\log t} dt}{\frac{x}{\log x}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{\log x}}{\frac{\log x - 1}{\log^2 x}} = 1.$$

Finally, (10), (11) and (12) give (9). The theorem is proved.

Corollary 2.13 *The series $\sum_{n=2}^{\infty} \frac{1}{c(n)}$ is divergent and $\sum_{i=2}^n \frac{1}{c(i)} = o(n)$.*

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

References

- [1] R. Jakimczuk, Sums of prime numbers, the zeta function and the π number, *International Mathematical Forum*, **3** (2008), 1383 - 1386.

Received: June, 2012