Goldbach's Conjecture II: Chasing Goldbach on the Circle

Yangkun Li

Summer 2025

1 Introduction

Some problems in mathematics are deceptively simple yet remarkably enduring, and Goldbach's conjecture is one of the most well-known. It states that **every even number no less than 6 can be written as the sum of two prime numbers.** It is also proposed that **every odd number no less than 9 can be written as the sum of three prime numbers**, which is now know as Goldbach's weak conjecture. Despite centuries of effort and many numerical evidence, a complete proof remains out of reach[1]. In my article from the previous issue of this magazine [1], we explored the conjecture's historical development and highlighted some of the key breakthroughs along the way. These seemingly irregular primes consistently sum to even numbers calls for deeper analytical insight. To that end, we now turn to one of the most influential analytic technique developed in the 20th century.

2 Goldbach's conjecture as an integral

In 1920, Hardy and Littlewood published a series of papers under the general title Some Problems of "Partitio Numerorum" in which they systematically created and developed an analytic method in additive number theory. Among them, the third paper published in 1923 were specifically devoted to discussing Goldbach's conjecture [4]. The core ideas behind this new method had already appeared in a 1918 paper by Hardy and Ramanujan, which were later referred to as the Hardy–Littlewood–Ramanujan *circle method* [3]. Building on this foundation, Soviet mathematician Ivan Vinogradov made significant improvements to the method leading to major breakthroughs in expressing odd numbers as sums of primes.

Generally speaking, the circle method turns the conjecture into studying an integral along the unit circle. Consider the equations

$$N = p_1 + p_2, \quad p_1, p_2 \text{ are odd primes.} \tag{1}$$

$$N = p_1 + p_2 + p_3, \quad p_1, p_2, p_3 \text{ are odd primes.} \tag{2}$$

Given N, we can solve these equations for p_1, p_2 and p_3. Then to prove Goldbach's conjecture, it is to prove that equation (1) has solutions for all even $N \geq 6$. To prove Goldbach's weak conjecture, it is to prove that equation (2) has solutions for all odd $N \geq 9$. It turns out that the number of solutions can be written in an interesting form.

Theorem 1. *The number of solutions to equation (1),*

$$D(N) = \int_0^1 S^2(\alpha, N)e(-N\alpha) \, \mathrm{d}\alpha, \tag{3}$$

and the number of solutions to equation (2)

$$T(N) = \int_0^1 S^3(\alpha, N)e(-N\alpha) \, \mathrm{d}\alpha \tag{4}$$

---

[1]Peruvian mathematician Harald Helfgott released a paper in 2013, which is now widely accepted as a proof of Goldbach's weak conjecture. [5]

where $e(x) = e^{2\pi i x}$, $S(\alpha, N) = \sum_{2 < p \leq N} e(\alpha p)$, *p is a prime.*

Proof. By direct expansion we have

$$\int_0^1 S^2(\alpha, N) e(-N\alpha)\, \mathrm{d}\alpha = \int_0^1 \left(\sum_{2 < p \leq N} e(\alpha p) \right)^2 e(-N\alpha)\, \mathrm{d}\alpha$$

$$= \int_0^1 \left(\sum_{2 < p_1, p_2 \leq N} e(\alpha(p_1 + p_2)) \right) e(-N\alpha)\, \mathrm{d}\alpha$$

$$= \int_0^1 \left(\sum_{2 < p_1, p_2 \leq N} e(\alpha(p_1 + p_2 - N)) \right) \mathrm{d}\alpha$$

$$= \sum_{2 < p_1, p_2 \leq N} \int_0^1 e(\alpha(p_1 + p_2 - N))\, \mathrm{d}\alpha \qquad (*)$$

Recall that for any $n \in \mathbb{Z}$ we have the orthogonality relation

$$\int_0^1 e(n\alpha)\, \mathrm{d}\alpha = \begin{cases} 1, & \text{if } n = 0 \\ 0, & \text{if } n \neq 0 \end{cases} \qquad (5)$$

Every integral in $(*)$ evaluates to 0 if $N \neq p_1 + p_2$, and to 1 if $N = p_1 + p_2$, so the integral in (3) gives the number of pairs of primes $\{p_1, p_2\}$ that solves (1). Similarly, we can show that the integral in (4) gives the number of solutions to equation (2). $\qquad \square$

We can restate the conjectures as follows.

Conjecture 2 (Goldbach's conjecture). *For all even $N \geq 6$, $D(N) > 0$.*

Conjecture 3 (Goldbach's weak conjecture). *For all odd $N \geq 9$, $T(N) > 0$.*

Hence, the problem of both conjectures is reduced to the analysis of the integrals in (3) and (4). Of course, that would require us to study the trigonometric sum $S(\alpha, N)$.

3   Major and minor arcs

3.1   Construction

Hardy and Littlewood believed that on certain small intervals (centered at irreducible fractions[2]), $S(\alpha, N)$ attains larger values, meaning that most of the contributions to $D(N)$ and $T(N)$ come from these intervals. Outside of them, $S(\alpha, N)$ are comparatively small and thus its contribution to the integrals is insignificant. Based on this, they proposed dividing the interval of integration into what we call the major and minor arcs[2].

Let us introduce some technical details to make the idea precise. Let $Q, r > 0$ with $1 \leq Q \leq r \leq N$. Consider the Farey sequence[3] of order Q,

$$F_Q = \left\{ \frac{a}{b} \ \middle| \ 0 \leq a \leq b, \ 1 \leq b \leq Q, \ \gcd(a, b) = 1 \right\},$$

and the corresponding set of intervals centered at $\dfrac{a}{b}$,

$$I(a, b) = \left[\frac{a}{b} - \frac{1}{r}, \frac{a}{b} + \frac{1}{r} \right], \quad a < b.^4$$

---

[2]We will define these terms more precisely later.

[3]A Farey sequence of order $n \in \mathbb{Z}^+$ is the set of irreducible fractions $\frac{a}{b}$ arranged in increasing order, with $0 \leq a \leq b$ and $1 \leq b \leq n$. We use Farey sequence here is because we want to construct intervals centered at irreducible fractions.

These intervals will be used to construct major and minor arcs mentioned earlier. We put a condition on the choice of Q and r to give $I(a, b)$ some nice properties.

Theorem 4. *If $2Q^2 < r$, then every $I(a, b)$ lies strictly inside $\left[-\dfrac{1}{r}, 1 - \dfrac{1}{r}\right]$ and they are pairwise disjoint.*

Proof. If $2Q^2 < r$ then $\dfrac{1}{Q^2} > \dfrac{2}{r}$. For any two distinct terms $\dfrac{a_1}{b_1}, \dfrac{a_2}{b_2} \in F_Q$, we have

$$\left|\frac{a_1}{b_1} - \frac{a_2}{b_2}\right| \geq \frac{1}{b_1 b_2} \geq \frac{1}{Q^2} > \frac{2}{r}.^5$$

By definition, the length of each $I(a, b)$ is $\frac{2}{r}$, the above inequality implies that they do not intersect with each other. Now by assumption $Q \geq 1$, so $2Q \leq 2Q^2 < r$. We can simply check that the endpoints of the "rightmost" and "leftmost" interval lies inside $\left[-\frac{1}{r}, 1 - \frac{1}{r}\right]$:

$$-\frac{1}{r} < \frac{1}{Q} - \frac{1}{r} \leq \frac{a}{b} - \frac{1}{r}, \quad \text{and} \quad \frac{a}{b} + \frac{1}{r} \leq \frac{Q-1}{Q} - \frac{1}{r} < 1 - \frac{1}{r}.$$

So, every $I(a, b)$ lies strictly inside $\left[-\dfrac{1}{r}, 1 - \dfrac{1}{r}\right]$. $\qquad\square$

Definition 1 (Major and minor arcs[6]). *The **major arcs is** the set*

$$E_1 = \bigcup_{1 \leq b \leq Q} \bigcup_{\substack{0 \leq a < b \\ \gcd(a,b)=1}} I(a, b),$$

*each $I(a, b)$ is called a **major arc**. **Minor arcs** is the set*

$$E_2 = \left[-\frac{1}{r}, 1 - \frac{1}{r}\right] \setminus E_1.$$

*It is easy to see that E_2 is also a (disjoint) union of intervals. Each of these intervals is a **minor arc**.*

We have divided the interval $\left[-\frac{1}{r}, 1 - \frac{1}{r}\right]$ into two parts, E_1 and E_2, as proposed in the circle method. Note that $\left[-\frac{1}{r}, 1 - \frac{1}{r}\right]$ is not the integration range $[0, 1]$ of $D(N)$ and $T(N)$. However, this is not an issue because both $S(\alpha, N)$ and the integrands of $D(N)$ and $T(N)$ are periodic functions with period 1, so the integrals over both intervals yield the same value. We can use the one that is more convenient depending on the context.

Definition 2. *The denominator of a fraction is **small** if it is no more than Q, and **large** otherwise. Two points are said to be **close** if they are no more than r^{-1} apart.*

We previously noted that $S(\alpha, N)$ tends to be larger on certain intervals. In fact, Hardy and Littlewood predicted (and later proved by Vinogradov) that $S(\alpha, N)$ is larger when α is close to an irreducible fraction with a small denominator, and smaller when α is close to one with a large denominator. The following theorem offers an intuitive justification for the way we defined the major and minor arcs.

Theorem 5. *Every point in E_1 is close to an irreducible fraction with small denominator; Every point in E_2 is close to an irreducible fraction with large denominator.*

---

[4]Strictly speaking, the definition of the Farey sequence allows $a = b$, but we omit this case for technical convenience, as it has no effect on the analysis of the integral.

[5]A Farey sequence has the property that the distance between any two distinct terms $\left|\frac{a_1}{b_1} - \frac{a_2}{b_2}\right| \geq \frac{1}{b_1 b_2}$.

[6]Major arcs are sometimes called basic intervals, and minor arcs, supplementary intervals.

Proof. Let $e_1 \in E_1$. By definition $e_1 \in \left[\frac{a}{b} - \frac{1}{r}, \frac{a}{b} + \frac{1}{r}\right]$, where $\frac{a}{b} \in F_Q$. Clearly, e_1 is close to $\frac{a}{b}$, an irreducible fraction with small denominator ($b \leq Q$). Now let $e_2 \in E_2$. By Dirichlet's approximation theorem[7], there exists $c, d \in \mathbb{Z}$ satisfying $\gcd(c, d) = 1$, $1 \leq d \leq r$ such that

$$|de_2 - c| < \frac{1}{r}, \quad \text{or equivalently,} \quad \left|e_2 - \frac{c}{d}\right| < \frac{1}{dr} \leq \frac{1}{r}.$$

Thus, every point in E_2 is close to an irreducible fraction $\frac{c}{d}$. Here, if $d \leq Q$ then $\frac{c}{d} \in F_Q$ and the above inequality would imply $e_2 \in I(c, d) \subseteq E_1$. But since $E_1 \cap E_2 = \emptyset$ by definition, this leads to a contradiction. Therefore, $d > Q$ and $\frac{c}{d}$ must have a large denominator. $\square$

Theorem 5 shows that the major arcs E_1 are precisely the intervals where α is close to irreducible fractions with small denominators, and the minor arcs E_2 correspond to those close to irreducible fractions with large denominators, exactly capturing the structure Hardy and Littlewood had in mind. We can now analyze $D(N)$ and $T(N)$ on E_1 and E_2 separately.

$$D(N) = \int_{-\frac{1}{r}}^{1-\frac{1}{r}} S^2(\alpha, N)e(-N\alpha) \, d\alpha = D_1(N) + D_2(N),$$

where

$$D_1(N) = \int_{E_1} S^2(\alpha, N)e(-N\alpha) \, d\alpha, \quad D_2(N) = \int_{E_2} S^2(\alpha, N)e(-N\alpha) \, d\alpha.$$

Similarly,

$$T(N) = \int_{-\frac{1}{r}}^{1-\frac{1}{r}} S^3(\alpha, N)e(-N\alpha) \, d\alpha = T_1(N) + T_2(N),$$

where $T_1(N)$, $T_2(N)$ are defined analogously.

3.2 Some remarks

The above concludes the setup for the circle method. Our goal is to compute $T_1(N)$ and $D_1(N)$, and show that they are the principal terms of $T(N)$ and $D(N)$, with $T_2(N)$ and $D_2(N)$ being of lower order, respectively.

As discussed in [1], proving both conjectures directly is hard, but their weaker forms are approachable: we do not need to prove the conjectures hold for all N, but only that they hold beyond some large N of our choice, which proves the conjectures for all but finitely many cases. It suffices to then verify the conjecture for the numbers below that threshold.

Why is this important? Recall that the construction of major and minor arcs depends on the choice of Q and r. In particular, Theorem 4 imposes the condition $2Q^2 < r$. If we work with the weaker versions of the conjectures, then we are free to choose N to be as large as we want. Under our assumption $1 \leq Q \leq r \leq N$, Q and r can also be as large as we want, so the condition in Theorem 4 can be easily satisfied. Theorem 4 becomes relevant later, as we will see.

4 Analysis of $T(N)$ and $D(N)$

Estimating $T(N)$ and $D(N)$ is no easy task. It involves long chains of complex lemmas and theorems. Including all of them would make the article far too long, while including only a subset would complicate the things without added insight. For this reason, we omit the detailed proofs and instead focus on explaining the ideas and structure of the arguments. We will however give references to theorems and their proofs.

---

[7]A full proof can be found on page 90 of [3]. Some versions of the theorem does give a coprime pair, but diving two numbers by their GCD always makes them coprime.

4.1 Singular Series

Definition 3 (Singular Series[8]). *The singular series for Goldbach's conjecture is*

$$\mathfrak{S}_2(N) = \prod_{p \nmid N} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p \mid N} \left(1 + \frac{1}{p-1}\right)$$

The sigular series for Goldbach's weak conjecture is

$$\mathfrak{S}_3(N) = \prod_{p \mid N} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{p \nmid N} \left(1 + \frac{1}{(p-1)^3}\right).$$

These terms have a nice property.

Theorem 6. *For sufficiently large positive integer N,*

$$\mathfrak{S}_2(N) \begin{cases} = 0, & \text{if } N \text{ is odd,} \\ > \frac{1}{2}, & \text{if } N \text{ is even,} \end{cases} \quad , \text{and} \quad \mathfrak{S}_3(N) \begin{cases} = 0, & \text{if } N \text{ is even,} \\ > \frac{1}{2}, & \text{if } N \text{ is odd,} \end{cases}$$

Proof. Clearly, when N is odd/even for $\mathfrak{S}_2(N)$, $\mathfrak{S}_3(N) = 0$ respectively. Next we notice that when N is even,

$$\mathfrak{S}_2(N) > \prod_{p \nmid N} \left(1 - \frac{1}{(p-1)^2}\right) > \prod_{n \geq 3} \left(1 - \frac{1}{(n-1)^2}\right) = \frac{1}{2}.$$

and that when N is odd,

$$\mathfrak{S}_3(N) > \prod_{p \mid N} \left(1 - \frac{1}{(p-1)^2}\right) > \prod_{n \geq 3} \left(1 - \frac{1}{(n-1)^2}\right) = \frac{1}{2}.$$

$\square$

As we will see later, $\mathfrak{S}_2(N)$ and $\mathfrak{S}_3(N)$ act as "multipliers" in the estimation of $D(N)$ and $T(N)$, filtering out values of N that fall outside our scope. For Goldbach's conjecture, we are only interested in even N. If $D(N)$ is given an odd N, then $\mathfrak{S}_2(N) = 0$, making $D_1(N)$ nearly zero. Likewise, in Goldbach's weak conjecture, we consider only odd numbers. $\mathfrak{S}_3(N) = 0$ for even N, rendering $T_1(N)$ negligible.

4.2 Estimation of $T_1(N)$ and $T_2(N)$

We begin by stating a weaker form of Goldbach's weak conjecture.

Theorem 7 (Vinogradov's theorem[9]). *For sufficiently large even N, $T(N) > 0$.*

Proof idea. To apply the circle method, we choose $Q = \log^\lambda(N)$ and $r = N \log^{-\gamma}(N)$ where λ and γ are appropriate positive real numbers depending on the approximation methods we use. Note that for sufficiently large N, the condition $2Q^2 < r$ is satisfied regardless of the choice of λ and γ, so we can always find appropriate major and minor arcs.

There are several ways to bound $T_1(N)$. Two proofs, in [2] and [3], both employ Siegel–Walfisz theorem. Section 6.3 of [3] offers another proof that uses Page's theorem. The process of bounding $T_1(N)$ is highly intricate, so we shall omit the details. Vinogradov gave that

$$T_1(N) = \frac{1}{2}\mathfrak{S}_3(N)\frac{N^2}{\log^3(N)} + O\left(\frac{N^2}{\log^4(N)}\right). \tag{6}$$

---

[8]Singular series have series forms. The series form has useful properties but a more complicated definition. The product form will suffice for our purposes. See Section 6.2 and 11.1 of [3] for details.

[9]See Chapter I, section 2 of [2] or Section 6.2 of [3] for full proofs.

By assumption and Theorem 6, $\mathfrak{S}_3(N) > \frac{1}{2}$. The rest of the terms in (6) are clearly positive for large N. Thus, $T_1(N) > 0$.

To bound $T_2(N)$, we need to estimate $S(\alpha, N)$ for $\alpha \in E_2$. Required methods are systematically presented in Chapter 5 of [3]. Here we will directly state the result:

$$S(\alpha, N) \ll \frac{N}{\log^3(N)}$$

For sufficiently large N, we can bound $T_2(N)$ as follows:

$$
\begin{aligned}
T_2(N) &= \int_{E_2} S^3(\alpha, N) e(-N\alpha)\, \mathrm{d}\alpha \\
&< \int_0^1 |S^3(\alpha, N)|\, \mathrm{d}\alpha \\
&\ll \frac{N}{\log^3(N)} \int_0^1 |S^2(\alpha, N)|\, d\alpha \\
&= \frac{N}{\log^3(N)} \sum_{2 < p_1,\, p_2 \leq N} \int_0^1 e\left((p_1 - p_2)\alpha\right)\, d\alpha \\
&= \frac{N}{\log^3(N)} \sum_{2 < p \leq N} 1 \qquad\qquad\qquad \text{(Equation (5))} \\
&\sim \frac{N^2}{\log^4(N)}. \qquad\qquad\qquad\qquad \text{(Prime number theorem)}
\end{aligned}
$$

Notice that

$$\frac{N^2}{\log^4(N)} \ll \frac{1}{2} \mathfrak{S}_3(N) \frac{N^2}{\log^3(N)}.$$

Hence $T_2(N)$ is insignificant compared to $T_1(N)$, $T(N) \sim T_1(N) > 0$ as $N \longrightarrow \infty$. $\qquad\square$

This concludes our brief overview of the circle method as applied to Goldbach's weak conjecture. We now turn to Goldbach's Conjecture.

4.3 Estimation $D_1(N)$ and $D_2(N)$

It is natural to apply what we did to bound $T(N)$ to $D(N)$. Let N be a sufficiently large (even) integer, a similar (but complex, see Section 11.1 of [3]) method of bounding can be used to achieve

$$D_1(N) = \int_{E_1} S^2(\alpha, N) e(-N\alpha)\, \mathrm{d}\alpha \sim \mathfrak{S}_2(N) \frac{N}{\log^2(N)} + O\left(\frac{n(\log\log(N))^2}{\log^3(N)}\right) \qquad (7)$$

However, as we saw when bounding $T_2(N)$,

$$
\begin{aligned}
D_2(N) &= \int_{E_2} S^2(\alpha, N) e(-N\alpha)\, \mathrm{d}\alpha \\
&< \int_0^1 |S^2(\alpha, N)|\, \mathrm{d}\alpha \sim \frac{N}{\log N},
\end{aligned}
$$

which means $D_2(N)$ is of larger order than $D_1(N)$. Although there are several other known ways[10] to achieve better bounds on T_2, those bounds are not tight enough for us to conclude $|D_2(N)| \ll |D_1(N)|$.

On a side note, we have the following properties of major arcs.

---

[10]One way is to use Cauchy-Schwarz, we leave this as an exercise. Chen Jingrun proposed to apply circle method on the minor arc E_2 (see Appendix I of [3]) which divides E_2 into E_2' and E_3 which over E_2', $D_2(N)$ is insignificant, so we only need to concern the integral over E_3, a much smaller interval than E_2, although we still cannot bound that integral.

Theorem 8. $\lim\limits_{N\to\infty} \mu(E_1) = 0$, *where μ is the Lebesgue measure on $\mathbb{R}$.*

Proof. Choose N such that $2Q^2 \ll N$ and choose $r = N$. Theorem 4 says the major arcs are disjoint. We have

$$
\begin{aligned}
\mu(E_1) &= \mu\left(\bigcup_{1 \le b \le Q} \bigcup_{\substack{0 \le a < b \\ \gcd(a,b)=1}} I(a,b) \right) \\
&= \sum_{1 \le b \le Q} \sum_{\substack{0 \le a < b \\ \gcd(a,b)=1}} \mu(I(a,b)) \\
&= \sum_{1 \le b \le Q} \sum_{\substack{0 \le a < b \\ \gcd(a,b)=1}} \frac{2}{N} \\
&\le \sum_{1 \le b \le Q} \frac{2b}{N} = \frac{Q(Q+1)}{N}
\end{aligned}
$$

Clearly, $0 \le \lim\limits_{N\to\infty} \mu(E_1) \le \lim\limits_{N\to\infty} \dfrac{Q(Q+1)}{N} = 0$. $\qquad\square$

This means that for sufficiently large N, the major arcs become very small, and most of the interval $[0,1]$ is made up of minor arcs. This weakens the case for the bounding of $D_1(N)$ and $D_2(N)$, which is already unconvincing: we cannot show $D_2(N)$ is of lower order compared to $D_1(N)$, and now Theorem 8 tells us that E_2 occupies almost the entire interval.

References

[1] Y. Li, *Introduction to Goldbach's Conjecture and its History*, $U(t)$-Mathazine. **9**(2024).

[2] W. Yuan, *The Goldbach Conjecture*, World Scientific, 2002.

[3] Pan, Cheng Dong, and Pan Cheng Biao, *Goldbach Conjecture*, Science Press, 2011. [潘承洞, 潘承彪, 哥德巴赫猜想. 科学出版社, 2011.]

[4] G. H. Hardy, J. E. Littlewood, *Some problems of 'Partitio numerorum; III: On the expression of a number as a sum of primes*, Acta Mathematica, 44(none) 1-70 1923, DOI 10.1007/BF02403921.

[5] H. A. Helfgott, *The ternary Goldbach conjecture is true*, arXiv:1205.5252, 2013.