Note for the Large and Small Gaps

Frank Vega

May 29, 2024

Article

Note for the Large and Small Gaps

Frank Vega [1]

[1] GROUPS PLUS TOURS INC., 9611 Fontainebleau Blvd, Miami, FL, 33172, USA; vega.frank@gmail.com

Abstract: A prime gap is the difference between two successive prime numbers. The nth prime gap, denoted g_n is the difference between the (n + 1)st and the nth prime numbers, i.e. $g_n = p_{n+1} - p_n$. A twin prime is a prime that has a prime gap of two. On the one hand, the twin prime conjecture states that there are infinitely many twin primes. There isn't a verified solution to twin prime conjecture yet. In this note, using the Chebyshev function, we prove that

$$\liminf_{n\to\infty} \frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n + 2)} \geq 1,$$

under the assumption that the twin prime conjecture is false. It is well-known the proof of Daniel Goldston, János Pintz and Cem Yildirim which implies that $\liminf_{n\to\infty} \frac{g_n}{\log p_n} = 0$. In this way, we reach an intuitive contradiction. Consequently, by reductio ad absurdum, we can conclude that the twin prime conjecture is true. On the other hand, the Andrica's conjecture deals with the difference between the square roots of consecutive prime numbers. While mathematicians have showed it true for a vast number of primes, a general solution remains elusive. We consider the inequality $\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$ for two successive prime numbers p_n and p_{n+1}, where $\theta(x)$ is the Chebyshev function. In this note, under the assumption that the inequality $\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$ holds for all $n \geq 1.3002 \cdot 10^{16}$, we prove that the Andrica's conjecture is true. Since $\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$ holds indeed for large enough prime number p_n, then we show that the statement of the Andrica's conjecture can always be true for all primes greater than some threshold.

Keywords: prime gaps; prime numbers; Chebyshev function; primorial numbers

MSC: 11A41; 11A25

---

1. Introduction

Prime numbers, the building blocks of integers, have fascinated mathematicians for centuries. Their irregular distribution, with gaps of seemingly random size between them, is a source of ongoing intrigue. A twin prime is a prime that has a prime gap of two. The twin prime conjecture states that there are infinitely many twin primes. The question of whether there exist infinitely many twin primes has been one of the great open questions in number theory for many years. In 1849, de Polignac made the more general conjecture that for every natural number k, there are infinitely many primes p such that $p + 2 \cdot k$ is also prime [1]. The case $k = 1$ of de Polignac's conjecture is the twin prime conjecture. There is a stronger form of the twin prime conjecture, the Hardy-Littlewood conjecture, postulates a distribution law for twin primes [2].

In May 2013, the popular Yitang Zhang's paper was accepted by the journal Annals of Mathematics where it was announced that for some integer N, that is less than 70 million, there are infinitely many pairs of primes that differ by N [3]. A few months later, James Maynard gave a different proof of Yitang Zhang's theorem and showed that there are infinitely many prime gaps with size of at most 600 [4]. A collaborative effort in the Polymath Project, led by Terence Tao, reduced to the lower bound 246 just using Zhang and Maynard results. Moreover assuming the Elliott-Halberstam conjecture and its generalized form, the Polymath Project wiki states that the bound is 12 and 6, respectively. As of August 2022, the current largest twin prime pair known is $2996863034895 \cdot 2^{1290000} \pm 1$ [5].

In this work, using a proof by contradiction, we prove that the twin prime conjecture is true. The resolution of the twin prime conjecture would undoubtedly inspire mathematicians to tackle even more challenging unsolved problems. It could open doors to entirely new areas of inquiry, pushing the boundaries of human knowledge in number theory. A proven twin prime conjecture would be far more than just satisfying an intellectual curiosity. It would represent a significant leap forward in our understanding of prime numbers, potentially leading to advancements across various branches of mathematics with far-reaching consequences. The impact could be profound, both for theoretical knowledge and potentially for practical applications in the future.

Andrica's conjecture tackles this very irregularity, proposing a relationship between the sizes of these prime gaps and the primes themselves. Andrica's conjecture (named after Dorin Andrica) is a conjecture regarding the gaps between prime numbers [6]. The conjecture states that the inequality

$$\sqrt{p_{n+1}} - \sqrt{p_n} < 1$$

holds for all n, where p_n is the nth prime number. If $g_n = p_{n+1} - p_n$ denotes the nth prime gap, then Andrica's conjecture can also be rewritten as

$$g_n < 2 \cdot \sqrt{p_n} + 1.$$

Imran Ghory has used data on the largest prime gaps to confirm the conjecture for n up to $1.3002 \cdot 10^{16}$ [7].

Legendre's conjecture, proposed by Adrien-Marie Legendre, states that there is a prime number between n^2 and $(n+1)^2$ for every positive integer n [7]. The conjecture is one of Landau's problems (1912) on prime numbers. If Legendre's conjecture is true, the gap between any prime p and the next largest prime would be $O(\sqrt{p}\,)$, as expressed in big O notation. Oppermann's conjecture is another unsolved problem in mathematics on the distribution of prime numbers [7]. It is closely related to but stronger than Legendre's conjecture and Andrica's conjecture. It is named after Danish mathematician Ludvig Oppermann, who announced it in an unpublished lecture in March 1877 [8]. If the conjecture is true, then the gap size would be on the order of $g_n < \sqrt{p_n}$.

This seemingly simple statement has profound implications for our understanding of prime number distribution. Unfortunately, despite its apparent elegance, Andrica's conjecture remains unproven. Mathematicians have extensively verified it for a tremendous number of primes, but a universal solution proving its truth for all primes continues to be elusive. This lack of proof doesn't diminish the significance of the conjecture. It serves as a guidepost, nudging mathematicians towards a deeper understanding of prime number distribution. The quest to solve Andrica's conjecture pushes the boundaries of our knowledge and holds the potential to unlock new insights into the enigmatic world of primes. We also study the inequality $\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$ for two successive prime numbers p_n and p_{n+1} which has a close relation to Andrica's conjecture, where $\theta(x)$ is the Chebyshev function.

2. Materials and methods

In mathematics, the Chebyshev function $\theta(x)$ is given by

$$\theta(x) = \sum_{p \leq x} \log p$$

with the sum extending over all prime numbers p that are less than or equal to x, where log is the natural logarithm. We know the following properties of this function:

Proposition 1. *For every $x \geq 41$ [9, Corollary pp. 70]:*

$$\left(1 - \frac{1}{\log x}\right) \cdot x < \theta(x).$$

Proposition 2. *We have [10, pp. 1539]:*

$$\theta(x) \sim x \;\; as \;\; (x \to \infty).$$

A natural number N_n is called a primorial number of order n precisely when,

$$N_n = \prod_{k=1}^{n} p_k$$

where p_k is the kth prime number (We also use the notation p_n to denote the nth prime number). This implies that $\theta(p_n) = \log N_n$.

Proposition 3. *For $n \geq 25$ there is always a prime between n and $\left(1 + \frac{1}{5}\right) \cdot n$ [11].*

The definition of limit inferior is widely used in mathematics.

Definition 1. *The limit inferior of a sequence of real numbers x_n is the largest real number b such that, for any positive real number ε, there exists a natural number N such that $x_n > b - \varepsilon$ for all $n > N$. In other words, any number below the limit inferior is an eventual lower bound for the sequence. Only a finite number of elements of the sequence are less than $b - \varepsilon$.*

The following is a key Proposition:

Proposition 4. *If $g_n = p_{n+1} - p_n$ denotes the nth prime gap, then we know that [12]:*

$$\liminf_{n \to \infty} \frac{g_n}{\log p_n} = 0.$$

Putting all together yields two proofs related to large and small prime gaps.

3. Results

3.1. The twin prime conjecture

This is a main insight of this section.

Theorem 1. *If we assume that the twin prime conjecture is false, then*

$$\liminf_{n \to \infty} \frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n + 2)} \geq 1.$$

Proof. If p_n and p_{n+1} are twin primes, then

$$p_n \cdot (N_n + 2 \cdot N_{n-1}) = N_{n+1}.$$

Certainly, we have

$$\begin{aligned} p_n \cdot (N_n + 2 \cdot N_{n-1}) &= p_n \cdot N_{n-1} \cdot (p_n + 2) \\ &= N_n \cdot (p_n + 2) \\ &= N_{n+1} \end{aligned}$$

whenever $p_{n+1} = p_n + 2$. Suppose that the twin prime conjecture is false. Hence, there exists a large enough prime $p_{n_0} > 2996863034895 \cdot 2^{1290000} + 1$ such that

$$p_n \cdot (N_n + 2 \cdot N_{n-1}) < N_{n+1}$$

holds for all $n \geq n_0$. That is the same as

$$\log(p_n \cdot (N_n + 2 \cdot N_{n-1})) < \log N_{n+1}$$

after of applying the logarithm to the both sides. That is equivalent to

$$\log(p_n) + \log(N_n + 2 \cdot N_{n-1}) < \theta(p_{n+1})$$

which means that

$$1 < \frac{\theta(p_{n+1}) - \theta(p_{n-1})}{\log(p_n) + \log(p_n + 2)}$$

holds for all $n \geq n_0$, because of

$$\begin{aligned} \log(N_n + 2 \cdot N_{n-1}) &= \log(N_{n-1} \cdot (p_n + 2)) \\ &= \log(N_{n-1}) + \log(p_n + 2) \\ &= \theta(p_{n-1}) + \log(p_n + 2). \end{aligned}$$

By Proposition 2, we see that

$$\theta(p_{n+1}) \sim p_{n+1} \ \ as \ \ (n \to \infty)$$

and

$$\theta(p_{n-1}) \sim p_{n-1} \ \ as \ \ (n \to \infty).$$

In addition, we notice that

$$\begin{aligned} p_{n+1} - p_{n-1} &= (p_{n+1} - p_n) + (p_n - p_{n-1}) \\ &= g_n + g_{n-1}. \end{aligned}$$

Note that, the inequality

$$1 < \frac{\theta(p_{n+1}) - \theta(p_{n-1})}{\log(p_n) + \log(p_n + 2)}$$

implies that

$$1 < \frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n + 2)} + \varepsilon_n$$

where

$$\varepsilon_n = \frac{(\theta(p_{n+1}) - p_{n+1}) - (\theta(p_{n-1}) - p_{n-1})}{\log(p_n) + \log(p_n + 2)}$$

tends rapidly to zero when $n \to \infty$ and

$$\frac{\theta(p_{n+1}) - \theta(p_{n-1})}{\log(p_n) + \log(p_n + 2)} = \frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n + 2)} + \varepsilon_n.$$

By definition of limit inferior, we finally deduce that

$$\liminf_{n \to \infty} \frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n + 2)} \geq 1$$

when we assume that the twin prime conjecture is false. $\square$

This is the main theorem of this section.

Theorem 2. *The twin prime conjecture is true.*

Proof. This is a direct consequence of putting together Proposition 4 with Theorem 1 and using a proof by contradiction. Certainly, the sequence of positive real numbers $x_n = \frac{g_n + g_{n-1}}{\log(p_n)+\log(p_n+2)}$ is upper bounded by $y_n = max\left(\frac{g_n}{\log(p_n)}, \frac{g_{n-1}}{\log(p_{n-1})}\right)$ since

$$\frac{g_n + g_{n-1}}{\log(p_n) + \log(p_n+2)} < \left(\frac{g_n}{2 \cdot \log(p_n)} + \frac{g_{n-1}}{2 \cdot \log(p_{n-1})}\right) \leq max\left(\frac{g_n}{\log(p_n)}, \frac{g_{n-1}}{\log(p_{n-1})}\right),$$

where $max(\ldots, \ldots)$ is the maximum function. Since this implies that

$$1 \leq \liminf_{n\to\infty} x_n \leq \liminf_{n\to\infty} y_n = 0$$

by Proposition 4 and Theorem 1, we reach a contradiction. Consequently, by reductio ad absurdum, we can confirm that the twin prime conjecture is true. □

3.2. The Andrica's conjecture

The following is a key Lemma.

Lemma 1. *Let p_n and p_{n+1} be two successive prime numbers such that $n \geq 1.3002 \cdot 10^{16}$. Then,*

$$\theta(p_{n+1}) < \theta(p_n) \cdot \left(1 + \frac{1}{\sqrt{p_n}}\right).$$

Proof. The inequality

$$\theta(p_{n+1}) < \theta(p_n) \cdot \left(1 + \frac{1}{\sqrt{p_n}}\right).$$

would be

$$\log(\theta(p_{n+1})) - \log(\theta(p_n)) < \log\left(1 + \frac{1}{\sqrt{p_n}}\right).$$

after of applying the logarithm to the both sides and distributing the terms. By properties of the Chebyshev function, we have

$$\begin{aligned}
\log(\theta(p_{n+1})) - \log(\theta(p_n)) &= \log\log(N_{n+1}) - \log\log(N_n) \\
&= \log(\log(N_n) + \log(p_{n+1})) - \log\log(N_n) \\
&= \log\left((\log(N_n)) \cdot \left(1 + \frac{\log(p_{n+1})}{\log(N_n)}\right)\right) - \log\log(N_n) \\
&= \log\log(N_n) + \log\left(1 + \frac{\log(p_{n+1})}{\log(N_n)}\right) - \log\log(N_n) \\
&= \log\left(1 + \frac{\log(p_{n+1})}{\log(N_n)}\right) \\
&= \log\left(1 + \frac{\log(p_{n+1})}{\theta(p_n)}\right).
\end{aligned}$$

In this way, we obtain that

$$\log\left(1 + \frac{\log(p_{n+1})}{\theta(p_n)}\right) < \log\left(1 + \frac{1}{\sqrt{p_n}}\right)$$

which is

$$\left(1 + \frac{\log(p_{n+1})}{\theta(p_n)}\right) < \left(1 + \frac{1}{\sqrt{p_n}}\right)$$

and
$$\frac{\log(p_{n+1})}{\theta(p_n)} < \frac{1}{\sqrt{p_n}}$$
after simplifying the whole expression. We show that
$$\frac{\log(p_{n+1})}{\left(1 - \frac{1}{\log p_n}\right) \cdot p_n} < \frac{1}{\sqrt{p_n}}$$
since
$$\frac{1}{\left(1 - \frac{1}{\log p_n}\right) \cdot p_n} > \frac{1}{\theta(p_n)}$$
by Proposition 1. That is equivalent to
$$\frac{\log(p_n)}{\log(p_n) - 1} \cdot \log(p_{n+1}) < \sqrt{p_n}$$
because of
$$\sqrt{p_n} = \frac{p_n}{\sqrt{p_n}}.$$
That would be
$$2 \cdot \log(p_{n+1}) < \sqrt{p_n}$$
since the fraction $\frac{x}{x-1}$ decreases as x increases whenever $x > 1$ and so,
$$\frac{\log(p_n)}{\log(p_n) - 1} < \frac{2}{2-1} = 2.$$
Hence, it is enough to show that
$$2 \cdot \log\left(\left(1 + \frac{1}{5}\right) \cdot p_n\right) < \sqrt{p_n}$$
trivially holds for $n \geq 1.3002 \cdot 10^{16}$ according to the Preposition 3. Thus, the proof is done. $\square$

This is a main insight of this section.

Theorem 3. *For $n \geq 1.3002 \cdot 10^{16}$, the inequality*
$$\sqrt{p_{n+1}} - \sqrt{p_n} < 1$$
holds when
$$\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$$
holds as well.

Proof. There is not any natural number n' such that
$$\sqrt{p_{n'+1}} - \sqrt{p_{n'}} = 1$$
since this implies that $g_{n'} = 2 \cdot \sqrt{p_{n'}} + 1$. For every n, g_n is a natural number and $2 \cdot \sqrt{p_n} + 1$ is always irrational. In fact, all square roots of natural numbers, other than of perfect squares, are irrational [13]. Suppose that there exists a natural number $n_0 \geq 1.3002 \cdot 10^{16}$ such that
$$\sqrt{p_{n_0+1}} - \sqrt{p_{n_0}} > 1$$

under the assumption that the inequality

$$\frac{\theta(p_{n_0+1})}{\theta(p_{n_0})} \geq \sqrt{\frac{p_{n_0+1}}{p_{n_0}}}$$

holds. That is equivalent to

$$\sqrt{\frac{p_{n_0+1}}{p_{n_0}}} - 1 > \frac{1}{\sqrt{p_{n_0}}}$$

and

$$\sqrt{\frac{p_{n_0+1}}{p_{n_0}}} > 1 + \frac{1}{\sqrt{p_{n_0}}}$$

after dividing both sides by $\sqrt{p_{n_0}}$ and distributing the terms. We obtain that

$$\frac{\theta(p_{n_0+1})}{\theta(p_{n_0})} > 1 + \frac{1}{\sqrt{p_{n_0}}}$$

when we assume that

$$\frac{\theta(p_{n_0+1})}{\theta(p_{n_0})} \geq \sqrt{\frac{p_{n_0+1}}{p_{n_0}}}.$$

That would be the same as

$$\theta(p_{n_0+1}) > \theta(p_{n_0}) \cdot \left(1 + \frac{1}{\sqrt{p_{n_0}}}\right).$$

Since this implies that the Lemma 1 should be false for some $n_0 \geq 1.3002 \cdot 10^{16}$, we reach a contradiction. Consequently, by reductio ad absurdum, we conclude that the Theorem 3 is true. $\square$

This is the main theorem of this section.

Theorem 4. *There exists some natural number $n_0 \geq 1.3002 \cdot 10^{16}$ such that $g_n < 2 \cdot \sqrt{p_n} + 1$ for $n \geq n_0$. Moreover, the Andrica's conjecture is true if the inequality $\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$ holds for all $n \geq 1.3002 \cdot 10^{16}$.*

Proof. By Proposition 2, the inequality

$$\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$$

holds for large enough prime number p_n since

$$\frac{\theta(p_{n+1})}{\theta(p_n)} \sim \frac{p_{n+1}}{p_n} \ \ as \ \ (n \to \infty)$$

and

$$\frac{p_{n+1}}{p_n} \gg \sqrt{\frac{p_{n+1}}{p_n}}$$

where the symbol $\gg$ means "much greater than". Therefore, there exists some natural number $n_0 \geq 1.3002 \cdot 10^{16}$ such that the inequality

$$\frac{\theta(p_{n+1})}{\theta(p_n)} \geq \sqrt{\frac{p_{n+1}}{p_n}}$$

holds for all $n \geq n_0$. To sum up, the Theorem 4 is a direct consequence of Theorem 3. $\square$

4. Conclusion

Further exploration about this result may involve:

- Developing new techniques in analytic number theory, the branch of mathematics that studies the distribution of prime numbers.
- Leveraging advanced computational methods to test this result for even larger prime ranges and potentially uncover patterns.
- Investigating connections between this result and other unsolved problems in prime number theory.

This result could be a significant advancement in our understanding of prime number distribution.

References

1. de Polignac, A. *Recherches nouvelles sur les nombres premiers*; Vol. 29, 1849; pp. 397–401.
2. Bateman, P.T.; Diamond, H.G. *Analytic number theory: an introductory course*; Vol. 1, World Scientific, 2004.
3. Zhang, Y. Bounded gaps between primes. *Annals of Mathematics* **2014**, pp. 1121–1174. https://doi.org/10.4007/annals.2014.179.3.7.
4. Maynard, J. Small gaps between primes. *Annals of mathematics* **2015**, pp. 383–413. https://doi.org/10.4007/annals.2015.181.1.7.
5. Caldwell, C.K. The Prime Database: $2996863034895 \cdot 2^{1290000} - 1$. https://primes.utm.edu/primes/page.php?id=122213, 2016. Accessed 7 May 2024.
6. Andrica, D. Note on a conjecture in prime number theory. *Studia Univ. Babes-Bolyai Math* **1986**, *31*, 44–48.
7. Wells, D. *Prime Numbers: The Most Mysterious Figures in Math*; Turner Publishing Company, 2011.
8. Oppermann, L. Om vor Kundskab om Primtallenes maengde mellem givne Graendser. *Oversigt over det Kongelige Danske Videnskabernes Selskabs Forhandlinger og dets Medlemmers Arbejder* **1882**, pp. 169–179.
9. Rosser, J.B.; Schoenfeld, L. Approximate Formulas for Some Functions of Prime Numbers. *Illinois Journal of Mathematics* **1962**, *6*, 64–94. https://doi.org/10.1215/ijm/1255631807.
10. Platt, D.J.; Trudgian, T.S. On the first sign change of $\theta(x) - x$. *Mathematics of Computation* **2016**, *85*, 1539–1547. https://doi.org/10.1090/mcom/3021.
11. Nagura, J. On the interval containing at least one prime number. *Proceedings of the Japan Academy* **1952**, *28*, 177–181. https://doi.org/10.3792/pja/1195570997.
12. Goldston, D.A.; Pintz, J.; Yalçin Yıldırım, C. Primes in tuples II. *Acta Mathematica* **2010**, *204*, 1–47. https://doi.org/10.1007/s11511-010-0044-9.
13. Jackson, T. 95.42 Irrational square roots of natural numbers - a geometrical approach. *The Mathematical Gazette* **2011**, *95*, 327–330. https://doi.org/10.1017/S0025557200003193.

Short Biography of Authors

Frank Vega is essentially a Back-End Programmer and Mathematical Hobbyist who graduated in Computer Science in 2007. In May 2022, The Ramanujan Journal accepted his mathematical article about the Riemann hypothesis. The article "Robin's criterion on divisibility" makes several significant contributions to the field of number theory. It provides a proof of the Robin inequality for a large class of integers, and it suggests new directions for research in the area of analytic number theory.