The smallest gap between primes

Frank Vega[1*]

[1*]Software Department, CopSonic, 1471 Route de Saint-Nauphary, Montauban, 82000, Tarn-et-Garonne, France.

Corresponding author(s). E-mail(s): vega.frank@gmail.com;

Abstract

A prime gap is the difference between two successive prime numbers. Two is the smallest possible gap between primes. 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. This conjecture has been one of the great open problems in number theory for many years. 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 as the main theoretical background. In this note, using arithmetic operations, we prove that the twin prime conjecture is true. Indeed, this is a trivial and short note very easy to check and understand which is a breakthrough result at the same time.

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).

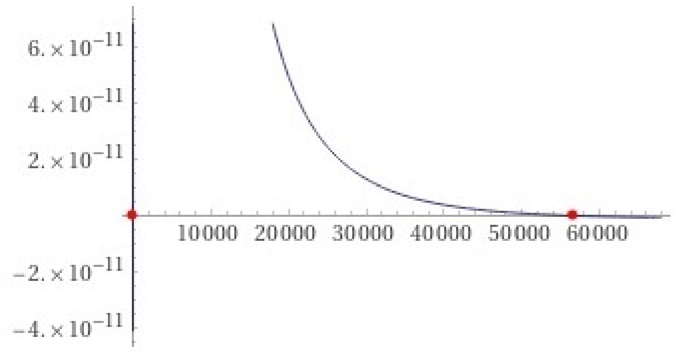

Fig. 1 Roots of $H_2(x)$ [4]

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.).

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 [2, (17.) pp. 54].

Proposition 2 *We have [2, (17.) pp. 54]:*

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

where log *is the natural logarithm.*

For $x \geq 2$, the function $u(x)$ is defined as follows [3, pp. 379]:

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

We use the following function:

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

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

We state the following Propositions:

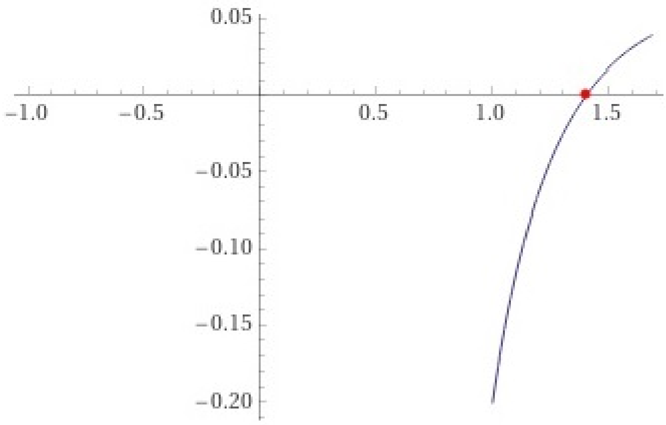

Fig. 2 Roots of $H_4(x)$ [5]

Proposition 3 *For a sufficiently large positive value x, we have $H_2(x) < 0$. Note that, for a sufficiently large positive value x the function $H_2(x)$ is negative and strictly decreasing (its derivative is lesser than 0 for large enough x) and its greatest root is between 50000 and 60000 (See Figure 1).*

Proposition 4 *For a sufficiently large positive value x, we have $H_4(x) > 0$. Note that, for a sufficiently large positive value x the function $H_4(x)$ is positive and strictly increasing (its derivative is greater than 0 for large enough x) and its unique root is between 1.4 and 1.5 (See Figure 2).*

The following property is based on natural logarithms:

Proposition 5 *[6, pp. 1]. For $x > -1$:*

$$\frac{x}{x+1} \leq \log(1+x) \leq x.$$

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 2. □

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

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

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

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

by Propositions 1 and 2. □

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

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

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

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

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

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

Proof 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, they are greater than or equal to 4. First, we need prove that

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

That is the same as

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

We know that

$$\log(1+\frac{1}{p_n}) \leq \frac{1}{p_n} = \frac{1}{p_n-1} - \frac{1}{p_n \cdot (p_n-1)}$$

by Proposition 5. Moreover, we have

$$\log(\frac{p_k^2 - \sqrt[\log(p_k)+1]{p_k}}{p_k^2 - 1}) = \log(1 - \frac{\sqrt[\log(p_k)+1]{p_k} - 1}{p_k^2 - 1})$$

$$\geq \frac{-\frac{\sqrt[\log(p_k)+1]{p_k}-1}{p_k^2-1}}{-\frac{\sqrt[\log(p_k)+1]{p_k}-1}{p_k^2-1} + 1}$$

$$= -\frac{\sqrt[\log(p_k)+1]{p_k} - 1}{-\sqrt[\log(p_k)+1]{p_k} + 1 + p_k^2 - 1}$$

$$= -\frac{\sqrt[\log(p_k)+1]{p_k} - 1}{p_k^2 - \sqrt[\log(p_k)+1]{p_k}}$$

by Proposition 5. Using the well-known properties of Geometric series, we infer that

$$\frac{1}{p_n - 1} = \sum_{i=0}^{\infty} \frac{1}{p_n^{i+1}}.$$

For all integers $i \geq 2$, we see that $\frac{\sqrt[\log(p_{n+i})+1]{p_{n+i}}-1}{p_{n+i}^2 - \sqrt[\log(p_{n+i})+1]{p_{n+i}}}$ is much greater than $\frac{1}{p_n^{i+1}}$ when p_n is large enough. Indeed, we have

$$\frac{\sqrt[\log(p_{n+i})+1]{p_{n+i}} - 1}{p_{n+i}^2 - \sqrt[\log(p_{n+i})+1]{p_{n+i}}} \gg \frac{1}{p_n^{i+1}}$$

for every natural number $i \geq 2$ and large enough p_n. The remaining values $i \in \{0, 1\}$ contains a very small expression deeply close to 0:

$$\frac{1}{p_n} + \frac{1}{p_n^2} - \frac{1}{p_n \cdot (p_n - 1)} - \sum_{i=0}^{1} \frac{\sqrt[\log(p_{n+i})+1]{p_{n+i}} - 1}{p_{n+i}^2 - \sqrt[\log(p_{n+i})+1]{p_{n+i}}}.$$

Consequently, we can assure that

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

for large enough p_n. In this way, we have

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

by Lemma 5. We verify that

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

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

since $-\log(1 + \frac{1}{p_{k+1}}) \geq -\frac{1}{p_{k+1}}$ by Proposition 5. Under our assumption, we notice that

$$\sum_{p_k \geq p_n} \left(\log(\frac{p_k}{p_k - 1}) - \frac{1}{p_{k+1}}\right) + \sum_{p_k \geq p_n} \log(\frac{p_k^2 - \sqrt[\log(p_k)+1]{p_k}}{p_k^2}) \geq \sum_{p_k \geq p_n} H_4(p_k)$$

since $-\frac{1}{p_{k+1}} \geq -\frac{1}{p_k+4}$. However, we know that

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

due to Proposition 4. Hence, the inequality

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

would not hold by transitivity. For that reason, we obtain a contradiction under the supposition that the twin prime conjecture is false. Certainly, we require the negative values of $H_2(p_k)$ for large enough prime numbers p_k in order to consider our satisfied inequality according to the Proposition 3. 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] 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

[3] J.L. Nicolas, Petites valeurs de la fonction d'Euler. Journal of number theory **17**(3), 375–388 (1983). https://doi.org/10.1016/0022-314X(83)90055-0

[4] 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 10 November 2022

[5] 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 10 November 2022

[6] L. Kozma. Useful Inequalities. http://www.lkozma.net/inequalities_cheat_sheet/ineq.pdf (2022). Accessed 10 November 2022