The Comb Sieve Method and Its Application to Conjectures on Prime Distribution

Shanzhong Zou Email: zoushanzhong@foxmail.com

Abstract: The Twin Prime Conjecture, the Goldbach Conjecture, and the Polignac Conjecture are all conjectures concerning the distribution of prime numbers. This paper introduces a novel sieve-based framework termed the " Comb Sieve Method ". This approach not only unifies the three conjectures by treating them as different manifestations of an underlying fundamental problem but also circumvents the need to estimate error terms, a common challenge in traditional sieve methods. We believe this framework will open new avenues for research in this area of number theory.

Key words: Comb Sieve Method; Order difference; Pure Distance; Mixed Distance Pairs
MR（2010） Subject classification: 11A41

1. Introduction

We recall that Hilbert presented the Twin Prime Conjecture [2] in 1900, which asserts the existence of infinitely many primes p such that p + 2 is also prime. Despite breakthroughs such as Zhang's [4] bounded gaps and Maynard's [5]multidimensional refinements, this conjecture remains unproven. Similarly, in the study of the Goldbach Conjecture, although the Chinese mathematician Chen Jingrun achieved the landmark "1+2" result (Chen's theorem) [7], the ultimate goal of proving "1+1" has yet to be attained.

To address these longstanding challenges, this paper introduces the Comb Sieve Method, which reframes the problems within a geometric framework. By strategically assuming the finiteness of twin primes and the potential falsehood of the Goldbach Conjecture [2], we establish the Comb Sieve Method for these three conjectures. This unified approach rigorously demonstrates that the Goldbach Conjecture, the Twin Prime Conjecture, and the Polignac Conjecture [1] belong to the same fundamental class of problems.

If $N_e = \omega(d)$, then N_e and d are called “Order difference”.

“Order difference” serves as a crucial basis for our study of the distribution of remaining elements in T_m and for identifying contradictions. During the analysis of the distribution of the remaining elements in T_m, we observed two instances where “Order difference” occurred.

Central to our proof strategy is the derivation of the geometric distribution of residual elements in a prime sieve. This derivation provides a rigorous proof for the conjectures under consideration. Crucially, as our approach avoids any use of probability densities, it eliminates the error terms typically associated with traditional sieve methods. We sincerely invite experts to review and critique this framework and its supporting analysis.

We will complete the proof in three steps.

Step1: Define two “order differences”.

Step2: The maximum distance between consecutive elements in sieved.

Step3: Taking Twin Prime Conjecture as an example, we will analyze the distribution of the remaining elements in T_m .

2. Definitions and Preliminaries

2.1 Prime and Composite Elements

Let N denote the natural numbers and $Q_{odd}\{q \mid q \equiv 1\ mod 2\}$ the odd integers.

For $q \in Q$, there exists a unique $n \in N$ such that $q = 2n + 1,\ n \geq 1$. Partition Q into:

2.1.1: Prime Set P: $P = \{p_i = 2s_i + 1 \mid s_i \in N, p_i$ is prime$\}$.
2.1.2: Composite Set A: $A = \{a_i = 2h_i + 1 \mid h_i \in N, a_i$ is composite$\}$
$\because P \cup A = N,\ P \cap A = \emptyset, \therefore s_i \cup h_i = n,\ s_i \cap h_i = \emptyset.$

Definition 1: s - Elements and h -Elements

For $p_i \in P$, *define* s_i *as the prime element where* $p_i = 2s_i + 1$. s_i *is* a s *- Elements*
For $a_i = pq \in A$ *define* h_i *as the composite element where* $a_i = 2h_i + 1$. h_i *is* a h *- Elements*

Lemma 1: Representation of Composite Elements

Every composite h_i can be expressed as $h_i = p_i\text{n} + s_i$ for some $p_i \in P, n \geq 1$.

Proof:

Let $a_i = pq = (2s_i + 1)(2n + 1) = 2(p_i n + s_i) + 1$ *hence* $h_i = p_i n + s_i$ *and* $n \geq 1$.■

3. The Comb Sieve Method

3.1 Sieving Process

The goal of this section is to demonstrate how to systematically exclude composite numbers and identify potential twin prime positions by constructing a screening model based on prime number "combs".

Definition 2 :

Comb Sieve: $\forall\ p_i$, and s_i one of his comb teeth, and add another comb tooth $n_i, 0 \leq n_i < p_i$ as shown in fig.1.

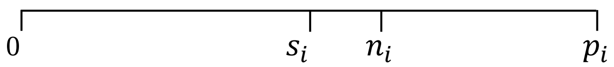

Fig.1.

3.2 Twin Prime Comb Sieve Method

For a sieve bound N_e *let* $P_{N_e} = \{p_1, p_2, \dots p_m\}$, *where* $P_m \le \sqrt{2N_e + 1}$, *each prime* p_i *acts as a "comb" removing residues* s_i *and* $n_i = s_i + 1$ *modulo* p_i. ■

The shape of such combs can be shown as fig.*2.*

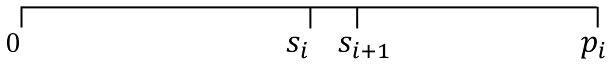

Fig2. a comb with length p_i

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 $0 \sim \mathrm{N_e}$ satisfying $p_i n + s_i$, $p_i n + s_i + 1, p_i \in p_{N_e}$,

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

Proof:

If there are remaining elements in $0 \sim N_e$, *set to* N_e, *then there is* $n_e \ne p_i n + s_i$, $n_e \ne p_i n + s_i + 1$, $\because n_e \ne p_i n + s_i, n_e - 1 \ne p_i n + s_i$, $\therefore n_e, n_e - 1$ *is s twins.*

Which contradict with the assumption of the lemma 2. Hence there is no element left within $0 \sim N_e$. ■

3.3 Barinak Conjecture Comb Sieve Method

Let p be a prime number such that $p + 2t$, t > 1 is also prime. We refer to such pairs $(p + 2t)$ as $2t$ ***-prime pairs***. The Barinak Conjecture asks: *Do infinitely many prime pairs with fixed even difference* $2t$ *exist?*

Proof by Contradiction:

Assume only finitely many such prime pairs exist. Let $s_z, s_z + t$ *be the last pair of prime elements corresponding to primes* $p_z = 2s_z + 1$, *and* $p_z + 2t = 2s_z + 2t + 1 = 2(s_z + t) + 1$

For any

sieve bound $N_e > (p_z^2 + 1)/2$, *after removing all elements satisfying :*

The interval $[0, N_e]$ *should contain no residual elements.*

The shape of such combs can be shown as fig.3.

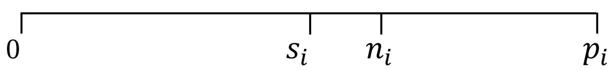

fig.3.

If $\nexists p_z + 2t \in P$, So after N_e is filtered by $p_i n + s_i$, $p_i n + n_i$, ($p_i n + n_i$ corresponds to $p_z + 2t \in P$). the interval $[0, N_e]$ should contain no residual elements.

Proof:

(Same as the proof method of Twin Prime Comb Sieve Method) ■

3.4 Goldbach Conjecture Comb Sieve Method

Goldbach Conjecture, $\forall\ 2n, n \geq 3$, $\exists 2n = p_i + p_j$,
that is $n - 1 = s_i + s_j \Longrightarrow N_e = s_u + s_v \Longrightarrow 2N_e + 2 = p_u + p_v$

Let $P_{N_e} = \{p_1, p_2, \dots p_m\}$, *where* $P_m \leq \sqrt{2N_e + 1}$

Two opposite directions number axis：$0 \sim N_e$, $0 \sim N_e'$ fig.4.

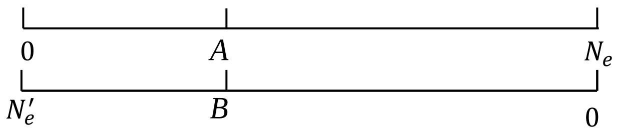

fig.4

$s_u = 0 \sim \mathrm{A} \neq p_i n + s_i, s_v = B \sim 0 \neq p_j n + s_j$, then A and B correspond to $N_e = s_u + s_v$.
Let $N_e' \sim B = p_i n + n_i, 0 \leq n_i < p_i$,
The shape of such combs can be shown as fig.5.

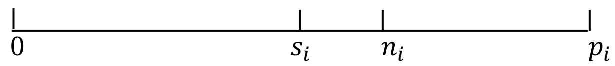

fig.5

If $N_e \neq s_u + s_v$, so after N_e is filtered by $p_i n + s_i$, $p_i n + n_i$, the interval $[0, N_e]$ should contain no residual elements.

Proof:

(*Same as the proof method of Twin Prime Comb Sieve Method*)■

The Goldbach Conjecture, Twin Prime Conjecture, and Barinak Conjecture all have a common point, that is, given N_e, there is a set $P_{N_e} = \{p_1, p_2, \dots p_m\}$, *where* $P_m \leq \sqrt{2N_e + 1}$, and each $p_i \in P_{N_e}$ in P_{N_e} has two comb teeth. After screening N_e with P_i, it is determined whether there are any remaining elements in N_e, and the difference is that the position of one comb tooth is different.

4. Order difference

if $N_e = \omega(d)$, then N_e and d are called “Order difference”.

Let $T_m = p_1 p_2 \dots p_m = \prod_{i=1}^{m} p_i$, $p_i \in P_{N_e}$.

If $P_m > \sqrt{2N_e+1}$, then P_{N_e} will exhibit a step1 and sttep2 "order difference" when screening T_m

$\forall \mathrm{N_e} \geq \mathrm{p_m^2}/2$, set $P_{N_e} = \{p_1, p_2, \ldots p_m\}$, $\mathrm{p_{m+1}^2}/2 > \mathrm{N_e}$, P_{N_e} is the set of all odd prime numbers less than p_m.

After using P_{N_e} to filter out the elements in T_m that satisfy $p_i n + s_i$, $p_i n + n_i$, according to the Chinese remainder theorem, the number of remaining elements in T_m is $\prod_{i=1}^{m}(p_i - 2)$, We obtain the average distance between the remaining elements d in Lemma 3 .

Lemma 3: $d = \prod_{i=1}^{m} \frac{p_i}{p_i-2} = \frac{p_1}{p_1-2}\frac{p_2}{p_2-2}\ldots\frac{p_{m-1}}{p_{m-1}-2}\frac{p_m}{p_m-2} < p_m$

Proof.

$\frac{p_1}{p_1-2}\frac{p_2}{p_2-2}\ldots\frac{p_{m-1}}{p_{m-1}-2}\frac{p_m}{p_m-2} = p_m\left(\frac{p_1 p_2 \ldots p_{m-1}}{(p_1-2)(p_1-2)\ldots(p_{m-1}-2)(p_m-2)}\right) = p_m A$

$\because p_1 - 2 = 3 - 2 = 1,\ \therefore A = \frac{p_1}{p_2-2}\frac{p_2}{p_3-2}\ldots\frac{p_{m-1}}{p_m-2}$

$p_i - 2 \geq p_{i-1}$, e.g.: $p_i = 11, p_{i-1} = 7$, $p_i - 2 = 11 - 2 = 9$, $9 > 7$.

Thus: $\frac{p_{i-1}}{p_i-2} \leq 1 = \frac{p_1}{p_2-2}\frac{p_2}{p_3-2}\ldots\frac{p_{m-1}}{p_m-2} < 1.$

$\because \frac{p_{i-1}}{p_i-2} \leq 1, \therefore A < 1, p_m A < p_m.$

$\because d < p_m$, if $N_e > \mathrm{p_{m+1}^2}/2$, then $d \ll N_e > \mathrm{p_{m+1}^2}/2$, $\therefore N_e = \omega(d)$. ■

produced its step1 " Order difference".

Lemma 4:

when $\mathbf{d} = p_m$, produced its Step 2 "Order difference"

Proof

Let $\prod_{i=1}^{m} \frac{p_i}{p_i-2} = \frac{p_1}{p_1-2}\frac{p_2}{p_2-2}\ldots\frac{p_{m-1}}{p_{m-1}-2}\frac{p_m}{p_m-2} = p_m$, when all $\frac{p_{i-1}}{p_i-2} = 1$ hold.

Thus: when $p_i - p_{i-1} > 2$, if $p_i - p_{i-1} = 2t, t > 1$, then in p_i, add $2t - 2$ screening elements to make $p_i - p_{i-1} = 2$, $\prod_{i=1}^{m} \frac{p_i}{p_i-\gamma_i} = \frac{p_1}{p_1-\gamma_1}\frac{p_2}{p_2-\gamma_2}\ldots\frac{p_{i-1}}{p_{i-1}-\gamma_i}\ldots\frac{p_m}{p_m-\gamma_m} = p_m$, $\gamma_i = 2t - 2$,

$\exists!\, p_i - p_{i-1} = 2, t = 2.$

Let $P_{N_e} = \{p_1, p_2, \ldots p_m\}$ be primes $<N_e$. We define:

Twin primes $\mathrm{A} = \{p_i - p_{i-1} = 2\}$

Non-twin primes $\mathrm{B} = \{p_i - p_{i-1} > 2\}$

By the Twin Prime Ratio Theorem, $|\mathrm{A}|/|\mathrm{B}| \to 0$ as $p_m \to \infty$.[3]

This means that p_i in B can increase $p_i + n_i$, $n_i = p_i - p_{i-1} \geq 2$. when $d < p_m$, $p_i + n_i$ the total number of M is: M= $\sum_{i=1}^{m} 2 = 2m$.

Let p_i in B can increase σ
$\because |A|/|B| \to 0$, $\therefore \sigma \gg M = 2m$, $\sigma = \omega(2m)$. ■

$\sigma = \omega(2m)$ indicates that if P_{N_e} is used to sieve T_m, then to maintain the average distance $d = p_m$ among the remaining elements in T_m, it is necessary to add infinitely more comb teeth to the set {Non-twin primes} of primes than the 2m comb teeth.

Thus if $d = p_m$ the **second "order difference"** has occurred. d is the average distance of the distribution of remaining elements and is an important parameter for future discussions on the distribution of remaining.

$N_e = \omega(d)$ and $\sigma = \omega(2m)$ this is an important basis for generating full-text analysis proofs, which are precise mathematical proofs and unrelated to Density probability and statistics.

Lemma 5: After P_{N_e}screening T_m, the distribution of remaining elements in each T_m/p_i cycle is basically uniformthe distribution of remaining elements.

Proof

$T_m = \prod_{i=1}^{m} p_i$, $P_{N_e} = \{p_1, p_2, \dots p_m\}$, where p_i are primes and m is sufficiently large.

After sieving out two elements from T_m, according to the Chinese remainder theorem, the remaining set G has size: $g_T = \prod_{i=1}^{m}(p_i - 2)$.
Goal: Prove that the elements of g_T are uniformly distributed in each T_m/p_i period within the range of T_m.
Inference: Because the Twin Prime Comb Sieve Method from 0 to $T_m - 1$ is the same as from $T_m - 1$ to 0 (using the same P_{N_e}criterion), g_T is symmetrically distributed about the midpoint $(T_m - 1)/2$ withinT_m. (Similarly, Goldbach Conjecture Comb Sieve Method and Barinak Conjecture Comb Sieve Method can also find symmetry points in T_m). ■

Let $P_{N_e} = \{p_1, p_2, \dots p_m\}$ be the set of the first m primes.

Let $A \subseteq \binom{P_{Ne}}{k}$, be a subset of k primes. Then there exists: $\frac{T_m}{A} = T_{m-t}$, $1 \leq t < k$.

For $t = 1$ take $A = \{p_i\}$ then: $T_{m-1} = T_m/p_i$,

Consider two distinct primes $p_i$and p_{i+n}, and their corresponding quotients:

Tmpiand $\frac{T_m}{p_i}$ and $\frac{T_m}{p_{i+n}}$.

Since T_m is a product of primes, its structure is periodic. In the residue classes modulo p_i or p_i, each class appears uniformly. Therefore, after sieving out one prime, the remaining elements are uniformly distributed in T_m. Due to the symmetry and periodicity of the residue classes, the distributions of remaining elements in the quotient structures corresponding to different primes are essentially the same.

For example: consider the structure modulo $\frac{T_m}{3}$. In a segment T_m, there are three periods of length $\frac{T_m}{3}$. The distribution of remaining elements within each such period is approximately uniform. Similarly, for a general prime $\frac{T_m}{p_i}$, there exist analogous $\frac{T_m}{p_i}$ periods, and the distribution within each period is also nearly uniform.

If we add the inference that the $(\frac{T_m}{p_i} - 1)/2$-symmetry points of each $\frac{T_m}{p_i}$ period are added, then the uniform distribution of g_T in T_m becomes even more apparent.

Because these various $\frac{T_m}{p_i}$ periods are interwoven throughout T_m, the remaining elements in T_m as a whole do not exhibit extreme clustering or significant deviation from a uniform distribution.

For example: $N_e = \omega(d)$, if $d \geq N_e$, exhibit extreme clustering or significant deviation from a uniform distribution. In 5, We will prove that this distribution is impossible.

For $t > 1$:

Expanding to the case of filtering multiple prime numbers, T_m contains numerous $\frac{T_m}{p_{i+n}}$ periods that are intertwined with each other. According to the Chinese remainder theorem, the distribution of residual elements in each $\frac{T_m}{p_{i+n}}$ period is basically the same. The Chinese remainder theorem ensures the independence and symmetry of the residual class structure of different prime numbers. Therefore, the remaining set G after screening multiple prime numbers is uniformly distributed in any of the same $\frac{T_m}{p_{i+n}}$ periods in T_m.■

From the above proof, it can be seen that after sieving $T_m$with respect to the primes p_i, the distances between the two remaining elements in T_m are not isolated; they are intertwined with the numerous T_{m-i} periods in T_m. For example, if a spacing of 0 to K appears in the $\frac{T_m}{p_i}$ period, then according to the Chinese Remainder Theorem, there are also $p_i - 2$ instances of the K-spacing in T_m. According to the inference, we can also find a large number of symmetric points with a distance of K.

The step1 and step2 order differences, together with Lemma 5, will provide us with a powerful toolkit for proving the propositions. Let's demonstrate the practical application of comb screening method through a specific example:

Let $P_{N_e} = \{p_1, p_2, p_3, p_4,\}$ $p_1 = 3, p_2 = 5, p_3 = 7, p_m = p_4 = 11, T_m = T_4 = p_1p_2p_3p_4 = 1155.$

$p_1p_2p_3p_4 = T_4,\ g_T = (p_1 - 2)(p_2 - 2)(p_3 - 2)(p_4 - 2) = 158$

$d = \frac{1155}{158} \approx 7.31$, if $d = 11$ then $p_4 \to 11 - 2 \to 11 - 4$

$g_T = (p_1 - 2)(p_2 - 2)(p_3 - 2)(p_4 - 4) = 105,\ d = \frac{1155}{105} = 11 = p_4$

$N_e = \frac{p_m^2 - 1}{2} = \frac{11^2 - 1}{2} = 60,$

When $\mathrm{d} < p_m$, $\mathrm{d} \ll \frac{p_m^2-1}{2}$. When $\mathrm{d} = p_m$, $\mathrm{d} \ll \frac{p_m^2-1}{2} = N_e$.

Obviously, if $p_m \to \infty$, $\frac{|A|}{|B|} \to 0$ to make $\mathrm{d} = p_m$ must increase σ and $\sigma = \omega(2m)$.

This means that on the basis of the original$\{P_{N_e}\}$, we add infinitely many identical$\{P_{N_e}\}$, and after screening T with the$\{P_{N_e}\}$, the average distance between the remaining elements is $d = p_m \ll N_e$

Because after P_{N_e}screening T_m, the distribution of remaining elements in each T_m/p_i cycle is basically uniformthe distribution of remaining elements , So it is impossible to generate $d \geq N_e$.

Although the step2-order difference produces an obvious result, this does not constitute a rigorous mathematical proof. The proof based on step1-order differences, however, provides insight into the variation patterns of the remaining elements within T_m. Below, we will employ step1-order differences to prove this proposition in a rigorous mathematical manner.

5. Prove the infinite existence of twin prime numbers using the concepts of distance and area in geometry.

Because the distribution of g_T is roughly uniform within a single period of $\frac{T_m}{p_i}$, so $\mathrm{d} \geq N_e$ is contradictory , this chapter analyzes the distribution of g_T inT_m to solve the contradiction.

Because there are phenomena of different orders in the distribution of remaining elements, so in the proof that follows, we will introduce three major approximations (During the proof process, each error is clearly indicated).

The error from these approximations will always be chosen to favor the statement we are proving. As some of the numerical bounds are deliberately loose, the reader need not concern themselves with optimizing these precision.

.**Lemma 6:** After using $P_{N_e}$to comb T_m, denote the number of remainder elements as $\mathrm{g_T}$, then when $m \to \infty, \exists \mathrm{j} < \mathrm{m}$, satisfying $(p_j^2 - 1)/2 \geq 2T_m/g_T$, and $m - \mathrm{j} \to \infty$

Proof.

$d = T_m/g_T < p_m$, (error 1: neglected when $\mathrm{d} = p_m$, $\sigma = \omega(2m)$)

let $\mathrm{p_d}$ *be the largest prime number less than* $\sqrt{4\mathrm{p}_m + 1}$*, that is*

$p_d < \sqrt{4p_m + 1}$ *and* $\mathrm{p}_{d+1} \geq \sqrt{4p_m + 1}$*, it is well known that the number of prime numbers less than* p_m

is $\pi(\mathrm{p}_m) \approx \mathrm{p}_m/log\mathrm{p}_m$, [1] *the number of prime numbers less than* p_d
is $\pi(\mathrm{p}_d) \approx \mathrm{p}_d/log\mathrm{p}_d$*. Obviously, when* $m \to \infty, \pi(\mathrm{p}_m) \gg \pi(\mathrm{p}_d)$, [2]

$let\ j = d + 1,\ then\ (p_j^2 - 1)/2 \geq 2p_m \geq 2T_m/g_T\ \ and\ m - j \to \infty$ …… (2) ■

Because after P screening T_m , the distribution of remaining elements in T_m is uniform.

SO we now come to analyze the changes of remainder elements' distributions within $0{\sim}T_m$ when it is combed by elements in set $\{p_1, p_2, p_3, \dots p_i \dots p_m\}$.

When $0{\sim}T_m$ hasn't been combed, element pairs within $0{\sim}T_m$ whose distance equals to 1 form the set.

$\mathrm{D(d = 1)} = \{(0,1),\ (1,2),\ (2,3),\ \dots (T_m - 1), T_m)\}$.The total number of these pairs is T_m; element pairs whose distance equals to 2 form the set

$\mathrm{D(d = 2)} = \{(0,2),\ (1,3),\ (2,4),\ \dots (T_m - 2), T_m)\}$ .The total number of these pairs is $T_m - 1$; element pairs whose distance equals to 3 form the set

$\mathrm{D(d = 3)} = \{(0,3),\ (1,4),\ (2,5),\ \dots (T_m - 3), T_m)\}$ .The total number of these pairs is $T_m - 2$; ……element pairs whose distance equals to K form the set

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

When using comb $p_1 = 3$ to comb out those points within $0{\sim}T_m$ 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 $0{\sim}T_m$ is $T_m/3 = p_2 p_3 \dots p_m = \prod_{i=2}^{m} p_i$, and the distance of any two remainder elements is $3n$, such as element pairs with distance 3 form the set

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

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

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

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

$$\mathrm{D(d = 3n)} = \{(0,3\mathrm{n}),\ (3,3\mathrm{n} + 3),\ (6,3\mathrm{n} + 6),\ \dots (T_m - 3\mathrm{n}), T_m)\}.$$

the total number of such pairs is$(T_m/3) - n + 1$;

Definition 3 : Pure Distance Pairs and Mixed Distance Pairs

Let the distance between two remaining elements be K, $\forall \mathrm{D(d = k)}$, 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))
$\forall \mathrm{D(d = K)}$, 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.$g(D(d = K)$

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

Obviously, $g(D(d = K)) > g(Z(d = K))$. (error 2: neglected $\mathrm{g}(c(d = k))$) …… (3)

We only discuss the distribution of pure distance pairs here.

We know that after $0\sim T_m$ is combed by $p_1$and p_2, there are only two different pure distance pairs within $0\sim T_m$, 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 $0\sim T_m$ are $0,6,9,15,21,24,30\ldots 15n, 15n+6, 15n+9$ ....The total number of such elements is $(p_1-2)(p_2-2)\,p_3\ldots\,p_m$. after $0\sim T_m$ 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 $0\sim p_1\,p_2\,p_3=105$,

$$g\big(c(d=3)\big)=3,\ g\big(c(d=6)\big)=8,\ g\big(c(d=9)\big)=2,\ g\big(c(d=15)\big)=2$$

and within $0\sim T$, $g\big(c(d=3)\big)=3p_4p_5\ldots p_m$, $g\big(c(d=6)\big)=8p_4p_5\ldots p_m$

$$g\big(c(d=9)\big)=2p_4p_5\ldots p_m,\ g\big(c(d=15)\big)=2p_4p_5\ldots p_m$$

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

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

$3\times g\big(c(d=3)\big)+6\times g\big(c(d=6)\big)+\cdots+mc\times g\big(c(d=mc)\big)=T_m$ and the average distance $d=T_m/g_T$.

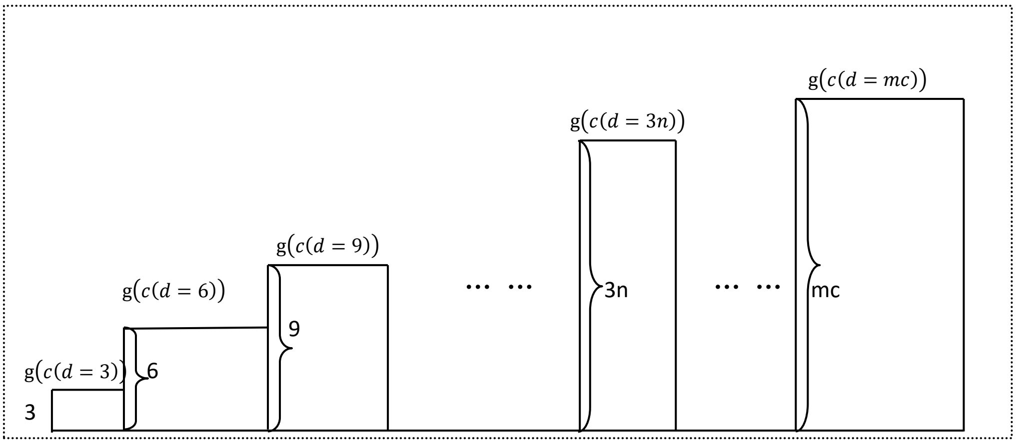

Fig6.

Fig6 the distribution of pure distance pairs within $0\sim T_m$ after combed by $p_1, p_2, \ldots, p_K$

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

Lemma 7:

$\forall C(d=k_a), \exists g(C(d=k_a))$, settle for $g\big(C(d=k_a)\big)\le g\big(D(d=k_a)\big)$

$$g\big(D(d=k_a)\big)=(p_1-\gamma_1)(p_2-\gamma_2)\ldots(p_i-\gamma_i)(p_{i+1}-4)\ldots(p_m-4)$$
$$=\prod_{i=1}^{e}(p_i-\gamma_i)\prod_{i=e+1}^{m}(p_i-4)\qquad (p_{e+1}>k_a+1)\qquad \ldots\ldots(3)$$

When $p_i \le k_{i+1}$, $2 \le \gamma_i \le 4$. When $p_i > k_i + 1, \gamma = 4$.

The proof in Chapter 7 relies only on the result $\gamma = 4$ from Lemma 7, ignoring $\gamma = 2$ and $\gamma = 3$. Consequently, the purpose of Lemma 7 is merely to assert the existence of $\gamma = 2$ and $\gamma = 3$.for completeness, a step that is logically incidental to the main proof's rigor. Thus, readers can omit reading through its proof steps.

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* $0\sim T_m$ *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* $0\sim T_m$*,*

set $E = \{(np_1\, p_2\, p_3 \ldots\, p_i + a, np_1\, p_2\, p_3 \ldots\, p_i + a + k) | n = 0,1,2, \ldots\}$.
$\because p_1\, p_2\, p_3 \ldots\, p_i + a \equiv a \bmod p_{i+1}$ *, and* $p_1\, p_2\, p_3 \ldots\, p_i + a + k \equiv (a+k) \bmod p_{i+1}$*,*
$\therefore 2 \le \gamma_i \le 4$*, 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* 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* 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* 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* 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* p_{i+1}*'s tooth .*
here $\gamma = 3$ *in the same way ,when* $k = n\, p_{i+1} - 1$*,* $\gamma = 3$*.*

(C) *When* $k \ne n\, p_{i+1}, k \ne n\, p_{i+1} \pm 1$ *or* $np_{i+1} > k + 1$*,*
since $n_u p_1\, p_2\, p_3 \ldots\, p_i + a$ *and* $n_u p_1\, p_2\, p_3 \ldots\, p_i + a + k$ *is no at one time combed out by comb* p_{i+1}*'s tooth , here* $\gamma = 4$*. Inductively, equation (3) is proved.* ■

6. 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 $\{p_1\, p_2\, p_3 \ldots\, p_z\, p_{z+1}\, p_{z+2} \ldots\, p_u\}$ is combed,
$0\sim (p_u^2 - 1)/2$, $0\sim (p_u^2 - 1)/2$ there is no remainder. [lemma2]

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

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

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

$\because (p_u^2 - 1)/2 \ge 2T_v/g_{T_v}$, $(Z + 1 < u < v)$, $\therefore (v - u)/u = n$, $v \gg u$. [lemma6]

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 7, and its area is T_v.

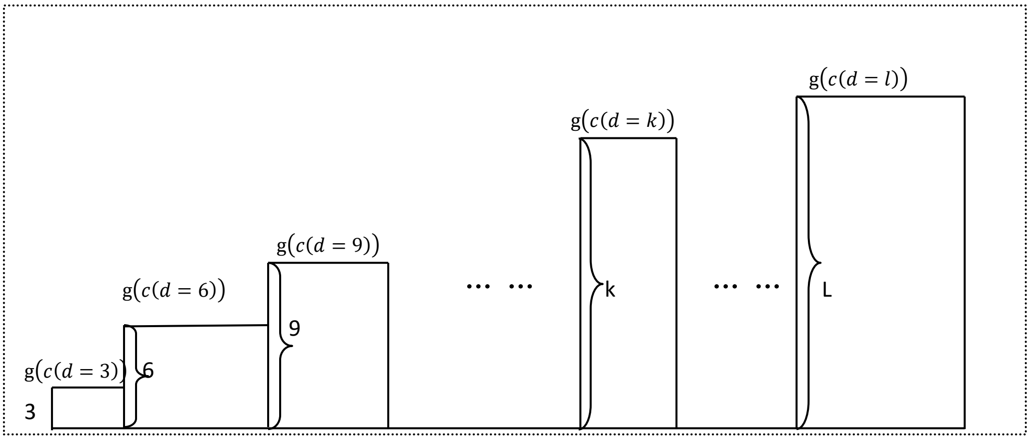

Fig7.

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 8, and its area is T_v

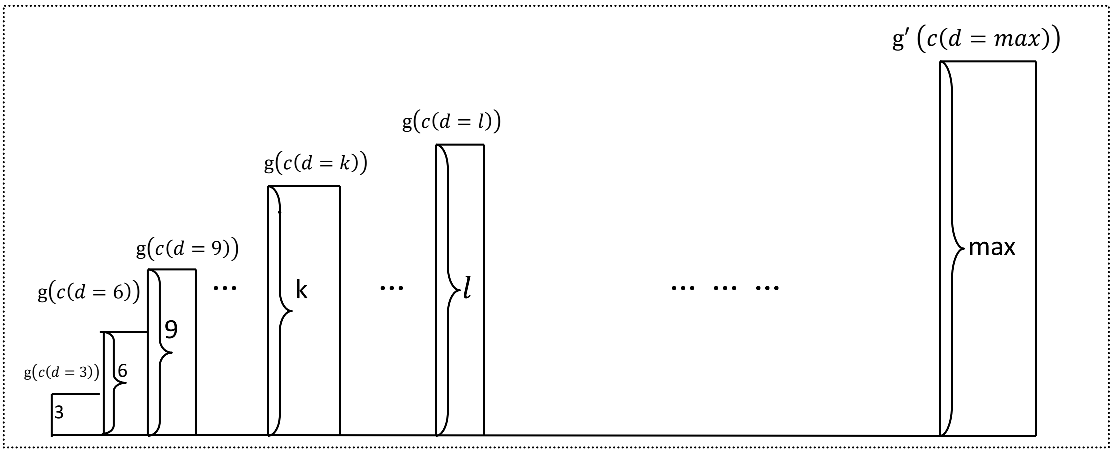

Fig8.

In Fig 8, 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 7 and Figure 8, 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 $M_c = g(C(d = k_c))$, 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* $g'(c(d=k_i))$ *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*

$g'(c(d=k_a)) \le \prod_{i=1}^{e}(p_i - \gamma_i) \prod_{i=e+1}^{v}(p_i - 4)$. [lemma7] *We have:*

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

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

Because $v \gg e$, [lemma4] so *in* (4), *we only need to analyze* $\frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)}$,

$\because \frac{P_i}{P_i-4} > 1, \therefore \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)} \xrightarrow{V\to\infty} \infty$. (error 3: neglected$(p_i-3)$ *and* (p_i-2))

$\because \frac{v-u}{u} = n, v \to \infty, n \to \infty$, [lemma6] $\therefore \frac{M_c}{\prod_{i=1}^{e}(p_i-\gamma_i)} \frac{\prod_{i=u+1}^{v} p_i}{\prod_{i=e+1}^{v}(p_i-4)} = n$.

n *increases with the increase of* u, v. ■

Fig. 9 is pure distance distribution after $0 \sim T_v$ *is combed by* $p_1, p_2, p_3 \dots p_u \dots p_v$

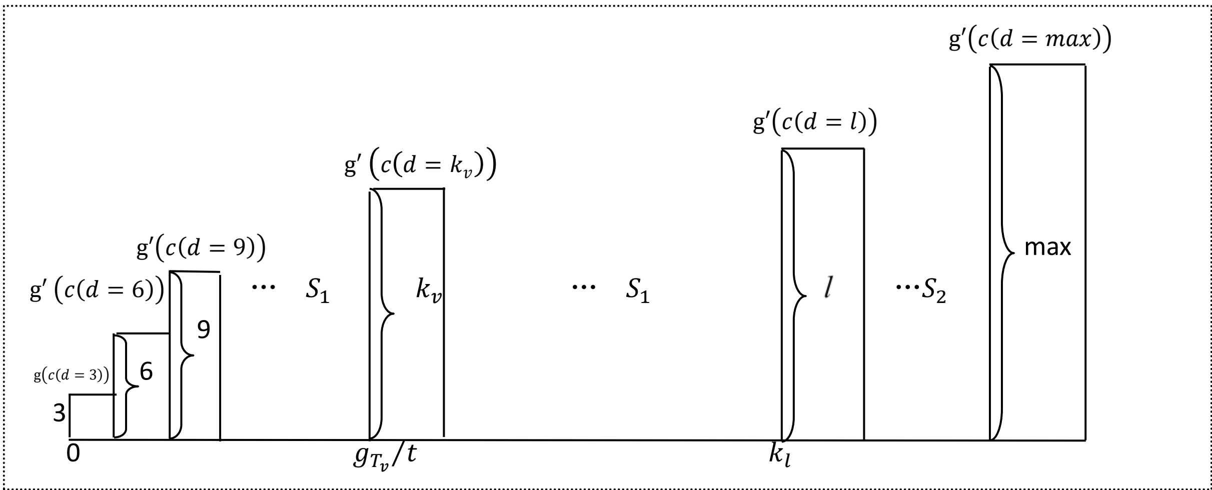

fig9

In fig9 S_1 is an area map made up of pure distance $g'(c(d \le L))$

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

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

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

We can gat $\frac{S_2}{S_1} \xrightarrow{v\to\infty} \infty$, that is $S_2 \gg S_1$. [lemma6]

In fig9, Point k_l corresponds to $d = l$, $\because l \geq 2T_v/g_{T_v}$, $\therefore$ 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}/t it's a point in the S_1 area, $t > 2$, the pure distance corresponding to point g_{T_v}/t is k_v, set $k_v = \frac{T_v}{g_{T_V}}\beta, 0 < \beta < 1$,

If the average number of remaining elements in $g(c(d = 3)) \sim g(c(d = k_v))$ is g_v, then

$g_v\frac{k_v}{3} = \sum_{i=3t}^{k_v} g(c(d = i))$ that is $k_v = \frac{3}{g_v}\sum_{i=3t}^{k_v} g(c(d = i))$,

$\because k_v = \frac{T_v}{g_{T_v}}\beta, \therefore \frac{3}{g_v}\sum_{i=3t}^{k_v} g(c(d = i)) = \frac{T_v}{g_{T_v}}\beta$ have:

$$\beta = \frac{3g_{T_v}}{T_v g_v}\sum_{i=3t}^{k_v} g(c(d = i)) \qquad \dots \dots 5$$

Comparison between (4) and (5), (4) 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) $\xrightarrow{v\to\infty} \infty$, so must have (5) $\xrightarrow{v\to\infty}$ 0.

Proof

$k_v = \frac{T_v}{g_{T_v}}$ If β is not convergent and tends to be zero, then $k_v \xrightarrow{v\to\infty} \epsilon$,

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

S_ϵ increases with the increase of V, make $\frac{S_2}{S_1} \xrightarrow{v\to\infty} \infty$ not tenable, so (5) $\xrightarrow{v\to\infty}$ 0. ■

In (5), although there is $\frac{g_{T_v}}{T_v} \xrightarrow{v\to\infty} 0$, $\sum_{i=3t}^{k_v} g(c(d = i)) > g_v$ is divergent, (5) is $\infty \times 0$ series, no definite trend,

That is: $\frac{1}{g_v}\sum_{i=3t}^{k_v} g(c(d = i)) = \frac{k_v - 3}{3} g_v$

obviously, $\frac{k_v - 3}{3} g_v$ increases with the increase of v.

In(5), although there is $\frac{g_{T_v}}{T_v} = \frac{(p_1 - 2)(p_2 - 2)(p_3 - 2)\dots(p_v - 2)}{p_1 p_2 p_3 \cdots p_v} \xrightarrow{v\to\infty} 0$

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. ■

The proposition has been proven!

follow-up:

7. The prospect of using step 2 "step difference" to solve such

This section only provides a rough overview of lemma4, without delving into in-depth proofs，Only as a prospect introduction for using "Comb Sieve Method," to solve problems related to prime number distribution.

In Lemma 3:

$$d = \prod_{i=1}^{m} \frac{p_i}{p_i-2} = \frac{p_1}{p_1-2}\frac{p_2}{p_2-2}\cdots\frac{p_{m-1}}{p_{m-1}-2}\frac{p_m}{p_m-2} < p_m$$

$$Let\ \prod_{i=1}^{m} \frac{p_i}{p_i-2} = \frac{p_1}{p_1-2}\frac{p_2}{p_2-2}\cdots\frac{p_{m-1}}{p_{m-1}-2}\frac{p_m}{p_m-2} = p_m$$

when all $\frac{p_{i-1}}{p_i-2} = 1$ hold, [lemma4]

In lemma4, $|A|/|B| \to 0$ Indicating sparse twin prime numbers, means that d equals pm only when the number of teeth added to the non twin prime set is infinitely many times greater than the original number of teeth(compared to a quantity of 2M comb teeth).

Lemma 5: After P screening T, the distribution of remaining elements in T is uniform If $d \geq p_{m/2}^2$, the twin prime density will be abnormally high, This is contradictory to $|A|/|B| \to 0$.

There is also key evidence:

If Twin prime numbers are finite, in$T_{\mathbf{m}}$, $T_{\textbf{Twin primes}} = \prod_{i=1}^{z} p_i\, p_{i-1} = C, p_i - p_{i-1} = 2$ is constant.

If $T_{\textbf{Twin primes}}$ is a constant, an infinite increase in ti will lead to a series of irreconcilable contradictions in subsequent transformations. We will use mathematical logic to prove the irreconcilability of these contradictions. However, since the first step of the proof of order in this article has already produced results consistent with the double prime conjecture, the second step of the proof of order is mainly to prepare for conjectures related to such problems, such as the Mason conjecture. This article will not delve into this aspect in depth for now.

Since the step 1"step difference" has already been used to prove the prime prime conjecture, it can likewise be employed to prove the Barinak conjecture and the Goldbach conjecture. By combining the first and second step differences, proving conjectures such as the prime prime conjecture becomes straightforward. The author is currently applying this method to address the Mason conjecture. Should this paper be accepted by a journal, the author will subsequently submission a paper demonstrating the infinitude of Mason primes.

1. The authors declare that there are no conflicts of interest regarding the publication of this paper.
2. All references cited in this paper are from publicly available online sources.

REFERENCES

[1] Green & Tao (2008). The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics.*

[2] Halberstam & Richert (1974). Sieve Methods. *Academic Press*.

[3] Hardy, G.H. & Littlewood, J.E. (1923). Some problems of ‘Partitio numerorum’; III: On the expression of a number as a sum of primes. *Acta Math*.

[4] Zhang, Y. (2014). Bounded gaps between primes. *Annals of Mathematics.*

[5] Maynard, J. (2015). Small gaps between primes. *Inventiones Mathematicae.*

[6]Tao, T. (2015). Polymath and the twin prime conjecture. *arXiv preprint.*

[7] Chen, J.R. (1973). *Acta Mathematica Sinica*