The smallest gap between primes

Frank Vega

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.

Abstract

A prime gap is the difference between two successive prime numbers. 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. 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. 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. A collaborative effort in the Polymath Project, led by Terence Tao, reduced to the lower bound 246 just using Zhang and Maynard results. In this note, using arithmetic operations, we prove that the twin prime conjecture is true.

Frank Vega

Research Department, NataSquad, 10 rue de la Paix, Paris, 75002, France

vega.frank@gmail.com

Keywords: Twin prime conjecture, Prime numbers, Prime gap.

MSC Classification: 11A41, 11A25

1. Introduction

Leonhard Euler studied the following value of the Riemann zeta function (1734).

Proposition 1. *It is known that [[1], (1) pp. 1070]:*

$$\zeta(2) = \prod_{k=1}^{\infty} \frac{p_k^2}{p_k^2 - 1} = \frac{\pi^2}{6},$$

where p_k is the kth prime number (We also use the notation p_n to denote the nth prime number).

Proposition 2. For $y > 24317$ [2]:

$$\sum_{p_k \geq y} \log\left(\frac{p_k^2}{p_k^2 - 1}\right) \geq \frac{1}{y \cdot \log y} - \frac{1}{y \cdot \log^2 y} + \frac{2}{y \cdot \log^3 y} - \frac{10.26}{y \cdot \log^4 y}.$$

Franz Mertens obtained some important results about the constants B and H (1874). We define $H = \gamma - B$ such that $B \approx 0.2614972128$ is the Meissel-Mertens constant and $\gamma \approx 0.57721$ is the Euler-Mascheroni constant [[3], (17.) pp. 54].

Proposition 3. *We have* [[4], Lemma 2.1 (1) pp. 359]:

$$\sum_{k=1}^{\infty} \left(\log\left(\frac{p_k}{p_k - 1}\right) - \frac{1}{p_k}\right) = \gamma - B = H,$$

where log *is the natural logarithm.*

For $x \geq 2$, the function $u(x)$ is defined as follows [[5], pp. 379]:

$$u(x) = \sum_{p_k > x} \left(\log\left(\frac{p_k}{p_k - 1}\right) - \frac{1}{p_k}\right).$$

We use the following function:

Definition 1. For all $x > 1$ and $a \geq 0$, we define the function:

$$H_a(x) = \log\left(\frac{x}{x-1}\right) - \frac{1}{x+a} + \log\left(\frac{x^2 - \frac{\log(x)+1}{\sqrt{x}}}{x^2}\right).$$

We state the following Propositions:

Proposition 4. *For a sufficiently large positive value x, we have* $H_2(x) < 0$. *Certainly,* $H_2(x)$ *is negative for all* $x \geq 60000$ *since it is negative for* $x = 60000$, *strictly decreasing for* $x \geq 60000$ *(because its derivative is lesser than* 0 *for* $x \geq 60000$*) and its greatest root is between* 50000 *and* 60000 *(See Figure 1).*

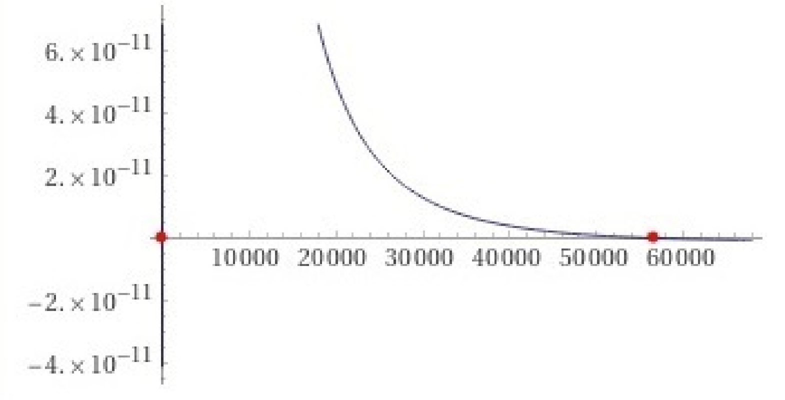

Figure 1. Roots of $H_2(x)$ [6]

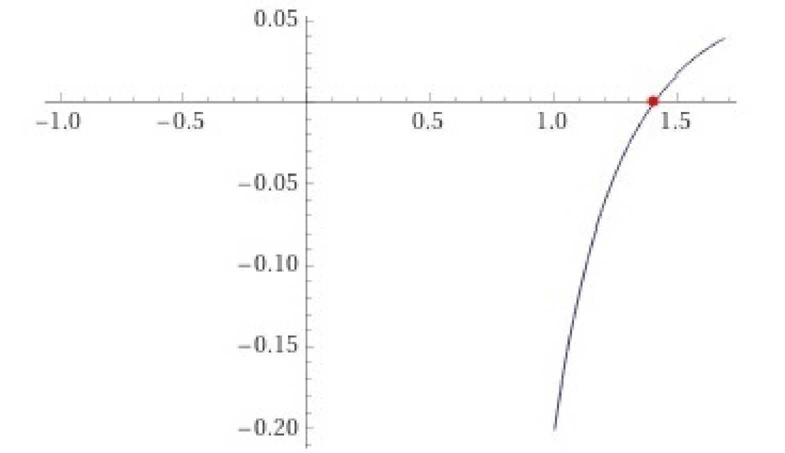

Figure 2. Roots of $H_4(x)$ [7]

Proposition 5. *For a sufficiently large positive value x, we have* $H_4(x) > 0$*. Certainly,* $H_4(x)$ *is positive for all* $x \geq 1.5$ *since it is positive for* $x = 1.5$ *and its unique root is between* 1.4 *and* 1.5 *(See Figure 2).*

The following property is based on natural logarithms:

Proposition 6. [8]*. For* $x > -1$*:*

$$\log(1 + x) \leq x.$$

Proposition 7. [[9], Theorem 1.1 (13) pp. 3]. For $x \geq 1$:

$$\frac{1}{x + 0.4} > \log\left(1 + \frac{1}{x}\right) > \frac{1}{x + 0.5}.$$

Putting all together yields the proof of the main theorem.

Theorem 1. *The twin prime conjecture is true.*

2. Infinite Sums

Lemma 1.

$$\sum_{k=1}^{\infty}\left(\frac{1}{p_k}-\log(1+\frac{1}{p_k})\right)=\log(\zeta(2))-H$$

Proof. We obtain that

$$\begin{aligned}
\log(\zeta(2))-H &= \log(\prod_{k=1}^{\infty}\frac{p_k^2}{p_k^2-1})-H\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k^2}{(p_k^2-1)})\right)-H\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k^2}{(p_k-1)\cdot(p_k+1)})\right)-H\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1})+\log(\frac{p_k}{p_k+1})\right)-H\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1})-\log(\frac{p_k+1}{p_k})\right)-H\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1})-\log(1+\frac{1}{p_k})\right)-\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1})-\frac{1}{p_k}\right)\\
&=\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1})-\log(1+\frac{1}{p_k})-\log(\frac{p_k}{p_k-1})+\frac{1}{p_k}\right)\\
&=\sum_{k=1}^{\infty}\left(\frac{1}{p_k}-\log(1+\frac{1}{p_k})\right)
\end{aligned}$$

by Propositions 1 and 3. □

Lemma 2.

$$\sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1}) - \log(1+\frac{1}{p_{k+1}})\right) = \log(\zeta(2)) + \log(\frac{3}{2}).$$

Proof. We obtain that

$$\begin{aligned}
\log(\zeta(2)) + \log(\frac{3}{2}) &= \log(\zeta(2)) - H + H + \log(\frac{3}{2}) \\
&= \sum_{k=1}^{\infty}\left(\frac{1}{p_k} - \log(1+\frac{1}{p_k})\right) + H + \log(\frac{3}{2}) \\
&= \sum_{k=1}^{\infty}\left(\frac{1}{p_k} - \log(1+\frac{1}{p_k})\right) + \sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1}) - \frac{1}{p_k}\right) + \log(\frac{3}{2}) \\
&= \sum_{k=1}^{\infty}\left(\frac{1}{p_k} - \log(1+\frac{1}{p_k}) + \log(\frac{p_k}{p_k-1}) - \frac{1}{p_k}\right) + \log(\frac{3}{2}) \\
&= \sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1}) - \log(1+\frac{1}{p_k})\right) + \log(\frac{3}{2}) \\
&= \sum_{k=1}^{\infty}\left(\log(\frac{p_k}{p_k-1}) - \log(1+\frac{1}{p_{k+1}})\right)
\end{aligned}$$

by Lemma 1. □

3. Partial Sums

Lemma 3.

$$\sum_{p_k \leq x}\left(\frac{1}{p_k} - \log(1+\frac{1}{p_k})\right) = \log(\prod_{p_k \leq x}\frac{p_k^2}{p_k^2-1}) - H + u(x).$$

Proof. We obtain that

$$\log(\prod_{p_k \leq x} \frac{p_k^2}{p_k^2 - 1}) - H + u(x) = \sum_{p_k \leq x} \left(\log(\frac{p_k^2}{(p_k^2 - 1)}) \right) - H + u(x)$$

$$= \sum_{p_k \leq x} \left(\log(\frac{p_k^2}{(p_k - 1) \cdot (p_k + 1)}) \right) - H + u(x)$$

$$= \sum_{p_k \leq x} \left(\log(\frac{p_k}{p_k - 1}) + \log(\frac{p_k}{p_k + 1}) \right) - H + u(x)$$

$$= \sum_{p_k \leq x} \left(\log(\frac{p_k}{p_k - 1}) - \log(\frac{p_k + 1}{p_k}) \right) - H + u(x)$$

$$= \sum_{p_k \leq x} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_k}) \right) - \sum_{p_k \leq x} \left(\log(\frac{p_k}{p_k - 1}) - \frac{1}{p_k} \right)$$

$$= \sum_{p_k \leq x} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_k}) - \log(\frac{p_k}{p_k - 1}) + \frac{1}{p_k} \right)$$

$$= \sum_{p_k \leq x} \left(\frac{1}{p_k} - \log(1 + \frac{1}{p_k}) \right)$$

by Propositions 1 and 3. □

Lemma 4.

$$\sum_{p_k < p_n} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}}) \right) = \log(\frac{3}{2}) + \log(\prod_{p_k \leq p_{n-1}} \frac{p_k^2}{p_k^2 - 1}) - \log(1 + \frac{1}{p_n}).$$

Proof. We obtain that

$$\log(\frac{3}{2}) + \log(\prod_{p_k \leq p_{n-1}} \frac{p_k^2}{p_k^2 - 1}) - \log(1 + \frac{1}{p_n})$$

$$= \log(\frac{3}{2}) + \log(\prod_{p_k \leq p_{n-1}} \frac{p_k^2}{p_k^2 - 1}) - H + u(p_{n-1}) + H - u(p_{n-1}) - \log(1 + \frac{1}{p_n})$$

$$= \log(\frac{3}{2}) + \sum_{p_k \leq p_{n-1}} \left(\frac{1}{p_k} - \log(1 + \frac{1}{p_k})\right) + H - u(p_{n-1}) - \log(1 + \frac{1}{p_n})$$

$$= \log(\frac{3}{2}) + \sum_{p_k \leq p_{n-1}} \left(\frac{1}{p_k} - \log(1 + \frac{1}{p_k})\right) + \sum_{p_k \leq p_{n-1}} \left(\log(\frac{p_k}{p_k - 1}) - \frac{1}{p_k}\right) - \log(1 + \frac{1}{p_n})$$

$$= \log(\frac{3}{2}) + \sum_{p_k \leq p_{n-1}} \left(\frac{1}{p_k} - \log(1 + \frac{1}{p_k}) + \log(\frac{p_k}{p_k - 1}) - \frac{1}{p_k}\right) - \log(1 + \frac{1}{p_n})$$

$$= \log(\frac{3}{2}) + \sum_{p_k \leq p_{n-1}} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_k})\right) - \log(1 + \frac{1}{p_n})$$

$$= \sum_{p_k < p_n} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}})\right)$$

by Lemma 3. □

4. Main Insight

Lemma 5.

$$\sum_{p_k \geq p_n} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}})\right) = \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n} \frac{p_k^2}{p_k^2 - 1}).$$

Proof. We obtain that

$$\sum_{p_k \geq p_n} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}}) \right)$$

$$= \sum_{k=1}^{\infty} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}}) \right) - \sum_{p_k < p_n} \left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}}) \right)$$

$$= \log(\zeta(2)) + \log(\frac{3}{2}) - \log(\frac{3}{2}) - \log(\prod_{p_k \leq p_{n-1}} \frac{p_k^2}{p_k^2 - 1}) + \log(1 + \frac{1}{p_n})$$

$$= \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n} \frac{p_k^2}{p_k^2 - 1})$$

by Lemmas 2 and 4. □

5. Proof of Theorem 1

Suppose that the twin prime conjecture is false. Then, there would exist a sufficiently large prime number p_n such that for all prime gaps starting from $p_n + 2$, this implies that they are greater than or equal to 4. In addition, we assume that $p_n + 2$ is also prime. We know that

$$\sum_{p_k > p_n} H_4(p_k) > 0$$

due to Proposition 5. By Proposition 7, we have

$$\left(\log(\frac{p_n}{p_n - 1}) - \frac{1}{p_n - 0.5} + \frac{1}{p_n + 4} \right) + \sum_{p_k > p_n} H_4(p_k) > 0.$$

That is equivalent to

$$-\frac{1}{p_n - 0.5} + \sum_{p_k \geq p_n} \left(\log(\frac{p_k}{p_k - 1}) - \frac{1}{p_{k+1}} \right) + \sum_{p_k > p_n} \log(\frac{p_k^2 - \sqrt{p_k}^{\log(p_k) + 1}}{p_k^2}) > 0$$

since $-\frac{1}{p_{k+1}} \geq -\frac{1}{p_k + 4}$ under our assumption. Moreover, we obtain that

$$-\frac{1}{p_n - 0.5} + \sum_{p_k \geq p_n}\left(\log(\frac{p_k}{p_k - 1}) - \log(1 + \frac{1}{p_{k+1}})\right) + \sum_{p_k > p_n}\log(\frac{p_k^2 - \sqrt[\log(p_k)+1]{p_k}}{p_k^2}) > 0$$

since $-\log(1 + \frac{1}{p_{k+1}}) \geq -\frac{1}{p_{k+1}}$ by Proposition 6. By Lemma 5, we deduce

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n}\frac{p_k^2}{p_k^2 - 1}) + \sum_{p_k > p_n}\log(\frac{p_k^2 - \sqrt[\log(p_k)+1]{p_k}}{p_k^2}) > 0.$$

So,

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n}\frac{p_k^2}{p_k^2 - 1}) + \sum_{p_k > p_n}\log(\frac{p_k^2 - (\sqrt[\log(p_k)+1]{p_k})^{\frac{\log(p_k)}{\log(p_k)}}}{p_k^2}) > 0$$

which is

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n}\frac{p_k^2}{p_k^2 - 1}) + \sum_{p_k > p_n}\log(\frac{p_k^2 - \sqrt[\frac{\log(p_k)+1}{\log(p_k)}]{e}}{p_k^2}) > 0$$

since $x^{\frac{1}{\log x}} = e$ for $x > -1$. Hence, it is enough to show that

$$\sum_{p_k > p_n}\log(\frac{p_k^2 - 2.54}{p_k^2}) > \sum_{p_k > p_n}\log(\frac{p_k^2 - \sqrt[\frac{\log(p_k)+1}{\log(p_k)}]{e}}{p_k^2})$$

for a prime number $p_n > 10^6$. Finally, we obtain that

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) + \log(\prod_{p_k \geq p_n}\frac{p_k^2}{p_k^2 - 1}) + \sum_{p_k > p_n}\log(\frac{p_k^2 - 2.54}{p_k^2}) > 0$$

which is

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) + \log(\frac{p_n^2}{p_n^2 - 1}) + \sum_{p_k > p_n} \log(\frac{p_k^2 - 2.54}{p_k^2 - 1}) > 0.$$

However, the previous inequality is trivially false, because of

$$\log(\frac{p_n^2}{p_n^2 - 1}) < \frac{1}{p_n^2 - 0.6} < \sum_{p_k > p_n} \log(\frac{p_k^2}{p_k^2 - 1}) < \sum_{p_k > p_n} \log(\frac{p_k^2 - 1}{p_k^2 - 2.54})$$

when $p_{n+1} = p_n + 2$ and

$$-\frac{1}{p_n - 0.5} + \log(1 + \frac{1}{p_n}) < -\frac{1}{p_n - 0.5} + \frac{1}{p_n + 0.4} < 0$$

by Proposition 2 and 7. Hence, the inequality

$$\sum_{p_k > p_n} H_4(p_k) > 0$$

would not hold by transitivity. For that reason, we obtain a contradiction under the supposition that the twin prime conjecture is false. By reductio ad absurdum, we prove that the twin prime conjecture is true. □

References

1. ^*R. Ayoub, Euler and the zeta function. The American Mathematical Monthly 81(10), 1067–1086 (1974). https://doi.org/10.2307/2319041*
2. ^*J.L. Nicolas, The sum of divisors function and the Riemann hypothe- sis. The Ramanujan Journal pp. 1–45 (2021). https://doi.org/10.1007/s11139-021-00491-y*
3. ^*F. Mertens, Ein Beitrag zur analytischen Zahlentheorie. J. reine angew. Math. 1874(78), 46–62 (1874). https://doi.org/10.1515/crll.1874.78.46*
4. ^*Y. Choie, N. Lichiardopol, P. Moree, P. Sol´e, On Robin's criterion for the Riemann hypothesis. Journal de Th´eorie des Nombres de Bordeaux 19(2), 357–372 (2007). https://doi.org/10.5802/jtnb.591*
5. ^*J.L. Nicolas, Petites valeurs de la fonction d'Euler. Journal of num- ber theory 17(3), 375–388 (1983). https://doi.org/10.1016/0022-314X(83)90055-0*
6. ^*Equation Solver - Wolfram Alpha. Roots for the function H in the value of a = 2. https://www.wolframalpha.com/input?i2d=true&i= log%5C%2840%29Divide%5BX%2C%5C%2840%29X+-+1%5C%2841% 29%5D%5C%2841%29+-*

+Divide%5B1%2C%5C%2840%29X+%2B+ 2%5C%2841%29%5D+%2B+log%5C%2840%29Divide%5B%5C%2840% 29Power%5BX%2C2%5D+- +Power%5BX%2C%5C%2840%29Divide% 5B1%2C%5C%2840%29log%5C%2840%29X%5C%2841%29+%2B+1% 5C%2841%29%5D%5C%2841%29%5D%5C%2841%29%2CPower%5BX% 2C2%5D%5D%5C%2841%29%3D0. Accessed 15 November 2022

7. ^*Equation Solver - Wolfram Alpha. Roots for the function H in the value of a = 4. https://www.wolframalpha.com/input?i2d=true&i= log%5C%2840%29Divide%5BX%2C%5C%2840%29X+-+1%5C%2841% 29%5D%5C%2841%29+- +Divide%5B1%2C%5C%2840%29X+%2B+ 4%5C%2841%29%5D+%2B+log%5C%2840%29Divide%5B%5C%2840% 29Power%5BX%2C2%5D+- +Power%5BX%2C%5C%2840%29Divide% 5B1%2C%5C%2840%29log%5C%2840%29X%5C%2841%29+%2B+1% 5C%2841%29%5D%5C%2841%29%5D%5C%2841%29%2CPower%5BX% 2C2%5D%5D%5C%2841%29%3D0. Accessed 15 November 2022*
8. ^*L. Kozma. Useful Inequalities. http://www.lkozma.net/inequalities_cheat_sheet/ineq.pdf (2022). Accessed 15 November 2022*
9. ^*A. Ghosh, An asymptotic formula for the Chebyshev theta function. Notes on Number Theory and Discrete Mathematics 25(4), 1–7 (2019). https://doi.org/10.7546/nntdm.2019.25.4.1-7*