Proof of Goldbach conjecture

By Toshihiko Ishiwata

Apr. 25, 2025

Abstract. This paper is a trial to prove Goldbach conjecture according to the following process.
1. We find that {the total number of ways to divide an even number n into 2 prime numbers} : $l(n)$ diverges to ∞ with $n \to \infty$.
2. We find that $1 \leq l(n)$ holds true in $4 * 10^{18} < n$ from the probability of $l(n) = 0$.
3. Goldbach conjecture is already confirmed to be true up to $n = 4 * 10^{18}$.
4. Goldbach conjecture is true from the above item 2 and 3.

1. Introduction

1.1 When an even number n is divided into 2 odd numbers x and y, we can express the situation as pair (x, y) like the following (1).

$$n = x + y = (x, y) \qquad (n = 6, 8, 10, 12, \cdots\cdots \quad x, y : \text{odd number}) \tag{1}$$

n has $n/2$ pairs like the following (2).

$$(1, n-1), (3, n-3), (5, n-5), \cdots\cdots, (n-5, 5), (n-3, 3), (n-1, 1) \tag{2}$$

We define as follows.

Prime pair : the pair where both x and y in (x, y) are prime numbers
Composite pair : the pair other than the above prime pair
$l(n)$: the total number of the prime pairs which exist in $n/2$ pairs shown by the above (2). (p, q) is regarded as the different pair from (q, p). $(p, q$: prime number$)$

1.2 Goldbach conjecture can be expressed as the following (3) i.e. any even number $n(\geq 6)$ can be divided into 2 prime numbers.

$$1 \leq l(n) \qquad (n = 6, 8, 10, 12, \cdots\cdots) \tag{3}$$

Since Goldbach conjecture is already confirmed to be true up to $n = 4 * 10^{18}$, we can try to prove Goldbach conjecture in the following condition.

$$4 * 10^{18} < n \tag{4}$$

2020 *Mathematics Subject Classification.* Primary 11P32.
Key Words and Phrases. Goldbach conjecture.

2. Investigation of $l(n)$

2.1 When an even number n is divided into 2 odd numbers x and y, we can find the pair of $\pi(n), l(n), m_{xx}, m_x, m_y$ and m_{xy} in $n/2$ pairs of (x, y) as shown in the following (Figure 1).

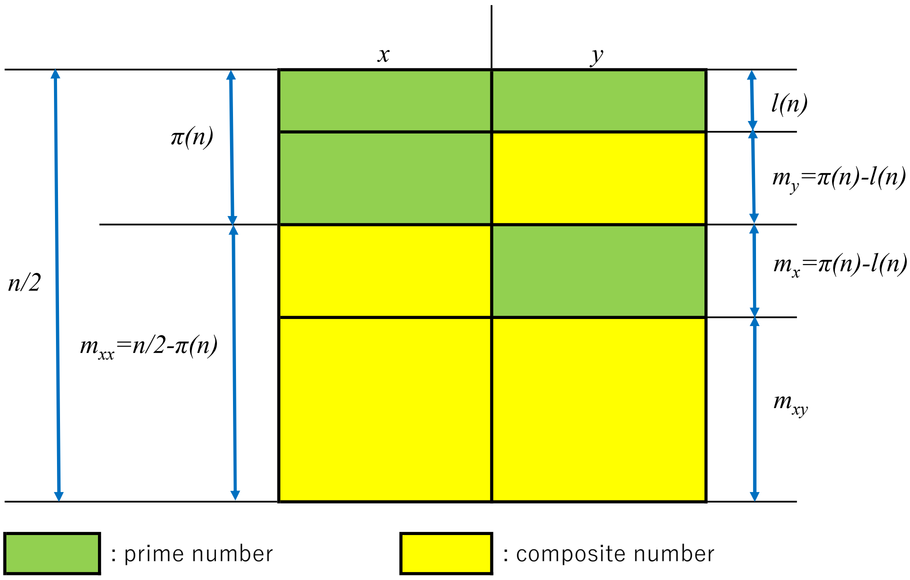


Figure 1 : Various pairs in $n/2$ pairs of (x, y)

We define as follows.

$\pi(n)$: $\pi(n)$ shows the total number of prime numbers which exist between 1 and n. But we use $\pi(n)$ in the above (Figure 1) for the total number of prime numbers which exist in $n/2$ odd numbers of $(1, 3, 5, \cdots\cdots,$ $n-5, n-3, n-1)$. Strictly speaking, this value must be $\pi(n-1)-1$. But we can say $\quad \pi(n-1)-1 = \pi(n)-1 \fallingdotseq \pi(n)$ because n is an even number and a large number as shown in (4).

m_{xx} : the total number of pairs where x is a composite number. 1 is regarded as a composite number.

m_x : the total number of pairs where x and y are composite number and prime number respectively

2.2 We have the following (5) from Prime number theorem.

$$\frac{\pi(n)}{n} \sim \frac{n/\log n}{n} = \frac{1}{\log n} \qquad (n \to \infty) \tag{5}$$

We have $\lim_{n\to\infty} \frac{\pi(n)}{n} = 0$ from the above (5). Then we have the following (6) from (Figure 1) and $\lim_{n\to\infty} \frac{\pi(n)}{n} = 0$

$$m_{xx} = n/2 - \pi(n) = (n/2)\{1 - 2\pi(n)/n\} \sim n/2 \qquad (n \to \infty) \tag{6}$$

When m_{xx} approaches $n/2$ with $n \to \infty$ as shown in the above (6), m_x approaches $\pi(n)$ with $n \to \infty$ due to the following reasons.

2.2.1 m_x shows the total number of prime numbers which exist in y of m_{xx} as shown in (Figure 1).

2.2.2 $n/2$ pieces of y, $(1, 3, 5, \cdots\cdots, n-5, n-3, n-1)$ have $\pi(n)$ prime numbers.

Then we can have the following (7) from (Figure 1).

$$m_x = \pi(n) - l(n) = \pi(n)\{1 - l(n)/\pi(n)\} \quad \sim \quad \pi(n) \qquad (n \to \infty) \qquad (7)$$

We have $\lim_{n\to\infty} \frac{l(n)}{\pi(n)} = 0$ from the above (7). We have the following (8) from the above (6) and (7).

$$\frac{\pi(n) - l(n)}{n/2 - \pi(n)} \quad \sim \quad \frac{\pi(n)}{n/2} \qquad (n \to \infty) \qquad (8)$$

We have the following (9) from the above (8) and Prime number theorem.

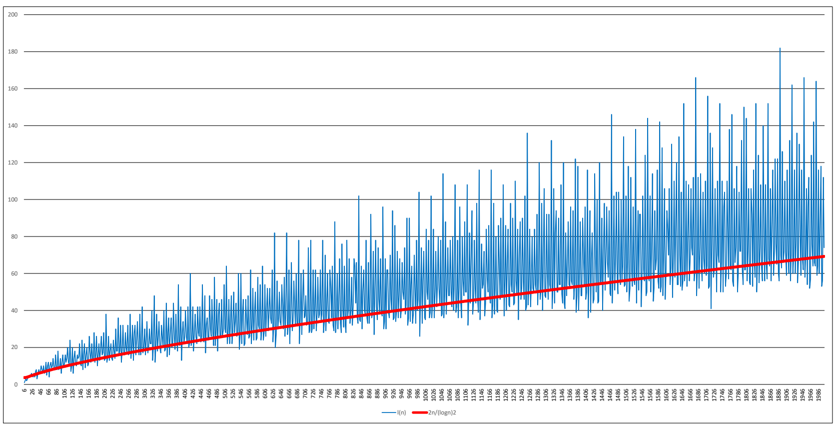
$$l(n) \quad \sim \quad \frac{\{\pi(n)\}^2}{n/2} \quad \sim \quad \frac{\{n/\log n\}^2}{n/2} = \frac{2n}{(\log n)^2} \qquad (n \to \infty) \qquad (9)$$

We can find that $l(n)$ has the following property from the above (9).

2.2.3 $l(n)$ repeats increases and decreases with increase of n as shown in the following (Graph 1). But overall $l(n)$ is an increasing function regarding n because $\frac{2n}{(\log n)^2}$ is an increasing function regarding n.

2.2.4 $l(n)$ diverges to ∞ with $n \to \infty$ because $\frac{2n}{(\log n)^2}$ diverges to ∞ with $n \to \infty$.

2.3 $\frac{2n}{(\log n)^2}$ seems to approximate $l(n)$ sufficiently well as shown in the following (Graph 1).



Graph 1 : $l(n)$(blue line)[1] and $\frac{2n}{(\log n)^2}$(red line) from $n = 6$ to $n = 2,000$

3. Investigation of zero point of $l(n)$

3.1 We can consider the probability that k or $(n-k)$ in pair $(x,y) = (k, n-k)$ is a prime number as follows.

$$(k = 3, 5, 7, 9, \ldots\ldots, n/2-4, n/2-2, n/2 \qquad n/2 : \text{odd number})$$
$$(k = 3, 5, 7, 9, \ldots\ldots, n/2-5, n/2-3, n/2-1 \qquad n/2 : \text{even number})$$

3.1.1 Since both k and $(n-k)$ in $(k, n-k)$ are always an odd number, we must consider the probability that k or $(n-k)$ is a prime number in the world where only odd numbers exist.

3.1.2 Prime number theorem shows that {the probability that randomly selected integer m is a prime number} approaches to $1/\log m$ with $m \to \infty$. Then we can have {the probability that randomly selected odd number N is a prime number} : $P(N)$ like the following (10) because an even mumber cannot be a prime number.

$$P(N) \quad \sim \quad \frac{2}{\log N} \qquad (N \to \infty \quad N : \text{odd number}) \tag{10}$$

3.1.3 {The average probability that odd numbers between 1 and N is a prime number} : $p(N)$ can be expressed like the following (11) from Prime number theorem.

$$\begin{aligned} p(N) &= \frac{(\text{The total number of prime numbers between 3 and } N)}{(\text{The total number of odd numbers between 1 and } N)} \\ &= \frac{\pi(N)-1}{(N+1)/2} \quad \sim \quad \frac{2*\pi(N)}{N} \quad \sim \quad \frac{2*N/\log N}{N} = \frac{2}{\log N} \\ & \qquad (N \to \infty \quad N : \text{odd number}) \end{aligned} \tag{11}$$

Since $P(N)$ decreases with increase of N as shown in [Appendix 1 : Investigation of $P(N)$], we have the following (12).

$$P(N) < p(N) \tag{12}$$

3.2 Since the probability that $(k, n-k)$ or $(n-k, k)$ is a prime pair is $P(k)*P(n-k)$, the probability that $(k, n-k)$ or $(n-k, k)$ is a composite pair is $\{1 - P(k) * P(n-k)\}$. Therefore the probability that all of $n/2$ pairs are a composite pair i.e. {the probability of $l(n) = 0$} : $A(n)$ can be expressed like the following (13).
Since $(1, n-1)$ and $(n-1, 1)$ are always a composite pair, we don't include these pairs in (13). Then k does not include 1 and (13) has $(n/2-2)$ terms of $\{1 - P(k) * P(n-k)\}$ altogether.

$$\begin{aligned} A(n) &= \{1-P(3)*P(n-3)\}^2\{1-P(5)*P(n-5)\}^2\{1-P(7)*P(n-7)\}^2\ldots\ldots \\ &\quad \{1-P(k)*P(n-k)\}^2\ldots\ldots\{1-P(n/2+4)*P(n/2-4)\}^2 \\ &\quad \{1-P(n/2+2)*P(n/2-2)\}^2\{1-P(n/2)^2\} \qquad (n/2 : \text{odd number}) \\ &= \{1-P(3)*P(n-3)\}^2\{1-P(5)*P(n-5)\}^2\{1-P(7)*P(n-7)\}^2\ldots\ldots \end{aligned}$$

$$\{1-P(k)*P(n-k)\}^2 \ldots\ldots \{1-P(n/2+5)*P(n/2-5)\}^2$$
$$\{1-P(n/2+3)*P(n/2-3)\}^2\{1-P(n/2+1)*P(n/2-1)\}^2$$
$$(n/2 : \text{even number}) \quad (13)$$

$$(k=3,5,7,9,\ldots\ldots,n/2-4,n/2-2,n/2 \qquad n/2 : \text{odd number}) \quad (13\text{-}1)$$
$$(k=3,5,7,9,\ldots\ldots,n/2-5,n/2-3,n/2-1 \qquad n/2 : \text{even number}) \quad (13\text{-}2)$$

3.3 We have the following (14) from (13-1) and (13-2).

$$3 \le k \le n/2 \le n-k < n+1 \ll 10^{18}*n+1 \quad (14)$$

Since $P(N)$ decreases with increase of N as shown in [Appendix 1], if n is large enough, we have the following (15) from (10) and (14).

$$1 > P(k) \ge P(n-k) > P(n+1) \fallingdotseq \frac{2}{\log n}$$
$$> P(10^{18}*n+1) \fallingdotseq \frac{2}{\log(10^{18}*n)} = \frac{2}{\log n + 41.4} \quad (15)$$

We have the following (16) from (15).

$$0 < 1-P(k)*P(n-k) < 1-\{P(10^{18}*n+1)\}^2 \quad (16)$$

We have the following (17) from (13), (15) and (16).

$$0 < A(n) < B(n) = [1-\{P(10^{18}*n+1)\}^2]^{n/2-2}$$
$$\sim \quad \{1-\frac{4}{(\log n+41.4)^2}\}^{n/2}$$
$$= [\{1-\frac{1}{\{(\log n+41.4)/2\}^2}\}^{\{(\log n+41.4)/2\}^2}]^{(n/2)/\{(\log n+41.4)/2\}^2}$$
$$\sim \quad (\frac{1}{e})^{(n/2)/\{(\log n+41.4)/2\}^2} = \frac{1}{e^{(n/2)/\{(\log n+41.4)/2\}^2}} \qquad (n\to\infty) \quad (17)$$

We have the following (18) from the above (17).

$$\lim_{n\to\infty} A(n) = 0 \quad (18)$$

If n is large enough, i.e. if $4*10^{18} \le n$ is satisfied, $B(n)$ can be approximated to $\frac{1}{e^{(n/2)/\{(\log n+41.4)/2\}^2}}$ from the above (17) and $\frac{1}{e^{(n/2)/\{(\log n+41.4)/2\}^2}}$ decreases with increase of n in $4*10^{18} \le n$. Therefore we have the following (19).

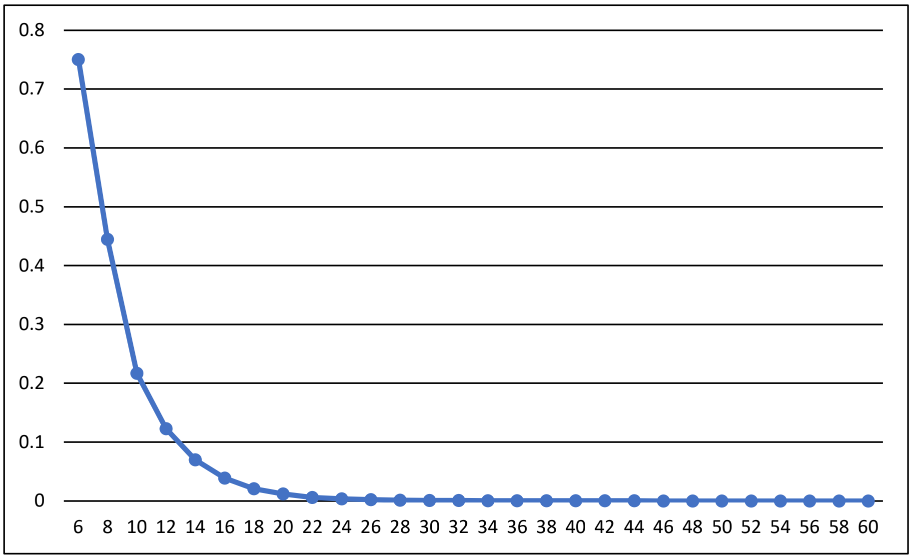
$$0 < A(n) < B(n) < B(4*10^{18}) \qquad (4*10^{18} < n) \quad (19)$$

3.4 Here we make another {the probability of $l(n)=0$} : $a(n)$ by substituting $p(k)$ and $p(n-k)$ for $P(k)$ and $P(n-k)$ in (13) respectively. Because when N is small calculating $p(N)$ is easier than calculating $P(N)$ as shown in item 3.1.2 and 3.1.3.

We have the following (20) from (12) and (13).

$$a(n) < A(n) \tag{20}$$

3.5 The following (Graph 2) shows that $a(n)$ decreases with increase of n in $n \leq 60$. $A(n)$ exists above $a(n)$ in (Graph 2) from (20).



Graph 2 : $a(n)$ from $n = 6$ to $n = 60$

n	6	8	10	12	14	16	18	20	30	60
$a(n)$	0.75	0.4444	0.217	0.1225	0.07	0.0386	0.0207	0.0117	0.0008	3E-06

Table 1 : the values of $a(n)$

3.6 $A(n)$ and $a(n)$ have the following property from the above item 3.4 and 3.5.

3.6.1 $a(n)$ decreases with increase of n at least in $n \leq 60$.

3.6.2 The above (19) holds true.

3.6.3 $A(n)$ converges to zero with $n \to \infty$.

3.7 When $l(n_0) = 0$ holds true we define n_0 as {zero point of $l(n)$}. We defined $A(n)$ as {the probability of $l(n) = 0$} in item 3.2. But we can also call $A(n)$ {the probability of zero point occurrence of $l(n)$}.
Possible zero point distribution of $l(n)$ is limited to 4 cases which are classified according to location of zero point as shown in the following (Table 2).

	Location of zero point		Contradiction with	Can this case exist as real $l(n)$?
	$n \leqq 4*10^{18}$	$4*10^{18}<n$		
Case 1	●	●	item 3.7.2	NO
Case 2	●	X	item 3.7.2	NO
Case 3	X	●	item 3.7.1	NO
Case 4	X	X	nothing	YES

● : zero points exist. X : no zero points exist.

Table 2 : 4 cases of zero point distribution of $l(n)$

Distribution of zero point of $l(n)$ is affected by the following facts.

3.7.1 $A(n)$ and $a(n)$ have the property shown in item 3.6.

3.7.2 Goldbach conjecture is already confirmed to be true up to $n = 4*10^{18}$ as shown in item 1.2. Therefore a zero point of $l(n)$ does not exist in $n \leq 4*10^{18}$.

Case 1 and Case 2 cannot exist because they contradict item 3.7.2.
Case 3 cannot exist because it contradicts item 3.7.1 as shown in the following item 3.8.

3.8 From (19) we have the following (21) which shows that $A(n)$ is extremely small in $4*10^{18} < n$. $B(n)$ is defined in (17).

$$A(n) < B(4*10^{18}) \fallingdotseq \frac{1}{e^{(2*10^{18})/[\{\log(4*10^{18})+41.4\}/2]^2}} = \frac{1}{e^{(2*10^{18})/1774}} = e^{-1.1*10^{15}}$$
$$= (e^{1.1})^{-10^{15}} = (10^{0.47})^{-10^{15}} = 10^{-4.7*10^{14}} \qquad (4*10^{18} < n) \qquad (21)$$

We can calculate the probability of zero point occurrence of $l(n)$ near $n = 6$ from (11), (13) and (20) as follows.

$$A(6) > a(6) = 1 - \{p(3)\}^2 = 1 - \{\frac{\pi(3)-1}{(3+1)/2}\}^2 = 1 - (1/2)^2 = 0.75 \qquad (22)$$

Since Case 3 has zero points only in $4*10^{18} < n$, Case 3 contradicts $A(n)$ as follows.

3.8.1 The situation where a zero point can exist in $A(n) < 10^{-4.7*10^{14}}$ as (21) shows contradicts the situation where a zero point cannot exist in $A(n) > 0.75$ as (22) shows. Because the larger $A(n)$ is, the more likely a zero point will appear. In other words, Case 3 shows the situation that is completely opposite to the situation expected from $A(n)$ as shown in the following item 3.8.2 and 3.8.3.

3.8.2 0.75 is extremely larger than $10^{-4.7*10^{14}}$ and zero points already exist in $A(n) < 10^{-4.7*10^{14}}$. Therefore a new zero point must exist near $n = 6$. But Case 3 does not have any zero point in $n \leq 4*10^{18}$.

3.8.3 $10^{-4.7*10^{14}}$ is extremely smaller than 0.75 and zero points do not exist near $n = 6$. Therefore zero points must not exist in $4*10^{18} < n$. But Case 3 has zero points in $4*10^{18} < n$.

By the way Case 2 and Case 4 are consistent with $A(n)$. The following (Figure 2) shows the contradiction between Case 3 and $A(n)$.

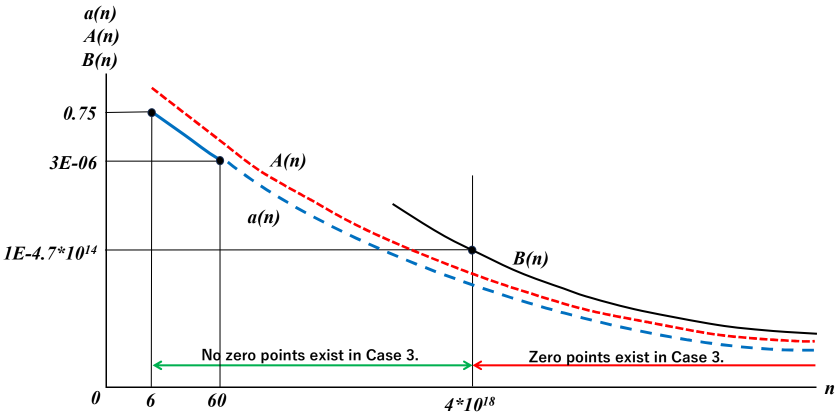


Figure 2 : the contradiction between Case 3 and $A(n)$

3.9 Among 4 cases of zero point distribution of $l(n)$ shown in (Table 2), only Case 4 is consistent with both item 3.7.1 and 3.7.2. Therefore Case 4 is the real $l(n)$. We have the following (23) from Case 4 because Case 4 does not have any zero point in $4 * 10^{18} < n$.

$$1 \leq l(n) \qquad\qquad (4 * 10^{18} < n) \tag{23}$$

4. Conclusion

Goldbach conjecture is true from the following item 4.1 and 4.2.

4.1 Goldbach conjecture is already confirmed to be true up to $n = 4 * 10^{18}$.

4.2 Goldbach conjecture is true in $4 * 10^{18} < n$ from the above (23).

Appendix 1. : Investigation of $P(N)$

1.1 When odd number N is a composite number, N is divisible by any of $\{\pi(\lfloor\sqrt{N}\rfloor)-1\}$ prime numbers of $\{p_2 = 3, p_3 = 5, p_4 = 7, \cdots\cdots, p_k, \cdots\cdots, p_{\pi(\lfloor\sqrt{N}\rfloor)-1}, p_{\pi(\lfloor\sqrt{N}\rfloor)}\}$. The above $\{\pi(\lfloor\sqrt{N}\rfloor)-1\}$ prime numbers satisfy $3 \leq p \leq \sqrt{N}$. (p : prime number) Then {the probability that N is a composite number} i.e. {the probability that N is divisible by any of the above $\{\pi(\lfloor\sqrt{N}\rfloor)-1\}$ prime numbers} : $Q(N)$ can be expressed like the following (24).

$$Q(N) = Q_2 + Q_3 + Q_4 + \cdots\cdots + Q_k + \cdots\cdots + Q_{\pi(\lfloor\sqrt{N}\rfloor)-1} + Q_{\pi(\lfloor\sqrt{N}\rfloor)} \qquad (2 \leq k \leq \pi(\lfloor\sqrt{N}\rfloor)) \tag{24}$$

Q_2 : the probability that N is divisible by $p_2 = 3$
Q_3 : the probability that N is divisible by $p_3 = 5$ but not by $p_2 = 3$
Q_4 : the probability that N is divisible by $p_4 = 7$ but not by $p_3 = 5$ or $p_2 = 3$
Q_k : the probability that N is divisible by p_k but not by any of ($p_{k-1}, p_{k-2}, \cdots\cdots, p_4 = 7, p_3 = 5$ or $p_2 = 3$)

1.2 We have the values of Q_2, Q_3 and Q_4 as follows.

1.2.1 We have $Q_2 = 1/3$ because the probability that randomly selected odd number N is divisible by $p_2 = 3$ is $1/3$.

1.2.2 We have $Q_3 = 1/5 - 1/(5*3)$ because the probability that randomly selected odd number N is divisible by $p_3 = 5$ is $1/5$ and the probability that randomly selected odd number N is divisible by both $p_3 = 5$ and $p_2 = 3$ is $1/(5*3)$.

1.2.3 Similarly we have $Q_4 = 1/7 - \{1/(7*5) + 1/(7*3) - 1/(7*5*3)\}$. Here $1/(7*5*3)$ is necessary because both {the probability that N is divisible by both $p_4 = 7$ and $p_3 = 5$} and {the probability that N is divisible by both $p_4 = 7$ and $p_2 = 3$} contain {the probability that N is divisible by all of $p_4 = 7$, $p_3 = 5$ and $p_2 = 3$}.

1.3 Then we can have the following (25).

$$Q_k = 1/p_k - C_k$$
$$(2 \leq k \leq \pi(\lfloor\sqrt{N}\rfloor) \quad 0 \leq C_k < 1/p_k \quad 0 = C_k \text{ only at } k = 2) \tag{25}$$

C_k : the correction value for the fact that N is not divisible by any of ($p_{k-1}, p_{k-2}, \cdots\cdots, p_4 = 7, p_3 = 5, p_2 = 3$)

We have the following (26) from the above (25).

$$0 < Q_k \leq 1/p_k \qquad (Q_k = 1/p_k \text{ only at } k = 2) \tag{26}$$

1.4 $Q(N)$ increases with increase of N due to the following reasons.

1.4.1 $\pi(\lfloor\sqrt{N}\rfloor)$ increases with increase of N.

1.4.2 Since $Q(N)$ has $\{\pi(\lfloor\sqrt{N}\rfloor)-1\}$ terms as shown in (24), the total number of term of $Q(N)$ increases with increase of N.

1.4.3 Q_k has positive value as shown in (26) because Q_k is probability.

1.4.4 Since the value of Q_k depends on k but not N as shown in item 1.2 and 1.3, even if the total number of term of $Q(N)$ increases by 1 after increase of N, the value of each Q_k which existed before increase of N does not change.

1.4.5 When N increases from $N = N_0 - 2$ to $N = N_0$, if a prime number does not exist in the range of $\sqrt{N_0 - 2} < r \leq \sqrt{N_0}$ (r : real number), $Q(N)$ does not change. But if a prime number $p_{\pi(\lfloor\sqrt{N_0}\rfloor)} = p_{\pi(\lfloor\sqrt{N_0-2}\rfloor)+1}$ exists in the range of $\sqrt{N_0 - 2} < r \leq \sqrt{N_0}$, $Q(N)$ increases by $Q_{\pi(\lfloor\sqrt{N_0}\rfloor)}(> 0)$.

Since $Q(N)$ increases with increase of N, $P(N) = 1 - Q(N)$ decreases with increase of N. We have the following (27) from (10).

$$\lim_{N\to\infty} P(N) = 0 \tag{27}$$

References

[1] THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES

Toshihiko Ishiwata
E-mail: toshihiko.ishiwata@gmail.com
https://www.toshihiko-ishiwata.com/