Prime twins is unending existence

Shanzhong Zou Email:zoushanzhong@foxmail.com

Abstract

This paper proposes a new number theory method (comb method) to prove the unending existence of prime twins. The natural numbers set is viewed as a union of two sets that don't intersect –s elements sets and h elements set. These elements are selected by a series of combs. By analyzing the distribution of these elements, we get our result.

Key words: *s* elements; *h* elements; comb method ; Pure Distance ; Mixed Distance

MR（2010） Subject classification 11A41

1. Introduction

The twin prime conjecture is a famous unsolved problem in number theory. This conjecture was formally proposed by Hilbert in the 8th problem of the 1900 International Congress of Mathematicians, and can be described as "the existence of an infinite number of twin prime numbers". Twin prime numbers are a pair of prime numbers that differ by 2. For example, 3 and 5, 5 and 7, 11 and 13,..., 10016957 and 10016959 are all twin prime numbers. But to this day, no one has provided a mathematical proof for the infinite existence of twin prime numbers.

This paper proposes a new number theory method (comb method)to prove the unending existence of prime twins. The natural numbers set is viewed as a union of two sets that don't intersect –s elements sets and h elements set. These elements are selected by a series of combs. By analyzing the distribution of these elements, It was found that if twin prime numbers do not exist infinitely, the distribution of these elements is extremely uneven, with differences of different orders from the theoretical distribution pattern. This identified the focal point of the contradiction, this paper focuses on resolving this contradiction.

This paper prove the proposition using methods A and B

A: Proof Method 1, (using advanced mathematics, concise)

B: Proof method 2, (using elementary mathematics to understand the variation pattern of numbers in T)

Firstly, using A to prove the proposition concisely and clearly. Then use B to transform the problem of double prime numbers into a distance problem, then transform the distance into an area problem, and finally prove that if double prime numbers exist in a finite number, they cannot form the area shape they should theoretically form.

2. Preparatory work

2.1 Definition of s elements and h elements

Denote natural numbers set as N, odd numbers set as Q,

$\forall q \in Q$, $\exists! \, n \in N$ satisfies $q = 2n + 1$, that is $n = (q - 1)/2$.

Next, divide the set Q set into two parts:

1) The odd prime number set P,
   $\mathrm{P} = \{p_1 = 3,\ p_2 = 5, p_3 = 7, \cdots p_i = 2s_i + 1 \cdots\}$, s_i is prime number element.
2) The odd composite number set A ,
   $\mathrm{A} = \{a_1 = 9, a_2 = 15, a_3 = 21, \cdots a_i = 2h_i + 1 = p_i q \cdots\}$, $p_i \in P, q \geq 3$.

DEF1. s elements and h elements

Set $p_i = 2s_i + 1$, $a_i = 2h_i + 1$, *then* s_i *is called prime number element,* h_i *is called compound number element.*
$p_i \in P, a_i \in A$. *Obviously* , $P \cup A = N$, $P \cap A = \emptyset$. **#**

DEF2. Representation of h elements

$\forall h \in A, \exists p \in P, q \in Q$, *make* $A = Pq$ *that* $2h + 1 = (2s + 1)(2n + 1)$, [DEF1]
then $h = (2s + 1)n + s, n \geq 1$.
we represent h as following: $h = pn + s$ …… (1) **#**

2.2 Comb Method

LEMMA 1.

$\forall N_e \in N$, *Record the set of h elements in* $0 \sim N_e$ *as* h_e, $\forall h_e \in A$ [DEF1]
$\because h_e = p_i n + s_i$, [DEF2]

$\therefore p_i$ *has the maximum value* p_m, *satisfying* $p_m \leq \sqrt{2N_e + 1}$.

Proof

$\because \mathrm{h}_e \leq N_\mathrm{e}$, $\therefore 2\mathrm{h}_e + 1 \leq 2\, N_\mathrm{e} + 1$,

$\because 2\mathrm{h}_e + 1 = p_i(2\mathrm{n} + 1)$, $p_i \leq (2\mathrm{n} + 1)$, $\therefore$ $p_i \leq \sqrt{2\, N_\mathrm{e} + 1}$, and $p_m \leq \sqrt{2\, N_\mathrm{e} + 1}$.**#**

We denote $P_{N_e} = \{p_1 p_2 p_3 \cdots p_i \cdots p_m\}$, $p_1 = 3, p_2 = 5$,

$p_m = \max(P_{N_e}) \leq \sqrt{2N_e + 1}$,

then $p_i n + s_i, i = 1,2,3 \cdots, n = 1,2,3 \cdots$ can represent all the h elements no larger than N_e.

DEF3. s elements prime twins

If p_i, $p_i + 2$ *are both prime numbers, they are called prime twins,*

$p_i = 2s_i + 1, p_i + 2 = 2(s_i + 1) + 1,\ s_i,\ s_i + 1$ *are both s elements，we define them s twins.* **#**

The unending existence proof of prime twins then can be translated into the unending existence proof of *s* twins.

LEMMA 2.

Suppose that the twin prime is not infinite, if $s_Z,\ s_{Z+1} = s_Z + 1$ *are the last s twins,*

then for any $N_e > p_{Z+1}^2/2$ *satisfying* $\max(p_{N_e}) > p_{z+1}$, ($p_{z+1} = 2(s_Z + 1) + 1$)

take out all the numbers within $0 \sim N_e$ *satisfying* $p_i n + s_i,\ p_i n + s_i + 1, p_i \in p_{N_e}$,

$n = 0,1,2,3 \ldots$, *there is no remainder.*

Proof

If there are remaining elements in $0 \sim N_e$*, set to* N_e*, then there is* $n_e \neq p_i n + s_i$, $n_e \neq p_i n + s_i + 1$, $\because n_e \neq p_i n + s_i, n_e - 1 \neq p_i n + s_i$, $\therefore n_e, n_e - 1$ *is s twins. Which contradict with the assumption of the lemma 2. Hence there is no element left within* $0 \sim N_e$. [LEMMA1] **#**

DEF 4. Comb Method

View prime number p_i *as a comb with two teeth* $s_i, s_i + 1$, [DEF3] *whose length is* p_i*. It moves on the natural numbers axis starting from point 0 and with a length period of* p_i*. The natural numbers that the teeth passed through are combed out. The shape of such combs can be shown as fig.1.* **#**

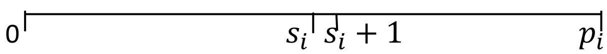

fig1. a comb with length p_i

Thus, the conclusion of lemma2 can be described as：When using elements of set p_{N_e} *as a series of combs to comb* $0 \sim N_e$*, there is no element left within* $0 \sim N_e$. [lemma2]

3. Primary Lemmas and Conclusions

LEMMA3. (Remainder Elements Lemma)

$\forall N_e \in N$, there is $P_{N_e} = \{p_1, p_2, p_3, \ldots p_i \ldots p_m\}$ [lemma1]

Let $T = p_1 p_2 p_3 \ldots p_i \ldots p_m = \prod_{i=1}^{m} p_i$, within set $\{p_1, p_2, p_3, \ldots p_i \ldots p_m\}$ as a series of combs to comb $0 \sim T$, then the number of remainder elements within $0 \sim T$ is g_T,

$$g_T = (p_1 - 2)(p_2 - 2)(p_3 - 2) \ldots (p_m - 2) = \prod_{i=1}^{m}(p_i - 2) \qquad \ldots\ldots \quad (2)$$

g_T is the total number of remaining elements and serves as the basis for future discussions on the distribution of remaining elements

Proof

If n_e are remainder elements, thus $n_e \not\equiv s_i \bmod p_i$ *and* $n_e \not\equiv (s_i+1) \bmod p_i$, $1 \le i \le m$. Inductively, equation (2) is proved. ($\not\equiv$—no congruence) #

Inference 1:

Let $T = p_1 p_2 p_3 \ldots p_i \ldots p_m$, *comb* $0 \sim \mathrm{T}$ *with* $p_1, p_2, p_3, \ldots p_i \ldots p_m$ *denote that the number of remainder elements within* $0 \sim \mathrm{T}$ *as* g_T*, the average distance of these remainder elements as d, then it follows that* $d < p_m$.

Proof

According to the remainder element lemma

$d = T/g_T = p_1 p_2 p_3 \ldots p_m/(p_1 - 2)(p_2 - 2)(p_3 - 2) \ldots (p_m - 2)$

$= \prod_{i=1}^{p_i} p_i/(p_i - 2).$

$\because p_{i+1} - 2 \ge p_i, \exists!\, p_{i+1} = p_i + 2,\ p_{i+1} - 2 = p_i, \quad \therefore d < p_m.$ #

d is the average distance of the distribution of remaining elements and is an important parameter for future discussions on the distribution of remaining elements

Important notes:

$\because$ average distance of these remainder elements as d$< p_m$, but there is no element left within $0 \sim \mathrm{N_e} = (\mathrm{p_m^2} - 1)/2$,

$\therefore (\mathrm{p_m^2} - 1)/2 \gg p_m > d$.

$\because$ in $\mathrm{g_T}$, the distribution of remaining elements in the T/p_i cycle is basically uniform,

but $(\mathrm{p_m^2} - 1)/2$ and $d < p_m$ is an extremely uneven distribution.

$\therefore$ the contradiction between $d < p_m$ and $(\mathrm{p_m^2} - 1)/2 \gg p_m > d$ is the focus of our problem-solving efforts.

This article analyzes the distribution of g_T in T to solve the contradiction,

$\because$ there are phenomena of different orders in the distribution of remaining elements,

$\therefore$ in the future proof process, there may be some calculations with low accuracy, but the resulting errors will be borne by the reverse direction that is not conducive to the proof, therefore, readers do not need to worry about calculation accuracy when reading.

4. Proposition proof

Prove the proposition using methods A and B

A: Proof Method 1, (using advanced mathematics, concise)

B: Proof method 2, (If elementary mathematics is used to solve the problem, one can understand the variation pattern of the remaining elements in T. But it will take a lot of time)
Whether using A or B, it is aimed at addressing the contradiction of extremely uneven distribution of residual elements caused by $(\mathrm{p}_\mathrm{m}^2 - 1)/2 \gg p_m > d$.

A: Proof Method 1

$\because$ when $m \to \infty$, $(\mathrm{p}_\mathrm{m}^2 - 1)/2 \gg p_m > d$,
$\therefore$ If twin prime numbers are not infinite, then the maximum L between the remaining elements in T must be $\mathrm{L} \geq (\mathrm{p}_\mathrm{m}^2 - 1)/2$. [LEMMA2]

The focus of contradictions:

$$(\mathrm{p}_\mathrm{m}^2 - 1)/2 \gg p_m > d, m \to \infty.$$

1. Coverage Analysis:

Each prime p_i excludes at most two residue classes. Within the interval [0,T) , the number of excluded positions per p_i is $2 \cdot \frac{T}{p_i}$ Using the product formula for independent sieves (Halberstam & Richert, 1974):

Due to overlaps, the actual excluded positions are fewer. The density of remaining elements is:

$$\rho = \prod_{i=1}^{m} (1 - \frac{2}{p_i})$$ [2]

1-1 Upper Bound for Maximum Gap:

Theorem Citation (Maximum Gap Theorem in Sieve Theory):

According to Theorem 2.1 in Ford, K., Green, B., Konyagin, S., & Tao, T. (2016), for a sieve density ρ, the maximum gap L satisfies:

$$\mathrm{L} \ll \frac{\log T}{\rho}$$

Where ρ is the density of remaining elements.

1-1.1 Estimation of Sieve Density ρ:

Estimation of Sieve Density ρ each prime p excludes two residue classes (corresponding to s_i and ($s_i + 1$)) leaving a survival probability of $\left(1 - \frac{2}{p_i}\right)$.

Using the product formula for independent sieves (Halberstam & Richert, 1974):
Further applying the generalized form of Mertens' Third Theorem

$$\rho \sim \rho = \prod_{i=1}^{m} (1 - \frac{2}{p_i}) \sim e^{-2\sum_{i=1}^{m}\frac{1}{p_i}}[5]$$

Combining this with the estimate for the sum of reciprocals of primes

$$\sum_{i=1}^{m} \frac{1}{p_i} \sim \mathrm{loglog} p_m$$

(Hardy & Wright, 1979), we obtain:

$$\rho \sim \frac{1}{(\log p_m)^2}$$

1-1.2 Asymptotic Behavior of the Maximum Gap:

Substituting into the sieve theorem and noting that $\log T \sim p_m$ (since $\sum_{i=1}^{m} p_i \sim e^{p_m}$ by the Prime Number Theorem):

$$\mathrm{L} \ll \frac{\log T}{\rho} \sim \frac{p_m}{(\log p_m)^{-2}} = p_m (\log p_m)^2$$

1-1.3 Comparison of Growth Orders:

For sufficiently large p_m the polynomial growth order $p_m (\log p_m)^2$ is significantly smaller than the quadratic function $\frac{{p_m}^2-1}{2}$ Specifically:

$$\lim_{p_m \to \infty} \frac{p_m (\log p_m)^2}{\frac{p_m^2}{2}} = \lim_{p_m \to \infty} \frac{2(\log p_m)^2}{p_m} = 0$$

Thus, there exists p_m such that $p_m\ (\log p_m)^2 < \frac{{p_m}^2-1}{2}$ leading to $\mathrm{L} < \frac{{p_m}^2-1}{2}$

2. Coverage Analysis:

Assume there exists a completely excluded interval$[a, a + \frac{{p_m}^2-1}{2}]$

Every position x in this interval must be excluded by at least one primep_i By the Covering theorem (Maynard, 2016),

Let each prime $p_i$can cover at most $\mathrm{L} = \frac{{p_m}^2-1}{2}$,

The number of covered positions is: $\left\lfloor \frac{L}{p_i} \right\rfloor + 1$

$\because \left\lfloor \frac{L}{p_i} \right\rfloor + 1 \leq \frac{L}{p_i} + 1,\ \therefore\ \left\lfloor \frac{({p_m}^2-1)/_2}{p_i} \right\rfloor + 1 \leq \frac{{p_m}^2-1}{2p_i}$

According to the coverage theorem（Maynard, 2016），The total coverage is bounded by: $\sum_{i=1}^{m} \left(\frac{({p_m}^2-1)/2}{p_i} + 1 \right) \leq \frac{{p_m}^2-1}{2} \sum_{i=1}^{m} \left(\frac{1}{p_i} \right) + m$,

Derivation of Contradiction:

Using $\sum_{i=1}^{m}\frac{1}{p_i}+m\sim \text{loglog}p_m$ and $m\sim\frac{p_m}{logp_m}$, (Prime Number Theorem), the total coverage satisfies:

$\frac{{p_m}^2-1}{2}$ loglog$p_m+\frac{p_m}{logp_m}\ll\frac{{p_m}^2-1}{2}$

However, the coverage required is $\frac{{p_m}^2-1}{2}+1$. Clearly, the total coverage is insufficient to exclude the entire interval, leading to a contradiction.

$\because$ This contradicts the requirement to cover all $\frac{{p_m}^2-1}{2}+1$ positions.

$\therefore \nexists L\geq\frac{{p_m}^2-1}{2}$, Prime twins is unending existence **#**

B: Proof method 2

LEMMA 4.

Let $T=p_1p_2p_3\dots p_i\dots p_m$, comb $0\sim T$ with $p_1,p_2,p_3,\dots p_i\dots p_m$, denote the number of remainder elements as g_T [lemma3], then when $m\to\infty,\exists j<m$, satisfying

$(p_j^2-1)/2\geq 2T/g_T$, and $m-j\to\infty$

Proof

$\because d=T/g_T<p_m$, [Inference 1] *let* p_d *be the largest prime number less than* $\sqrt{4p_m+1}$, *that is* $p_d<\sqrt{4p_m+1}$ *and* $p_{d+1}\geq\sqrt{4p_m+1}$,

$\therefore$ *it is well known that the number of prime numbers less than* p_m *is* $\pi(p_m)\approx p_m/logp_m$, [1] *the number of prime numbers less than* p_d *is* $\pi(p_d)\approx p_d/logp_d$.
Obviously, when $m\to\infty,\pi(p_m)\gg\pi(p_d)$, [2] *let* $j=d+1$,

then $(p_j^2-1)/2\geq 2p_m\geq 2T/g_T$ *and* $m-j\to\infty$ #

We now come to analyze the changes of remainder elements' distributions within $0\sim T$ when it is combed by elements in set $\{p_1,p_2,p_3,\dots p_i\dots p_m\}$. When $0\sim T$ hasn't been combed, element pairs within $0\sim T$ whose distance equals to 1 form the set

$D(d=1)=\{(0,1),\ (1,2),\ (2,3),\ \dots(T-1),T)\}$.

The total number of these pairs is T; element pairs whose distance equals to 2 form the set

$D(d=2)=\{(0,2),\ (1,3),\ (2,4),\ \dots(T-2),T)\}$.

The total number of these pairs is $\mathrm{T}-1$; element pairs whose distance equals to 3 form the set

$\mathrm{D}(\mathrm{d}=3)=\{(0,3),\ (1,4),\ (2,5),\ \ldots(\mathrm{T}-3),\mathrm{T})\}$.

The total number of these pairs is $\mathrm{T}-2$; ......element pairs whose distance equals to K form the set

$\mathrm{D}(\mathrm{d}=\mathrm{K})=\{(0,\mathrm{K}),\ (1,\mathrm{K}+1),\ (2,\mathrm{K}+2),\ \ldots(\mathrm{T}-\mathrm{K}),\mathrm{T})\}$.

The total number of these pairs is $\mathrm{T}-\mathrm{K}+1$,

When using comb $p_1=3$ to comb out those points within $0\sim T$ equal to $p_1 n+1,\ \ p_1 n+2,\ \ n=0,1,2,3,\cdots$, we obtain that the number of remainder elements within $0\sim\mathrm{T}$ is $T/3=p_2p_3\ldots p_m=\prod_{i=2}^{m}p_i$, and the distance of any two remainder elements is $3n$, such as element pairs with distance 3 form the set

$$\mathrm{D}(\mathrm{d}=3)=\{(0,3),\ (3,6),\ (6,9),\ \ldots(\mathrm{T}-3),\mathrm{T})\}$$

the total number of such pairs is $T/3$;element pairs with distance 6 form the set

$$\mathrm{D}(\mathrm{d}=6)=\{(0,6),\ (3,9),\ (6,12),\ \ldots(\mathrm{T}-6),\mathrm{T})\}$$

the total number of such pairs is $(T/3)-1$;
element pairs with distance $3n$ form the set

$$\mathrm{D}(\mathrm{d}=3\mathrm{n})=\{(0,3\mathrm{n}),\ (3,3\mathrm{n}+3),\ (6,3\mathrm{n}+6),\ \ldots(\mathrm{T}-3\mathrm{n}),\mathrm{T})\}$$

the total number of such pairs is$(\mathrm{T}/3)-n+1$;

DEF 5. Pure Distance Pairs and Mixed Distance Pairs

Let the distance between two remaining elements be K, $\forall D(d=k)$*, If there are no other elements in d, we call this segment K - pure distance pairs, represent the set they form as* $C\ (d=K)$*, and represent the number as* $g\ (C\ (d=K))$

$\forall D(d=K)$*, If there are other elements in d, we call them K-mixture distance pairs. Represent the set they form as* $Z\ (d=K)$ *and the numbers as.*$g(D(d=K)$

so $\big(c(d=k)\big)+g\big(Z(d=K)\big)=\ g(D(d=K))$*.*

Obviously, $g(D(d=K))>g(Z(d=K))$*.* ......（3） #

We only discuss the distribution of pure distance pairs here.

We know that after $0\sim T$ is combed by $p_1$and p_2, there are only two different pure distance pairs within $0\sim T$, which respectively form the sets $c(d=3)$and $c(d=6)$, the other distance pairs are all mixed distance pairs. From the fact that all the remainder elements between $0\sim p_1\,p_2$ are 0,6,9,15 , we obtain that remainder elements between $0\sim T$ are $0,6,9,15,21,24,30\ldots15n,15n+6,15n+9$ ....The total number of such elements is $(p_1-2)(p_2-2)\,p_3 \dots p_m$

after $0\sim T$ is combed by p_1, p_2, and p_3, the remainder elements between $0\sim p_1\,p_2\,p_3 = 105$ are $0,6,9,15,21,30,36,51,54,69,75,84,90,96,99,105$..

We can deduced that within $0\sim p_1\,p_2\,p_3 = 105$,

$$\mathrm{g}\big(c(d=3)\big)=3,\ \mathrm{g}\big(c(d=6)\big)=8,$$

$$\mathrm{g}\big(c(d=9)\big)=2,\ \mathrm{g}\big(c(d=15)\big)=2$$

and within $0\sim T$, $\mathrm{g}\big(c(d=3)\big)=3p_4p_5 \dots p_m$,

$$\mathrm{g}\big(c(d=6)\big)=8p_4p_5 \dots p_m$$

$$\mathrm{g}\big(c(d=9)\big)=2p_4p_5 \dots p_m,$$

$$\mathrm{g}\big(c(d=15)\big)=2p_4p_5 \dots p_m$$

From the above discussion, it can be seen that if the maximum pure distance after P comb selection is D, then there are $\mathrm{P}-2$ D in T. As the comb length puii increases, one or both ends of the $\mathrm{P}-2$ D distances will be combed out, and the maximum pure distance D will also increase.

(This description is somewhat colloquial because the expression of mathematical patterns is cumbersome, so if readers need the author to provide mathematical expressions, the author can provide them.)

The maximum pure distance is represented as follows: the pure distance distribution within the range of 0~T is shown in Figure 2, where the area of the graph is T, namely

$$3\times \mathrm{g}\big(c(d=3)\big)+6\times \mathrm{g}\big(c(d=6)\big)+\dots+mc\times \mathrm{g}\big(c(d=mc)\big)=T$$

and the average distance $\mathrm{d}=T/g_T$. [lemma3]

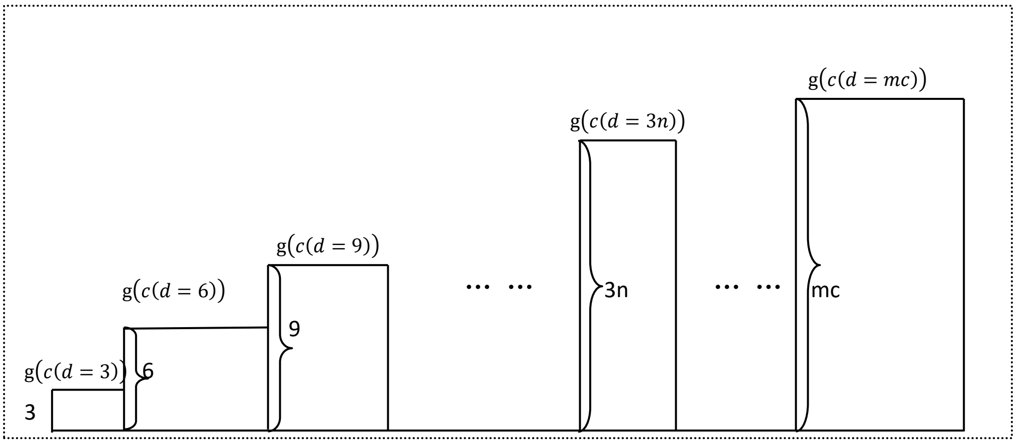

fig2.

fig2 the distribution of pure distance pairs within $0\sim T$ after combed by $p_1, p_2, \ldots, p_K$

We analyze after $0\sim T$ is combed by p_i, relation between of $c(d=k)$ and $D(d=k)$.

LEMMA 5

$\forall \mathrm{C}(\mathrm{d}=\mathrm{k_a}), \exists \mathrm{g}(\mathrm{C}(\mathrm{d}=\mathrm{k_a}))$, settle for $\mathrm{g}\big(\mathrm{C}(\mathrm{d}=\mathrm{k_a})\big) \leq \mathrm{g}\big(\mathrm{D}(\mathrm{d}=\mathrm{k_a})\big)$ [DEF5]

$$\mathrm{g}\big(\mathrm{D}(\mathrm{d}=\mathrm{k_a})\big) = (\mathrm{p_1}-\gamma_1)(\mathrm{p_2}-\gamma_2)\ldots(\mathrm{p_i}-\gamma_\mathrm{i})(\mathrm{p_{i+1}}-4)\ldots(\mathrm{p_m}-4)$$

$$= \prod_{\mathrm{i=1}}^{\mathrm{e}}(\mathrm{p_i}-\gamma_\mathrm{i})\prod_{\mathrm{i=e+1}}^{\mathrm{m}}(\mathrm{p_i}-4) \quad (\mathrm{p_{e+1}} > \mathrm{k_a}+1) \quad \ldots\ldots (3)$$

When $\mathrm{p_i} \leq \mathrm{k_{i+1}}$, $2 \leq \gamma_\mathrm{i} \leq 4$. When $\mathrm{p_i} > \mathrm{k_i}+1, \gamma = 4$.

Explanation: The value of $2 \leq \gamma_\mathrm{i} \leq 4$is only a statement of the range of γ, and there is no precise requirement in future proofs. Readers can skip reading.

Proof

Assume that $(a, a+k)$ is a k- distance pair within $0\sim p_1\, p_2\, p_3 \ldots p_i$, after combing $0\sim T$ is combed by $p_1, p_2, p_3, \ldots p_i$, since k - distance pairs appear periodically with period $p_1\, p_2\, p_3 \ldots p_i$ within $0\sim T$,

set $E = \{(np_1\, p_2\, p_3 \ldots p_i + a, np_1\, p_2\, p_3 \ldots p_i + a + k) | n = 0,1,2,\ldots\}$

$\because p_1\, p_2\, p_3 \ldots p_i + a \not\equiv a \bmod p_{i+1}$ *and*

$p_1\, p_2\, p_3 \ldots p_i + a + k \equiv (a+k) \bmod p_{i+1}$

$\therefore 2 \leq \gamma_i \leq 4$, *we come to discuss the bound of γ, we just discuss the maximal and minimal value of γ.*

(A) *When $k = np_{i+1}$, within $n_s p_1\, p_2\, p_3 \ldots p_i + a$ and $n_s p_1\, p_2\, p_3 \ldots p_i + a + k$*

$\because k = n\, p_{i+1}$, $\therefore a, a+k$ *is combed out by comb $\mathrm{p_{i+1}}$'s same tooth, then $\gamma = 2$.*

(B) *When $k = n\, p_{i+1} \pm 1$, assume $k = n\, p_{i+1} + 1$, if $n_t p_1\, p_2\, p_3 \ldots p_i + a$ is combed out by comb $\mathrm{p_{i+1}}$'s s_i tooth, $\because k = np_{i+1} + 1$*

$\therefore n_t p_1\, p_2\, p_3 \ldots p_i + a + k$ *is combed out by comb $\mathrm{p_{i+1}}$'s $s_i + 1$ tooth, and that $n_t p_1\, p_2\, p_3 \ldots p_i + a$ just is combed out by comb $\mathrm{p_{i+1}}$'s $s_i + 1$ tooth, then $n_t p_1\, p_2\, p_3 \ldots p_i + a + k$ is no combed out by comb $\mathrm{p_{i+1}}$'s tooth.*

here $\gamma = 3$ in the same way, when $k = n\, p_{i+1} - 1$, $\gamma = 3$.

(C) *When $k \neq n\, p_{i+1}$, $k \neq n\, p_{i+1} \pm 1$ or $np_{i+1} > k + 1$, $\because n_u p_1\, p_2\, p_3 \ldots p_i + a$ and $n_u p_1\, p_2\, p_3 \ldots p_i + a + k$ is no at one time combed out by comb $\mathrm{p_{i+1}}$'s tooth,*

here $\gamma = 4$. Inductively, equation (3) is proved. #

4-2 An Unending Existence Proof of Prime Twins

If prime twins don't exist endlessly, assume that p_Z, $p_{Z+1} = p_Z + 2$ are the last prime twins, then $s_Z = (p_Z - 1)/2$, $s_{Z+1} = s_Z + 1 = ((p_Z + 2) - 1)/2$ are the last s twins. [DEF2]

Let $p_u > p_{z+1}$, comb with set $\{p_1\, p_2\, p_3 \dots p_z\, p_{z+1}\, p_{z+2} \dots p_u\}$ is combed, $0 \sim (p_u^2 - 1)/2$, $0 \sim (p_u^2 - 1)/2$ there is no remainder. [lemma2]

Set $\{p_1, p_2, p_3 \dots p_u, p_{u+1}, p_{u+2} \dots p_v\}$, set $(p_u^2 - 1)/2 \geq 2T_v/g_{T_v}$ (p_u and p_v, are a pair of dependent variables), $T_v = p_1\, p_2\, p_3 \dots p_u\, p_{u+1}\, p_{u+2} \dots p_v$,

g_{T_v} is the number of remaining elements after $\{p_1, p_2, p_3 \dots p_u, p_{u+1}, p_{u+2} \dots p_v\}$ comb selects T_v 。

$g_{T_v} = (p_1 - 2)(p_2 - 2)(p_3 - 2) \dots (p_v - 2) = \prod_{i=1}^{v}(p_i - 2)$ [lemma3]

$\because (p_u^2 - 1)/2 \geq 2T_v/g_{T_v}$, $(\mathrm{Z} + 1 < \mathrm{u} < \mathrm{v})$,

$\therefore (v - u)/u = n$, $v \gg \mathrm{u}$. [lemma4]

Divide the set $\{p_1, p_2, p_3 \dots p_u, p_{u+1}, p_{u+2} \dots p_v\}$ set into $\{p_1, p_2, p_3 \dots p_u\}$ and $\{p_{u+1}, p_{u+2} \dots p_v\}$ two parts,

First, comb T_v with $p_1, p_2, p_3 \dots p_u$, in T_v, different pure distance pairs

$g(c(d = k_i))$, $k_i = 3, 6, \dots l$, (l is max pure distance), $l > (p_u^2 - 1)/2 \geq 2T_v/g_{T_v}$,

The pure distance distribution in $0 \sim T_v$ is shown in Figure 3, and its area is T_v.

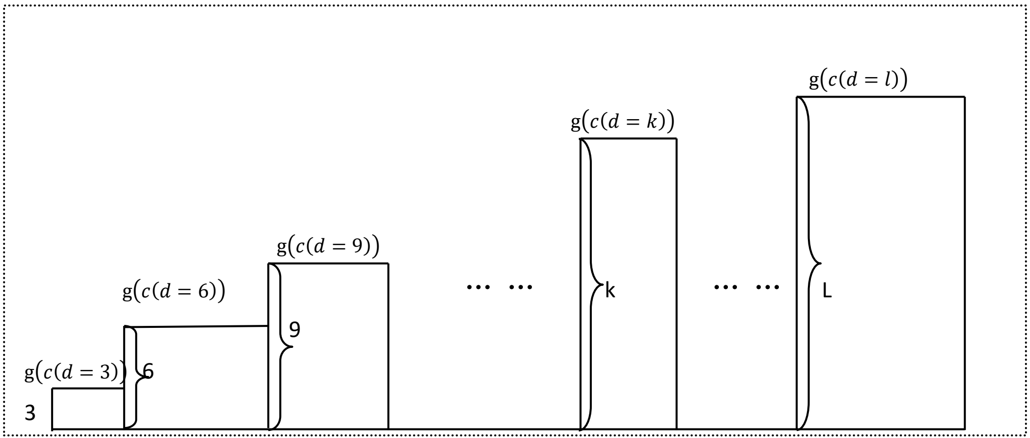

Fig3.

Next, comb T_v with $p_u, p_{u+1}, p_{u+2} \dots p_v$.

In this case, the number of different pure distances is $g'(c(d = k_i))$, k_i=3,6…l…m

$l \sim m$ is the number of new pure distances selected by $(p_{u+1}, p_{u+2}, \dots, p_v)$comb

The pure distance distribution in $0 \sim T_v$ is shown in Figure 4, and its area is T_v

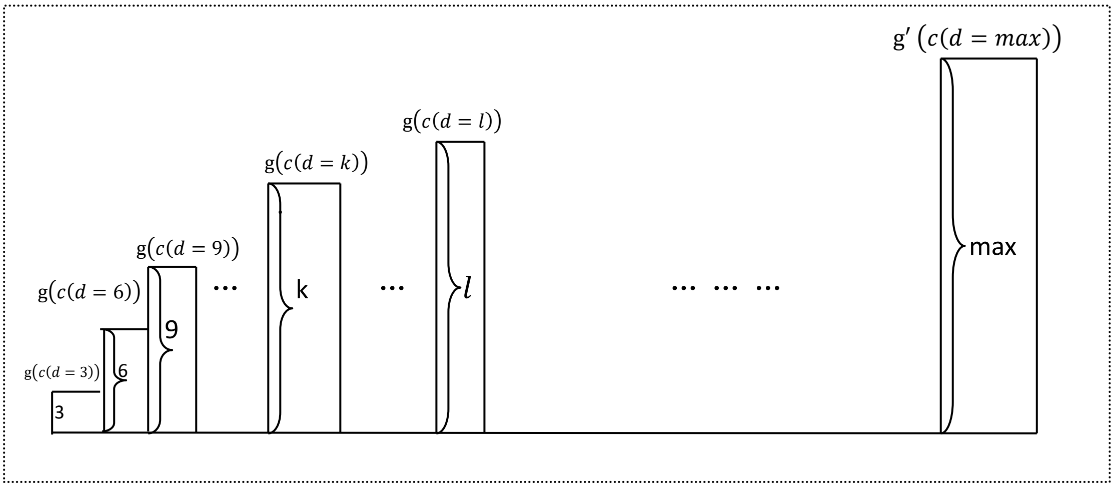

Fig4.

In Fig4 , values with pure distance greater than l are all generated by selecting $0 \sim T_v$ by $p_{u+1}, p_{u+2} \dots p_v$ comb.

Compare the area distribution of Figure 3 and Figure 4, change of area formed by pure distance less than l, that is $\frac{g(c(d=k_i))}{g'(c(d=k_i))} = n$, $(k_i = 3,6,9 \dots l)$, then n increases with the increase of u, v.

Proof

Let $M_c = g\big(C(d = k_c)\big)$, M_c *is the pure distance quantity after* $0 \sim T_v$ *is combed by* $p_1, p_2, p_3 \dots p_u$.

Set $g'(c(d = k_c)) = M_c \prod_{i=u+1}^{v} p_i$, $3 \le k_i \le L$, *If* $g'\big(c(d = k_i)\big)$ *is the pure distance quantity after* $0 \sim T_v$ *is combed by* $p_1, p_2, p_3 \dots p_u \dots p_v$,

then $g'\big(c(d = k_a)\big) \le \prod_{i=1}^{e}(p_i - \gamma_i) \prod_{i=e+1}^{v}(p_i - 4)$. [lemma5]

We have:

$$\frac{g\big(c(d = k_i)\big)}{g'\big(c(d = k_i)\big)} = \frac{M_c \prod_{i=u+1}^{v} p_i}{g'\big(c(d = k_a)\big)} \ge \frac{M_c \prod_{i=u+1}^{v} p_i}{\prod_{i=1}^{e}(p_i - \gamma_i) \prod_{i=e+1}^{v}(p_i - 4)}$$

$$= \frac{M_c}{\prod_{i=1}^{e}(p_i - \gamma_i)} \; \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i - 4)} \qquad \dots\dots \qquad (4)$$

$\because v \gg e$, [lemma4]

$\therefore$ *in (4), we only need to analyze* $\frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i - 4)}$.

$$\because \frac{P_i}{P_i - 4} > 1$$

$$\therefore \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)} \xrightarrow{V\to\infty} \infty$$

$$\because \frac{v-u}{u} = n\ v \to \infty, n \to \infty$$

$$\therefore \frac{M_c}{\prod_{i=1}^{e}(p_i-\gamma_i)}\ \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)} = n,$$

n increases with the increase of u, v. #

Fig. 5 is pure distance distribution after $0 \sim T_v$ is combed by $p_1, p_2, p_3 \ldots p_u \ldots p_v$

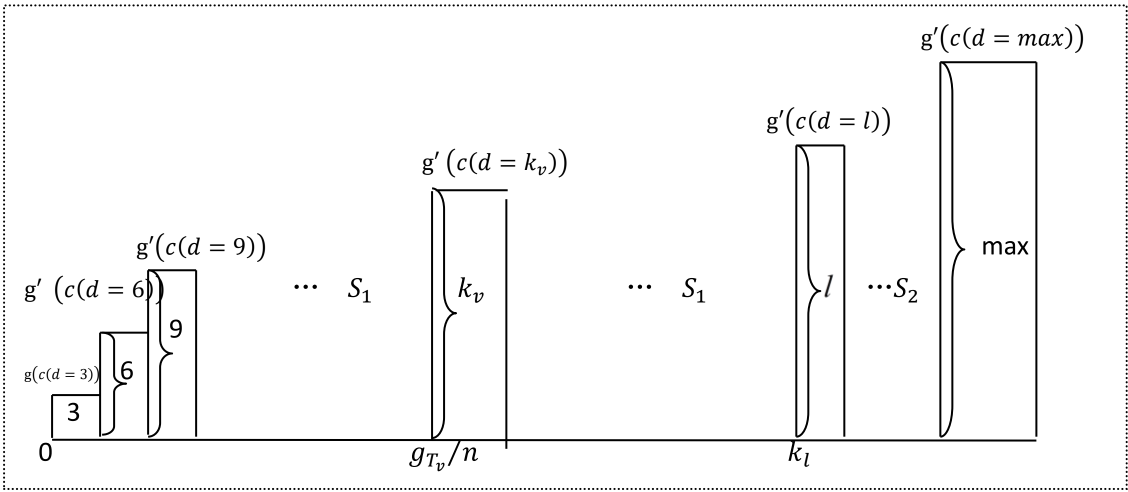

fig5

In fig5 S_1 is an area map made up of pure distance $g'(c(d \le L))$

$S_1 = 3 \times g'(c(d=3)) + 6 \times g'(c(d=6)) \ldots + l \times g'(c(d=l))$

S_2 is an area map made up of pure distance $g'(c(d > l))$

$S_2 = (l+3)\, g'(c(d=l+3)) + (l+6)\, g'(c(d=l+6)) \ldots + (mc)\, g'(c(d=mc))$

We can gat $\frac{S_2}{S_1} \xrightarrow{v\to\infty} \infty$, that is $S_2 \gg S_1$. [lemma4]

In fig5, Point k_l corresponds to $d = l$,

$\because l \ge 2T_v/g_{T_v}$,

$\therefore$ the distance from 0 to point k_l must be greater than $g_{T_v}/2$, if $k_l < \frac{g_{T_v}}{2}$,

then $l\frac{2T_v}{g_{T_v}} > T_v$.

Set g_{T_v}/n it's a point in the S_1 area , $n > 2$, the pure distance corresponding to point g_{T_v}/n is k_v, set $k_v = \frac{T_v}{g_{T_V}}\beta$, $0 < \beta < 1$,

If the average number of remaining elements in $g\big(c(d=3)\big) \sim g\big(c(d=k_v)\big)$ is g_v, then $g_v \frac{k_v}{3} = \sum_{i=3n}^{k_v} g\big(c(d=i)\big)$ that is $k_v = \frac{3}{g_v}\sum_{i=3n}^{k_v} g\big(c(d=i)\big)$,

$\because k_v = \frac{T_v}{g_{T_v}}\ \beta$,

$\therefore \frac{3}{g_v}\sum_{i=3n}^{k_v} g\big(c(d=i)\big) = \frac{T_v}{g_{T_v}}\ \beta$.

Have $\beta = \frac{3g_{T_v}}{T_v g_v}\sum_{i=3n}^{k_v} g\big(c(d=i)\big)$ ...... (5)

Comparison between (4) and (5), have :

$$\frac{M_C}{\prod_{i=1}^{e}(p_i-\gamma_i)}\ \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)},\ \frac{S_2}{S_1} \xrightarrow{v\to\infty} \infty,$$

$\because$ (4) $\xrightarrow{v\to\infty} \infty$,

$\therefore$ must have (5) $\xrightarrow{v\to\infty} 0$.

Proof

$\because k_v = \frac{T_v}{g_{T_v}}\beta$, *If β is not convergent and tends to be zero,*

$\therefore k_v \xrightarrow{v\to\infty} \epsilon$,

$\because$ $g_{T_v}/n \sim \frac{g_{T_v}}{2}$ *range is in the* S_1 *region, and* $S_\epsilon > (\frac{g_{T_v}}{2} - \frac{g_{T_v}}{n}) k_v = (\frac{g_{T_v}}{2} - \frac{g_{T_v}}{n})\,\epsilon$,

S_ϵ *increases with the increase of V, make* $\frac{S_2}{S_1} \xrightarrow{v\to\infty} \infty$ *not tenable,*

$\therefore$ *(5)* $\xrightarrow{v\to\infty} 0$. *#*

$\because$ *In (5), although there is* $\frac{g_{T_v}}{T_v} \xrightarrow{v\to\infty} 0$, $\sum_{i=3n}^{k_v} g\big(c(d=i)\big) > g_v$ *is divergent,*

$\therefore$ *(5) is* $\infty \times 0$ *series, no definite trend,*

That is $(\frac{1}{g_v}\sum_{i=3n}^{k_v} g\big(c(d=i)\big) = \frac{k_v-3}{3} g_v$, *obviously,* $\frac{k_v-3}{3} g_v$ *increases with the increase of v.*

In(5), although there is

$$\frac{g_{T_v}}{T_v} = \frac{(p_1-2)(p_2-2)(p_3-2)\ldots(p_v-2)}{p_1 p_2, p_3 \ldots p_v} \xrightarrow{v\to\infty} 0,$$

But $\frac{(p_v-2)}{p_v} \xrightarrow{v\to\infty}$ *1, so the convergence rate is slower than the divergence rate of* $\frac{k_v-3}{3} g_v$.

So (5) will not converge to 0, so the original assumption is not true, so the twin prime number is infinite. #

REFERENCES

[1] Chaohao Gu, Mathematics Dictionary, Shanghai Dictionary Press, (1992)

[2] Halberstam, H., & Richert, H.-E. *Sieve Methods* （1974）

[3] Ford, K., Green, B., Konyagin, S., & Tao, T. "Long gaps between primes"Journal: Journal of the American Mathematical Society （2016）

[4] Ford, K. "The distribution of integers with divisors in arithmetic progressions" Journal: Acta Arithmetica （2002）

[5] Ireland, K., & Rosen, M. A Classical Introduction to Modern Number TheoryPublisher: Springer-Verlag (1990)

[6] Maynard, J. "Large gaps between primes" Journal: Journal of the American Mathematical Society (2016)