Article

Not peer-reviewed version

An Asymptotic Approach to Twin Primes in a Prime-Dependent Interval

Ahmet F. Gocgen * and Emin M. Buyukyayla

Posted Date: 10 March 2025

doi: 10.20944/preprints202503.0585.v1

Keywords: twin prime conjecture; prime number theorem; hardy-littlewood estimate; montgomeryvaughan inequalities; analytic number theory; prime number distribution

Article

An Asymptotic Approach to Twin Primes in a Prime-Dependent Interval

Ahmet F. Gocgen [1,*] andEmin M. Buyukyayla [2]

[1] Department of Mathematics, Ordu University
[2] Independent Researcher
* Correspondence: ahmetfgocgen@gmail.com; Tel.: +90-552-750-66-35

Abstract: The Twin Prime Conjecture is a hypothesis that asserts the existence of infinitely many prime pairs of the form $(p, p+2)$ and is considered one of the most longstanding open problems in analytic number theory. In this study, we present a new approach to proving the infinitude of twin primes by examining specific prime number intervals. In particular, we demonstrate that there must be at least one twin prime pair within the interval $(p_k p_{k+1}, p_{k+1}^2)$. First, we prove that there is at least one prime in this interval using the Prime Counting Function, and then we establish that there are at least two primes. Subsequently, by employing the Hardy-Littlewood estimate and the Montgomery-Vaughan relations, we mathematically prove that at least one of these primes must form a twin prime pair. Additionally, we analyze the error term in this study, showing that the ratio of the error term to the main integral approaches zero, thereby ensuring the robustness of the obtained result.

Keywords: twin prime conjecture; prime number theorem,; hardy-littlewood estimate; montgomery-vaughan inequalities; analytic number theory,; prime number distribution

---

1. Introduction

The distribution of prime numbers has been one of the most fundamental and challenging problems in analytic number theory. Understanding the behavior of prime numbers is of great importance not only for theoretical mathematics but also for modern cryptography and computer science [1,2].

In particular, the question of whether there are infinitely many prime pairs of the form $(p, p+2)$, known as twin primes, has remained unsolved for nearly two centuries [3,4]. The Twin Prime Conjecture, which asserts the existence of infinitely many twin prime pairs, is one of the most famous unproven hypotheses.

In this study, we asymptotically analyze specific prime intervals to prove the infinitude of twin primes and demonstrate that at least one twin prime must exist within these intervals.

Specifically, the interval $(p_k \cdot p_{k+1}, p_{k+1}^2)$ has been chosen for analysis due to previous studies on the subject [5].

2. Methods

Definition 1. (Prime Counting Function)
$\pi(x)$ = It is the number of prime numbers up to x.
The following formula is known for the distribution of prime numbers:

$$\pi(x) \sim \frac{x}{\log x} \tag{1}$$

Definition 2. (Twin Prime Counting Function)
$\pi_2(x)$ = It is the number of twin prime numbers up to x.
Hardy-Littlewood twin prime estimate:

$$\pi_2(x) \sim 2C_2 \int_2^x \frac{dt}{(\log t)^2} \tag{2}$$

C_2 is a constant and is calculated as follows:

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2025 doi:10.20944/preprints202503.0585.v1

$$C_2 = \prod_{p \geq 3} \left(1 - \frac{1}{(p-1)^2}\right) \tag{3}$$

This is approximately:

$$C_2 \approx 0.6601618 \tag{4}$$

Definition 3. (Montgomery-Vaughan inequalities.)

Montgomery and Vaughan developed a powerful asymptotic relation that examines the distribution of prime numbers in specific intervals, particularly useful for predicting the average density of twin prime pairs.

Montgomery-Vaughan inequalities:

$$\sum_{n \leq x} (\pi(n+2) - \pi(n)) \approx \frac{x}{\log^2 x} \tag{5}$$

3. Theorems and Proofs

Lemma 1. There is at least one prime number in the interval $(p_k \cdot p_{k+1}, p_{k+1}^2)$.

Proof. According to the Prime Number Theorem:

$$\pi(x) \sim \frac{x}{\log x} \tag{6}$$

The number of prime numbers in this interval is:

$$\pi(p_{k+1}^2) - \pi(p_k p_{k+1}) \tag{7}$$

Approximately:

$$\pi(p_{k+1}^2) \approx \frac{p_{k+1}^2}{2\log p_{k+1}}, \quad \pi(p_k p_{k+1}) \approx \frac{p_k p_{k+1}}{\log p_k + \log p_{k+1}} \tag{8}$$

Hence:

$$\pi(p_{k+1}^2) - \pi(p_k p_{k+1}) \approx \frac{p_{k+1}^2}{2\log p_{k+1}} - \frac{p_k p_{k+1}}{\log p_k + \log p_{k+1}} \tag{9}$$

By showing that this is ≥ 1, we can prove that there is at least one prime:

$$\frac{p_{k+1}^2}{2\log p_{k+1}} \geq 1 + \frac{p_k p_{k+1}}{2\log p_{k+1}} \tag{10}$$

$$\frac{p_{k+1}^2 - p_k p_{k+1}}{2\log p_{k+1}} \geq 1 \tag{11}$$

In this case, we must examine the following expression to verify the inequality:

$$p_{k+1}^2 - p_k p_{k+1} \geq 2\log p_{k+1} \tag{12}$$

It is evident that this expression is true.

Lemma 2. There are at least two prime numbers in the interval $(p_k \cdot p_{k+1}, p_{k+1}^2)$.

Proof. Based on Lemma 1, we can obtain the following expression:

$$\frac{p_{k+1}^2}{2\log p_{k+1}} \geq 2 + \frac{p_k p_{k+1}}{2\log p_{k+1}} \tag{13}$$

Therefore:

$$\frac{p_{k+1}^2 - p_k p_{k+1}}{2\log p_{k+1}} \geq 2 \tag{14}$$

As a result:

$$p_{k+1}^2 - p_k p_{k+1} \geq 4\log p_{k+1} \tag{15}$$

It is also evident that this expression is true.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2025 doi:10.20944/preprints202503.0585.v1

Theorem 1.

$$\forall k, \quad \text{there is at least one twin prime pair in the interval } (p_k \cdot p_{k+1}, p_{k+1}^2). \tag{16}$$

Propositon 1. We obtain the same expression using the Hardy-Littlewood and Montgomery-Vaughan relations.

Proof. Using the Hardy-Littlewood formula, we obtain the following expression for the interval we have defined:

$$\pi_2(p_{k+1}^2) - \pi_2(p_k \cdot p_{k+1}) \approx 2C_2 \int_{p_k \cdot p_{k+1}}^{p_{k+1}^2} \frac{dt}{(\log t)^2} \tag{17}$$

Similarly, by using the Montgomery-Vaughan relations, we can obtain the same expression:

$$\sum_{n=p_k \cdot p_{k+1}}^{p_{k+1}^2} (\pi(n+2) - \pi(n)) \approx \int_{p_k \cdot p_{k+1}}^{p_{k+1}^2} \frac{dt}{\log^2 t} \tag{18}$$

Proposition 2. The ratio of the error term to the integral converges to 0 for the obtained expression.

Proof. The Hardy-Littlewood formula, including the error term, is expressed as follows:

$$\pi_2(x) = 2C_2 \int_2^x \frac{dt}{(\log t)^2} + O\left(\frac{x}{(\log x)^3}\right) \tag{19}$$

The error term here is:

$$H_1(x) = O\left(\frac{x}{\log^3 x}\right) \tag{20}$$

The Montgomery-Vaughan formula, including the error term, is expressed as follows:

$$\sum_{n=a}^{b} (\pi(n+2) - \pi(n)) = \int_a^b \frac{dt}{\log^2 t} + O\left(\frac{b}{\log^3 b}\right) \tag{21}$$

The error term here is:

$$H_2(b) = O\left(\frac{b}{\log^3 b}\right) \tag{22}$$

Let us recall our integral:

$$I_k = \int_{p_k p_{k+1}}^{p_{k+1}^2} \frac{dt}{\log^2 t} \tag{23}$$

To calculate the error term, let's compute the integral approximately:

$$\int_a^b \frac{dt}{(\log t)^2} \approx \frac{b-a}{(\log b)^2} \tag{24}$$

$$I_k \approx \frac{p_{k+1}^2 - p_k p_{k+1}}{(\log p_{k+1}^2)^2} \tag{25}$$

Here,

$$\log(p_{k+1}^2) = 2 \log p_{k+1} \tag{26}$$

$$(\log p_{k+1}^2)^2 = (2 \log p_{k+1})^2 = 4 \log^2 p_{k+1} \tag{27}$$

Therefore:

$$I_k \approx \frac{p_{k+1}^2 - p_k p_{k+1}}{4 \log^2 p_{k+1}} \tag{28}$$

To simplify further:

$$p_k \approx k \log k \tag{29}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2025 doi:10.20944/preprints202503.0585.v1

From here:

$$p_{k+1} - p_k = O(\sqrt{p_k}) \tag{30}$$

$$p_{k+1}^2 - p_k p_{k+1} = p_{k+1}(p_{k+1} - p_k) \tag{31}$$

$$= p_{k+1} O(\sqrt{p_k}) \tag{32}$$

$$p_{k+1}^2 - p_k p_{k+1} \approx O(p_{k+1}\sqrt{p_k}) \tag{33}$$

Therefore, our integral is:

$$I_k \approx \frac{p_{k+1}\sqrt{p_k}}{4\log^2 p_{k+1}} \tag{34}$$

The error term is:

$$H_k = O\left(\frac{p_{k+1}\sqrt{p_k}}{4\log^3 p_{k+1}}\right) \tag{35}$$

To determine whether the error term can be neglected in the integral calculation, we need to compute the ratio of the error term to the integral:

$$\frac{H_k}{I_k} = \frac{O\left(\frac{p_{k+1}\sqrt{p_k}}{4\log^3 p_{k+1}}\right)}{\frac{p_{k+1}\sqrt{p_k}}{4\log^2 p_{k+1}}} \tag{36}$$

$$\frac{H_k}{I_k} = O\left(\frac{1}{\log p_{k+1}}\right) \tag{37}$$

Since we are doing asymptotic analysis, we need to look at the limit at infinity:

$$\lim_{p_{k+1}\to\infty} \frac{H_k}{I_k} = 0 \tag{38}$$

As a result, the error term can be neglected.

Proposition 3. The value of the integral is greater than or equal to 1.

Proof. At this stage, we need to calculate the integral in more detail.

$$I_k = \int_{p_k \cdot p_{k+1}}^{p_{k+1}^2} \frac{dt}{\log^2 t} \tag{39}$$

According to prime number theory:

$$p_k \approx k \log k \tag{40}$$

$$p_k p_{k+1} \approx (k \log k) \cdot ((k+1)\log(k+1)) \tag{41}$$

$$p_{k+1}^2 \approx (k+1)^2 \log^2(k+1) \tag{42}$$

$$\log(p_k p_{k+1}) \approx \log k + \log(k+1) + \log\log k + \log\log(k+1) \tag{43}$$

$$\log(p_{k+1}^2) \approx 2(\log(k+1) + \log\log(k+1)) \tag{44}$$

We can express the integral as follows:

$$I_k \approx \int_{\log(p_k p_{k+1})}^{\log(p_{k+1}^2)} \frac{e^u\, du}{u^2} \tag{45}$$

$$\int e^u u^{-2} du \approx \frac{e^u}{u^2} \tag{46}$$

 doi:10.20944/preprints202503.0585.v1

Hence:

$$I_k \approx \left[\frac{e^u}{u^2}\right] \log(p_k pk + 1)^{\log(p_{k+1}^2)} \tag{47}$$

Here, the upper limit will be as follows:

$$\frac{p_{k+1}^2}{(2\log(k+1))^2} \tag{48}$$

The lower limit will be as follows:

$$\frac{p_k p_{k+1}}{(\log k \log(k+1))^2} \tag{49}$$

Therefore, the integral becomes:

$$I_k \approx \frac{p_{k+1}^2}{(2\log(k+1))^2} - \frac{p_k p_{k+1}}{(\log k \log(k+1))^2} \tag{50}$$

The proposition is true if the following condition is satisfied:

$$\frac{p_{k+1}^2}{(2\log p_{k+1})^2} \geq 1 + \frac{p_k p_{k+1}}{(\log p_k + \log p_{k+1})^2} \tag{51}$$

For this reason, the following expression must be incorrect:

$$\frac{p_{k+1}^2}{(2\log p_{k+1})^2} < 1 \tag{52}$$

Again, according to prime number theory:

$$p_{k+1} \approx p_k + O(\sqrt{p_k}) \tag{53}$$

$$\log p_k \approx \log p_{k+1} \tag{54}$$

$$\log p_k + \log p_{k+1} \approx 2\log p_{k+1} \tag{55}$$

Thence:

$$\frac{p_k p_{k+1}}{(2\log p_{k+1})^2} \tag{56}$$

$$= \frac{p_k p_{k+1}}{4(\log p_{k+1})^2} \tag{57}$$

The following expression can be formed:

$$p_k p_{k+1} < p_{k+1}^2 \tag{58}$$

From here:

$$\frac{p_k p_{k+1}}{4(\log p_{k+1})^2} < \frac{p_{k+1}^2}{4(\log p_{k+1})^2} \tag{59}$$

Hence:

$$\frac{p_{k+1}^2}{4(\log p_{k+1})^2} \geq 1 + \frac{p_k p_{k+1}}{4(\log p_{k+1})^2} \tag{60}$$

Then:

$$\frac{p_{k+1}^2 - p_k p_{k+1}}{4(\log p_{k+1})^2} \geq 1 \tag{61}$$

For this reason:

$$p_{k+1}^2 - p_k p_{k+1} \geq 4(\log p_{k+1})^2 \tag{62}$$

 doi:10.20944/preprints202503.0585.v1

It is evident that this expression is true.

As a result:

$$\frac{p_{k+1}^2}{(2\log p_{k+1})^2} \geq 1 + \frac{p_k p_{k+1}}{(\log p_k + \log p_{k+1})^2} \tag{63}$$

This situation can also be demonstrated using a limit. The limit expression will be as follows:

$$\lim_{p_k \to \infty} \left(\frac{p_{k+1}^2}{(2\log p_{k+1})^2} - \left(1 + \frac{p_k p_{k+1}}{(\log p_k + \log p_{k+1})^2} \right) \right) \tag{64}$$

From here, it is easy to conclude:

$$\lim_{p_k \to \infty} \left(\frac{p_{k+1}^2}{(2\log p_{k+1})^2} - \left(1 + \frac{p_k p_{k+1}}{(\log p_k + \log p_{k+1})^2} \right) \right) \geq 0 \tag{65}$$

Corollaries 1. Proposition 3 naturally shows that there must be at least one twin prime pair in the asymptotic interval. This completes the proof of Theorem 1.

4. Results

Firstly, we showed that there must be at least one prime number in the interval $(p_k p_{k+1}, p_{k+1}^2)$.

We used the asymptotic formula of the prime counting function to obtain this result:

$$\pi(x) \sim \frac{x}{\log x} \tag{66}$$

According to this relation, the number of prime numbers in the specified interval was calculated as follows:

$$\pi(p_{k+1}^2) - \pi(p_k p_{k+1}) \approx \frac{p_{k+1}^2}{2\log p_{k+1}} - \frac{p_k p_{k+1}}{\log p_k + \log p_{k+1}} \tag{67}$$

From here, it was shown that this difference is greater than or equal to 1, meaning that there must be at least one prime number in this interval. Secondly, we demonstrated that there must be at least two primes in the specified interval.

Finally, we showed that one of the at least two prime numbers in the interval must necessarily be a twin prime.

At this stage, the Hardy-Littlewood twin prime conjecture was used in the following form:

$$\pi_2(p_{k+1}^2) - \pi_2(p_k p_{k+1}) \approx 2C_2 \int_{p_k p_{k+1}}^{p_{k+1}^2} \frac{dt}{(\log t)^2} \tag{68}$$

The result of this integral is always greater than 1 for large values of k. This proves that there must be at least one twin prime pair in every specified interval.

One of the most important points in this study was the analysis of the error term.

By denoting the error term as H_k and the integral value as I_k, we showed that the error term converges to zero proportionally:

$$\lim_{p_{k+1} \to \infty} \frac{H_k}{I_k} = 0 \tag{69}$$

5. Discussion

This study presents a systematic approach to proving the infinitude of twin primes. Our work provides a fresh perspective by directly analyzing the existence of twin primes in specified prime number intervals.

By using the Hardy-Littlewood twin prime conjecture and Montgomery-Vaughan relations directly within an analytical framework, we obtained definitive results. Specifically, by driving the limit of the error term to zero, we controlled the error margin, thus providing a mathematically rigorous proof rather than a probabilistic or heuristic estimate.

The results obtained not only apply to twin primes but could also facilitate similar analyses for other prime pairs such as $(p, p+4)$, $(p, p+6)$, and others. Moreover, this could contribute to the development of new approaches to the Riemann hypothesis and the broader distribution of prime numbers.

 doi:10.20944/preprints202503.0585.v1

The distribution of twin primes is significant not only in theoretical mathematics but also in cryptography and computer science. The RSA algorithm relies on the difficulty of factoring large prime numbers. A better understanding of twin prime distribution could allow for the more efficient selection of prime numbers. Understanding the density of large primes in certain intervals could optimize random prime generation methods for cryptographic security. Determining the exact distribution of twin primes could lead to the development of faster and more effective algorithms for discovering prime numbers.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: Not applicable.

referencias.bib

1. Fernando Peralta Castr. The Evolution of Cryptography through Number Theor *arXiv* **2024**.
2. Mohan Satvik Adusumalli. A study of the importance of prime numbers in cryptographic algorithms *IJUSE* **2000**, *10*, 1–10.
3. Harvey Dubner. Twin prime conjectures *Journal Abbreviation* **2000**, *30*, 199–205.
4. Sadegh Nazardonyavi. Some history about twin prime conjecture *arXiv* **2012**.
5. Ahmet F. Gocge. Gocgen Approach for Bounded Gaps Between Odd Composite Numbers *Preprints* **2024**.