Proof of Goldbach conjecture

By Toshihiko ISHIWATA

Dec. 30, 2024

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 $1 \le l(n)$ 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 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 $(6 \le)n$ can be divided into 2 prime numbers.

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

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

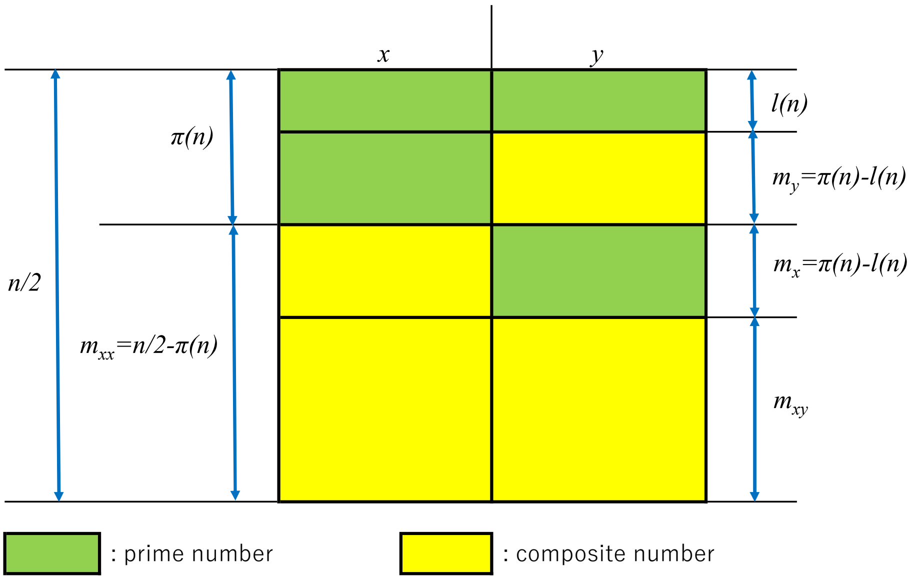


Figure 1 : Various pairs in $n/2$ pairs of (x, y)

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 $\pi(n-1)-1 = \pi(n)-1 \fallingdotseq \pi(n)$ 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).

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

Then we have $\lim_{n\to\infty} \dfrac{l(n)}{\pi(n)} = 0$ 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.

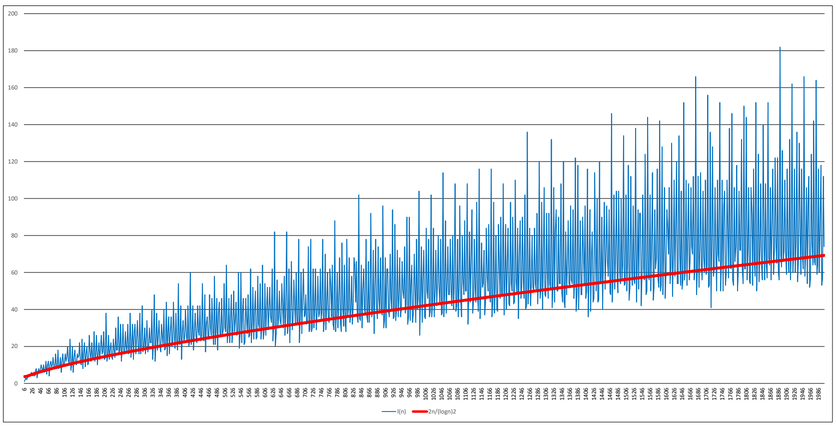
$$l(n) \quad \sim \quad \frac{\{\pi(n)\}^2}{n/2} \quad \sim \quad \frac{\{n/\log n\}^2}{n/2} = \frac{2n}{(\log n)^2} \qquad (n \to \infty) \qquad (9)$$

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 $\dfrac{2n}{(\log n)^2}$ is an increasing function regarding n.

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

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



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

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

3.1 According to Prime number theorem, the probability that randomly selected integer N is a prime number is $1/\log N$ as shown in (5). If N is an even number, the probability that N is a prime number is zero. Then we have the following equation. P_o is the probability that N is a prime number when N is an odd number.

$$(1/2)*0+(1/2)*P_o = 1/\log N \qquad \longrightarrow \qquad P_o = 2/\log N_o$$
$$(N : \text{randomly selected integer} \qquad N_o : \text{odd number})$$

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. Then the probability that $(k, n-k)$ or $(n-k, k)$ is a prime pear is $4/\{\log k \log(n-k)\}$.
$(k = 3, 5, 7, 9, \ldots\ldots, n/2-4, n/2-2, n/2 \qquad n/2 :$ odd number)
$(k = 3, 5, 7, 9, \ldots\ldots, n/2-5, n/2-3, n/2-1 \qquad n/2 :$ even number)
Since $(1, n-1)$ and $(n-1, 1)$ are always a composite pair, k does not include 1. The probability that $(k, n-k)$ or $(n-k, k)$ is a composite pair is

$$1-\frac{4}{\log k \log(n-k)}$$

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 as the following (10). Since $(1, n-1)$ and $(n-1, 1)$ are always a composite pair, we don't have to include them in (10). Then (10) has $(n/2-2)$ terms altogether.

$$\begin{aligned}
&\{\text{the probability of } l(n)=0\} \ : \ a(n)\\
&=\{1-\frac{4}{\log 3\log(n-3)}\}^2\{1-\frac{4}{\log 5\log(n-5)}\}^2\{1-\frac{4}{\log 7\log(n-7)}\}^2\cdots\cdots\\
&\quad \{1-\frac{4}{\log k\log(n-k)}\}^2\cdots\cdots\{1-\frac{4}{\log(n/2+4)\log(n/2-4)}\}^2\\
&\quad \{1-\frac{4}{\log(n/2+2)\log(n/2-2)}\}^2\{1-\frac{4}{(\log n/2)^2}\} \qquad (n/2:\text{odd number})\\
&=\{1-\frac{4}{\log 3\log(n-3)}\}^2\{1-\frac{4}{\log 5\log(n-5)}\}^2\{1-\frac{4}{\log 7\log(n-7)}\}^2\cdots\cdots\\
&\quad \{1-\frac{4}{\log k\log(n-k)}\}^2\cdots\cdots\{1-\frac{4}{\log(n/2+5)\log(n/2-5)}\}^2\\
&\quad \{1-\frac{4}{\log(n/2+3)\log(n/2-3)}\}^2\{1-\frac{4}{\log(n/2+1)\log(n/2-1)}\}^2\\
&\qquad\qquad (n/2:\text{even number})
\end{aligned} \tag{10}$$

3.2 We have the following (12) from the above (10) and the following (11).

$$0 \le 1-\frac{4}{\log k\log(n-k)} = 1-\frac{4}{\log(n/2+K)\log(n/2-K)} \le 1-\frac{4}{(\log n/2)^2} \tag{11}$$
$(k = 3, 5, 7, 9, \ldots\ldots, n/2-4, n/2-2, n/2 \qquad n/2 :$ odd number)

$(k = 3, 5, 7, 9, \ldots\ldots, n/2-5, n/2-3, n/2-1 \qquad n/2 : \text{even number})$
$(K = 0, 2, 4, 6, \ldots\ldots, n/2-7, n/2-5, n/2-3 \qquad n/2 : \text{odd number})$
$(K = 1, 3, 5, 7, \ldots\ldots, n/2-7, n/2-5, n/2-3 \qquad n/2 : \text{even number})$

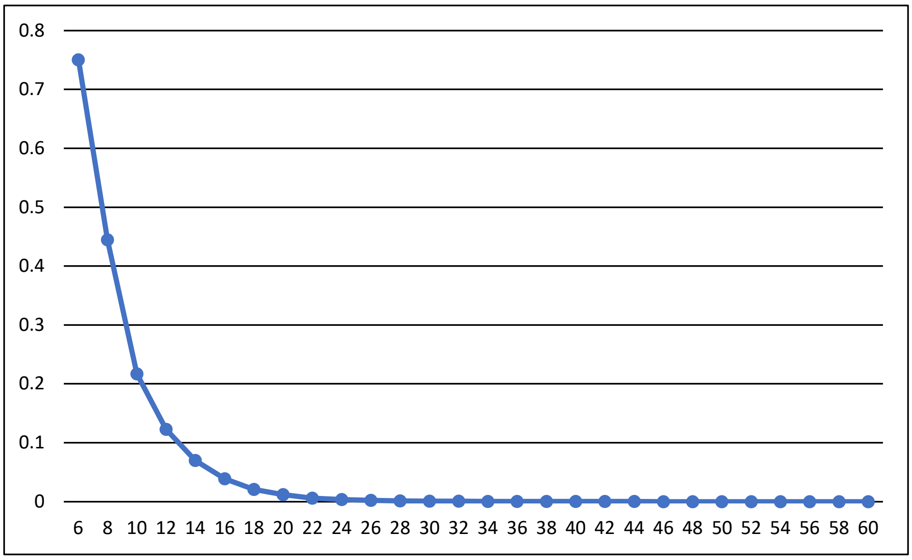
Please refer to [Appendix 1 : Verification of (11)] for verification of the above (11).

$$0 \leq a(n) \leq A(n)$$
$$= \{1 - \frac{4}{(\log n/2)^2}\}^{n/2-2} = [\{1 - \frac{1}{\{(\log n/2)/2\}^2}\}^{\{(\log n/2)/2\}^2}]^{(n/2-2)/\{(\log n/2)/2\}^2}$$
$$\sim \quad (\frac{1}{e})^{(n/2-2)/\{(\log n/2)/2\}^2} \fallingdotseq \frac{1}{e^{(n/2)/\{(\log n/2)/2\}^2}} \qquad (n \to \infty) \qquad (12)$$

We have the following (13) from the above (12).

$$\lim_{n\to\infty} a(n) = 0 \qquad (13)$$

3.3 The following (Graph 2) shows that $a(n)$ decreases with increase of n.



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 $a(n)$

If we calculate $a(n)$ by (10) in $n < 60$, $a(n)$ has a negative value or fluctuates wildly with increase of n. This situation seems to be due to the fact that the smaller n is, the larger the error[%] in approximating $\pi(n)$ to $n/\log n$ becomes. Then we

calculated for the above (Graph 2) using $[1 - \frac{\{\pi(k)-1\}\{\pi(n-k)-1\}}{\{(k+1)/2)\}\{(n-k+1)/2\}}]$ instead of $[1 - \frac{4}{\log k \log(n-k)}]$.
From the above (Graph 2) and (13) we can find that $a(n)$ decreases with increase of n and converges to 0 with $n \to \infty$.

3.4 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 3.1. 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 3.4.2	NO
Case 2	●	X	item 3.4.2	NO
Case 3	X	●	item 3.4.1	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.

3.4.1 $a(n)$ decreases with increase of n and converges to 0 with $n \to \infty$ as shown in item 3.3. Therefore the larger n is, the smaller the probability of zero point occurrence of $l(n)$ is.

3.4.2 Zero point of $l(n)$ does not exist in $n \leq 4{*}10^{18}$ as shown in item 1.2. Goldbach conjecture can be expressed as $l(n)$ does not have any zero point in $6 \leq n$.

Case 1 and Case 2 cannot exist because they contradict item 3.4.2.
Case 3 cannot exist because it contradicts item 3.4.1 as shown in the following item 3.5.

3.5 From (12) we have the following (14) which shows that $a(4*10^{18})$ is extremely small. $A(n)$ is defined in (12).

$$a(4*10^{18}) < A(4*10^{18}) \fallingdotseq \frac{1}{e^{(2*10^{18})/\{\log(2*10^{18})/2\}^2}} = \frac{1}{e^{(2*10^{18})/444}} = e^{-4.5*10^{15}}$$
$$= (e^{4.5})^{-10^{15}} = (10^{2.0})^{-10^{15}} = 10^{-2.0*10^{15}} \tag{14}$$

We can calculate the probability of zero point occurrence of $l(n)$ near $n = 6$ as follows.

$$a(6) = 1 - \{\frac{\pi(3)-1}{(3+1)/2}\}^2 = 1 - (1/2)^2 = 0.75 \tag{15}$$

In Case 3 zero points exist only in $4*10^{18} < n$. Case 3 contradicts $a(n)$ as follows.

3.5.1 The situation where a zero point can exist in $a(n) < 10^{-2.0*10^{15}}$ contradicts the situation where a zero point cannot exist at $a(n) = 0.75$. 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 $a(n)$.

3.5.2 Since 0.75 is extremely larger than the probability of $10^{-2.0*10^{15}}$ that zero points already exist, a new zero point must exist near $n = 6$. But in Case 3 no zero points exist in $n \leq 4 * 10^{18}$.

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

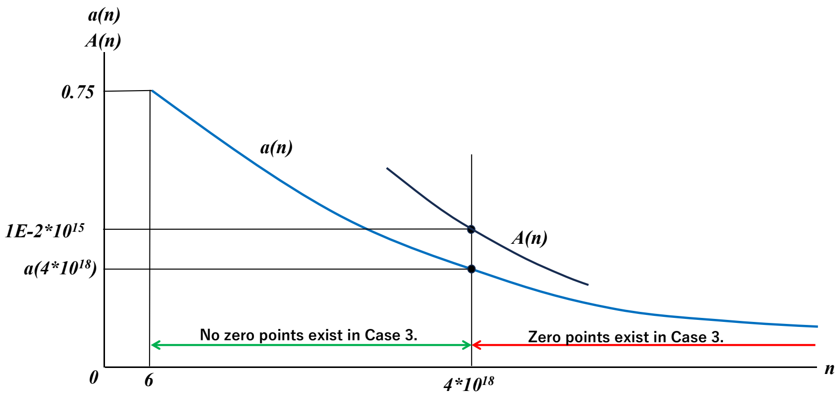


Figure 2 : the contradiction between Case 3 and $a(n)$

3.6 Case 4 is consistent with item 3.4.1 and 3.4.2. Because it is reasonable from item 3.4.1 and 3.4.2 that no zero points exist in $4 * 10^{18} < n$. Among 4 cases of zero point distribution of $l(n)$ shown in (Table 2), only Case 4 can exist. Therefore Case 4 shows the real $l(n)$. We have the following (16) from Case 4 because Case 4 does not have any zero point in $4 * 10^{18} < n$.

$$1 \leq l(n) \qquad (4 * 10^{18} < n) \tag{16}$$

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 (16).

Appendix 1. : Verification of (11)

We have the following (11) in the text.

$$0 \leq 1 - \frac{4}{\log k \log(n-k)} = 1 - \frac{4}{\log(n/2+K)\log(n/2-K)} \leq 1 - \frac{4}{(\log n/2)^2} \tag{11}$$

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

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

$(K = 0, 2, 4, 6, \ldots\ldots, n/2-7, n/2-5, n/2-3 \qquad n/2 : \text{odd number})$

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

In order for the above (11) to hold true, it is sufficient for the following (17) to hold true.

$$\log(n/2+K)\log(n/2-K) \leq (\log n/2)^2 \tag{17}$$

Here we define the following (18) as shown in the following (Figure 3).

$$\log n/2 = A \qquad \log(n/2-K) = A - B \qquad \log(n/2+K) = A + C \tag{18}$$

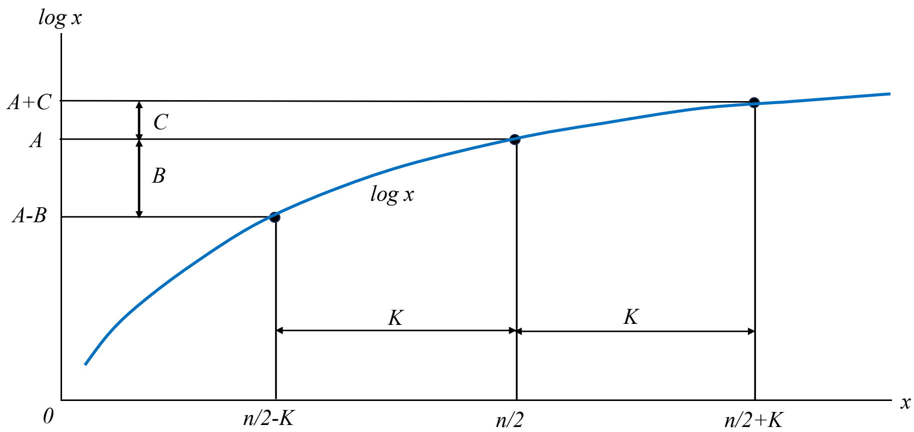


Figure 3 : Relationship among A, B, C and K

Since $\log x$ is a monotonically increasing and districtly concave function regarding x, the following (19) holds true.

$$0 < C < B \quad (1 \leq K) \qquad\qquad 0 = C = B \quad (K = 0) \tag{19}$$

The above (17) holds true from the following (20). $\geq$ is satisfied by the above (19).

$$(\log n/2)^2 - \log(n/2+K)\log(n/2-K)$$
$$= A^2 - (A+C)(A-B) = A(B-C) + BC \geq 0 \tag{20}$$

References

[1] THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES

Toshihiko ISHIWATA
E-mail: toshihiko.ishiwata@gmail.com