ON THE EXISTENCE OF TWIN PRIME IN AN INTERVAL

SHAON SAHHOO

Abstract. Let \( S_{[x,y]} = \{ \frac{p_n}{p_{n+1}} : n \in I \} \), where \( I = \{ n : x < p_n \leq y \} \), \( p_n \) is the \( n \)-th prime and \( x, y \in \mathbb{R}_{>0} \). If \( M_\alpha(x,y) \) denotes the \( \alpha \)-power mean of the elements of \( S_{[x,y]} \), it is shown that the existence of a twin prime pair in \( (x, y] \) is implied if \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,y) > 1 - 2/y + O(y^{-2}) \) for a sufficiently large \( y \). For a special choice of \( y \), we also find a lower bound for the mean: \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,x^\beta) > 1 - c/x^\beta + O(x^{-\beta \log^{-1} x}) \), where the constant \( c > 0 \) and \( \beta = 1 + c/\log^2 x \) or equivalently, \( x^\beta = x + cx/\log x + O(x/\log^2 x) \). With \( c < 2 \), the lower bound for \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,x^\beta) \) satisfies the inequality on the existence of a twin prime in the interval \((x,x^\beta]\).

1. Introduction

Let \( p_n \) be the \( n \)-th prime and \( x, y \in \mathbb{R}_{>0} \) with \( x < y \). Let \( S_{[x,y]} = \{ r_n : n \in I \} \), where \( r_n = \frac{p_n}{p_{n+1}} \) and \( I = \{ n : x < p_n \leq y \} \). The power mean or Hölder mean of the elements of \( S_{[x,y]} \) is given by \( M_\alpha(x,y) = \left( \frac{1}{\pi(x,y)} \sum_{n \in I} r_n^\alpha \right)^{\frac{1}{\alpha}}, \) where \( \alpha \in \mathbb{R} \) and \( \pi(x,y) \) denotes the number of primes in \((x,y]\). By construction, \( r_n \leq 1 \) for all \( n \in I \), and \( r_n = 1 \) only when \( p_{n+1} - p_n = 2 \). By Theorem 3, \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,y) \) converges to the largest element in \( S_{[x,y]} \). Therefore, \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,y) = 1 \) implies that there exists a twin prime in \((x,y]\). In fact, as stated in Theorem 2, the existence of a twin prime pair in the interval \((x,y]\) is implied if \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,y) > 1 - 2/y + O(y^{-2}) \) for a sufficiently large \( y \).

In this work we show that, for a special \( x \)-dependent choice of \( y \), a lower bound for the mean \( \lim \alpha \rightarrow \infty \) \( M_\alpha(x,y) \) is actually larger than \( 1 - 2/y + O(y^{-2}) \) for a sufficiently large \( y \) (or \( x \)). The result on the lower bound is stated as the following theorem.

**Theorem 1.** Let \( S_{[x,y]} = \{ \frac{p_n}{p_{n+1}} : n \in I \} \) for \( I = \{ n : x < p_n \leq y \} \), where \( p_n \) is the \( n \)-th prime. With \( r_n = \frac{p_n}{p_{n+1}} \), the mean of the elements of \( S_{[x,y]} \) is given by \( M_\alpha(x,y) = \left( \frac{1}{\pi(x,y)} \sum_{n \in I} r_n^\alpha \right)^{\frac{1}{\alpha}} \), where \( \alpha \) is a real number and \( \pi(x,y) \) denotes the number of primes in the interval \((x,y]\). We have

\[
(1.1) \quad \lim_{\alpha \rightarrow \infty} M_\alpha(x,x^\beta) > 1 - \frac{c}{x^\beta} + O\left(\frac{1}{x^\beta \log x}\right),
\]

where \( c \) is a fixed positive constant \( (c > 0) \) and \( \beta = 1 + c/\log^2 x \) or equivalently, \( x^\beta = x + cx/\log x + O(x/\log^2 x) \).

2010 Mathematics Subject Classification. 11N05.
Key words and phrases. Twin primes, Power mean.
The proof of Theorem 1 is given in Section 4. If we take \( c < 2 \) in Theorem 1, the inequality in Theorem 2 is satisfied for the interval \((x, x^2)\): 
\[
\lim_{\alpha \to \infty} M_{\alpha}(x, x^2) > 1 - \frac{c}{2^\alpha} + O\left(\frac{1}{x^{\log_2 x}}\right) > 1 - \frac{2}{x} + O\left(\frac{1}{x^{2\log x}}\right).
\]
This shows that there exists a twin prime in the interval \((x, x + \frac{c}{\log x} + O\left(\frac{x}{\log x}\right))\) for sufficiently large \( x \) and \( 0 < c < 2 \). If a twin prime exists in an interval, it also exists in an extension of the interval. We, accordingly, conclude that a twin prime exists in the aforementioned interval for any fixed \( c > 0 \).

### 2. Background

It is conjectured that there are infinite number of primes \( p \), for each of which \( p + 2 \) is also a prime. Many important contributions have been made in the efforts to prove the conjecture. In this regard we may note the Chen’s work [1] and more recent work by Tao (through Polymath projects) and Maynard brought down the upper bound of 70 million to 246 [3, 4]. It is still an open problem whether the upper bound of the limit can be brought down to the conjectured value of 2.

We take here a different approach to study the question on the twin primes.

### 3. Some useful results

**Theorem 2.** Let \( S_{(x,y)} \) and \( M_{\alpha}(x, y) \) respectively be the set and the power mean of the elements of the set, as defined in Theorem 1. The existence of a twin prime pair in the interval \((x, y)\) is implied if 
\[
\lim_{\alpha \to \infty} M_{\alpha}(x, y) > 1 - \frac{2}{y} + O\left(\frac{1}{y^2}\right)
\]
for a sufficiently large \( y \).

**Proof.** Taking \( g_n = p_{n+1} - p_n \), we rewrite the elements of \( S_{(x,y)} \) as \( r_n = \frac{1}{y(n+1)} \). We note that \( r_n \leq 1 \); the equality sign holds only when \( g_n = 2 \), i.e., when \( p_n \) represents (the first member of) a twin prime. We, therefore, see that 1 is the largest possible element of \( S_{(x,y)} \) and 1 \( \in S_{(x,y)} \) implies the existence of twin prime in \((x, y)\).

Let \( R \) be the largest element in \( S_{(x,y)} \) which is not 1. A non-trivial upper bound of \( R \) corresponds to \( g_n = 4 \) and the largest prime in \((x, y)\). If \( P \) is the largest prime in \((x, y)\), we have \( R \leq \frac{1-2\tau}{1+2\tau} = 1 - 2/(y + O(y^{-2})) \) for sufficiently large \( y \). By Theorem 3, \( \lim_{\alpha \to \infty} M_{\alpha}(x, y) \) converges to the largest element in \( S_{(x,y)} \). Therefore, the inequality \( \lim_{\alpha \to \infty} M_{\alpha}(x, y) > 1 - 2/y + O(y^{-2}) \) implies that there is an element in \( S_{(x,y)} \) which is larger than \( R \). But 1 can be the only element which is larger than \( R \). Hence \( \lim_{\alpha \to \infty} M_{\alpha}(x, y) > 1 - 2/y + O(y^{-2}) \) actually confirms that 1 \( \in S_{(x,y)} \). The statement of the lemma now follows. \( \square \)

**Theorem 3.** Let \( A = \{a_1, a_2, \ldots, a_N\} \) be a set of \( N \) positive real numbers. Let \( \max\{A\} \) and \( \min\{A\} \) denote respectively the largest and the smallest elements in \( A \). The \( \alpha \)-power mean of the elements of \( A \) is given by 
\[
M_{\alpha}(A) = \left(\frac{1}{N} \sum_{i=1}^{N} a_i^\alpha \right)^{1/\alpha},
\]
where \( \alpha \in \mathbb{R} \). We have
(1) \( M_\alpha(A) \geq M_\alpha'(A) \) for \( \alpha > \alpha' \); the equality sign holds iff all the elements in \( A \) are equal.

(2) \( \lim_{\alpha \to 0} M_\alpha(A) = (\prod_{i=1}^{N} a_i)^{1/N} \).

(3) \( \lim_{\alpha \to \infty} M_\alpha(A) = \max\{A\} \) and \( \lim_{\alpha \to -\infty} M_\alpha(A) = \min\{A\} \).

These are standard results; the proof can be found elsewhere (e.g. [8]).

**Theorem 4.** We have, 
\[
\sum_{p \leq x} \frac{1}{p} = \log \log x + M + O(\exp(-\sqrt{\log x})),
\]
where \( M \approx 0.261 \) is the Meissel-Mertens constant.

This result (Theorem 4) is due to Landau [9]. Results with improved error term are known but are not necessary for our work here. A corollary of this result is the following.

**Corollary 1.** We have, 
\[
\sum_{p \leq x} \frac{1}{p} = \log \log x + M + O(\frac{1}{\log x^{A}}),
\]
for any constant \( A > 0 \).

For our work we need the above result with \( A > 1 \); for definiteness, we will take \( A = 2 \).

**Lemma 1.** We have, 
\[
(1 - \frac{1}{2}) \prod_{2 < p \leq x} (1 - \frac{2}{p}) = \frac{e^{-D}}{\log^2 x} \left( 1 + O(\frac{1}{\log^2 x}) \right),
\]
where \( D \approx 0.877 \) is a constant.

**Proof.** Let \( G(x) = \prod_{2 < p \leq x} (1 - \frac{2}{p}) \). Therefore,
\[
\log G_x = \sum_{2 < p \leq x} \log (1 - \frac{2}{p}) = \sum_{2 < p \leq x} \left[ -\frac{2}{p} - \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k \right].
\]
Since \( \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k \leq \frac{1}{2} \sum_{k=2}^{\infty} \left( \frac{2}{p} \right)^k \leq \frac{2}{p^2(1-2/p)} \leq \frac{6}{p^2} \),
the infinite sum \( \sum_{p > 2} \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k \) converges absolutely to a constant, say, \( C \). By numerical calculation we estimate that \( C = -\sum_{p > 2} \left[ \log(1 - \frac{2}{p}) + \frac{2}{p} \right] \approx 0.660413 \). Moreover, \( \sum_{2 < p \leq x} \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k = \sum_{p > 2} \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k - \sum_{p > x} \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k \). Here the first term is the constant \( C \) and the second term \( \sum_{p > x} \sum_{k=2}^{\infty} \frac{1}{k} \left( \frac{2}{p} \right)^k \leq \sum_{p > x} \frac{6}{p^2} = O(\frac{1}{x}) \). With this, we have
\[
\log G(x) = -2 \sum_{2 < p \leq x} \frac{1}{p} - C + O(\frac{1}{x}).
\]
Lemma 1. Let $M$ be the smallest prime greater than $x$, then 

$$\frac{1}{2} \prod_{2<p\leq x} (1 - \frac{\beta}{p}) = \frac{1}{2} G_x = \frac{e^{-D}}{\log^2 x} (1 + O(\frac{1}{\log^2 x})).$$

Finally we get

$$T(1 + \frac{1}{2}) \prod_{2<p\leq x} (1 - \frac{e^{-D}}{\log^2 x}) = \frac{1}{2} G_x = \frac{e^{-D}}{\log^2 x} (1 + O(\frac{1}{\log^2 x})).$$

This gives us $G(1 + \frac{1}{2}) \prod_{2<p\leq x} (1 - \frac{\beta}{p}) = \frac{1}{2} G_x = \frac{e^{-D}}{\log^2 x} (1 + O(\frac{1}{\log^2 x})).$

Finally we get

$$T(1 + \frac{1}{2}) \prod_{2<p\leq x} (1 - \frac{\beta}{p}) = \frac{1}{2} G_x = \frac{e^{-D}}{\log^2 x} (1 + O(\frac{1}{\log^2 x})).$$

where $D = D' + \log 2 \approx 0.876554$. \hfill $\Box$

**Lemma 2.** Let $T(x, x^\beta) = \prod_{n\in I} \frac{p_n}{p_n+1 - 2}$, where $p_n$ is the $n$-th prime and $I = \{ n : x < p_n \leq x^\beta \}$. Here $\beta > 1$ and additionally, let $\beta$ also depend on $x$ in the following way: $\beta = O(1)$. We have $T(x, x^\beta) = \frac{e^{-D}}{x^{\beta-1}} \left( 1 + O(\frac{1}{\log^2 x}) \right)$.

**Proof.** Let $p_s$ be the smallest prime greater than $x$ and similarly, $p_c$ be the smallest prime greater than $x^\beta$.

Now, $T(x, x^\beta) = \prod_{n\in I} \frac{p_n}{p_n+1 - 2} = \left( \frac{p_s - 2}{p_s - 2} \right) \left( \frac{1}{2} \right) \prod_{2<p\leq x} \left( \frac{1 - \frac{\beta}{p}}{1 - \frac{\beta}{p}} \right)$. Using Lemma 1 we get,

$$T(x, x^\beta) = \left( \frac{p_s - 2}{p_s - 2} \right) \frac{e^{-D} \log^{-2} x (1 + O(\log^{-2} x))}{e^{-D} \log^{-2} x \beta (1 + O(\log^{-2} x))} = \frac{\beta^2 p_s - 2}{p_c - 2} \left( 1 + O(\frac{1}{\log^2 x}) \right).$$

We know from the prime number theorem that the average gap between the consecutive primes up to $x$ is $(1 + o(1)) \log x$. In fact it is conjectured, by Cramér [5] and later with some refinement by Granville [6], that, if $G(x)$ is the largest gap between two consecutive primes up to $x$, then $G(x) = O(\log^2 x)$. Therefore, one would expect that $p_s = x + O(\log^t x)$ and $p_c = x^\beta + O(\log^t x)$, for some $t \geq 1$.

In this regard, currently the best unconditional result is due to Baker, Harman and Pintz [7]. They showed that $G(x) = O(x^{\theta})$, where $\theta = 0.525$. This result is sufficient to prove the lemma here. Taking $p_s = x + O(x^{\theta})$ and $p_c = x^\beta + O(x^{\theta})$ in Equation 11 we get,

$$T(x, x^\beta) = \frac{\beta^2 x (1 + O(\frac{1}{\log^2 x}))}{x^\beta (1 + O(\frac{1}{\log^2 x}))} \left( 1 + O(\frac{1}{\log^2 x}) \right) = \frac{\beta^2}{x^{\beta-1}} \left( 1 + O(\frac{1}{\log^2 x}) \right).$$

\hfill $\Box$

4. **Proof of Theorem 1**

**Proof.** By Theorem 3 $M_\alpha(x, x^\beta) \geq M_{\alpha'}(x, x^\beta)$ for $\alpha > \alpha'$. For the time being, we take $\beta > 1$; its appropriate form will be mentioned soon. We are here, in particular, interested in the inequality $M_\infty(x, x^\beta) \geq M_0(x, x^\beta)$, where $M_\infty(x, x^\beta)$ and $M_0(x, x^\beta)$ respectively denote $\lim_{\alpha \to \infty} M_\alpha(x, x^\beta)$ and $\lim_{\alpha \to 0} M_\alpha(x, x^\beta)$. In the
following we first argue that a strict inequality holds between the two means, i.e., \( M\zeta(x, x^\beta) > M_0(x, x^\beta) \), and then we show that \( M_0(x, x^\beta) = 1 - c/x^\beta + O(x^{-\beta \log^{-1} x}) \), where \( c > 0 \) and \( \beta = 1 + c/\log^2 x \) or equivalently, \( x^\beta = x + cx/\log x + O(x/\log^2 x) \).

For the first part, we now note that \( p_n \leq p_{n+1} - 2 \) and, by Bertrand-Chebyshev theorem, \( 2p_n > p_{n+1} > p_{n+1} - 2 \). Therefore, \( p_n \leq p_{n+1} - 2 < 2p_n \). This implies that \( \gcd(p_n, p_{n+1} - 2) = 1 \), except when \( p_n = p_{n+1} - 2 \). This observation helps us to determine if two elements of the set \( S_{x, x^\beta} \) are equal or not. Since the elements of \( S_{x, x^\beta} \) are in the form \( \frac{p_n}{p_{n+1} - 2} (= r_n) \), the numerators of all the elements are different, except when the elements correspond to twin primes. Equivalently, we have \( r_n \neq r_m \) whenever \( n \neq m \), with an exception where both \( p_n \) and \( p_m \) are (first members of) twin primes. Since all the consecutive primes do not form twin prime pairs, we conclude that all elements in \( S_{x, x^\beta} \) are not equal, and hence by Theorem 3 \( M_\infty(x, x^\beta) > M_0(x, x^\beta) \).

Next we analyze the mean \( M_0(x, x^\beta) \). Let \( \pi(x, x^\beta) \) denote the number of primes in \( (x, x^\beta) \). If \( \pi(x) \) denotes the number of primes upto \( x \), we have \( \pi(x) = \frac{x}{\log x} (1 + O(\frac{1}{\log x})) \). Therefore, \( \pi(x, x^\beta) = \pi(x^\beta) - \pi(x) = \frac{x^\beta}{\beta \log x} (1 + O(\frac{1}{\log x})) \). Using Lemma 2 we now have

\[
M_0(x, x^\beta) = T(x, x^\beta) \frac{\beta}{\pi(x, x^\beta)} \frac{1}{\pi(x, x^\beta)}
\]

\[
= \left( \frac{\beta^2}{x^{\beta^2-1}} \right) \frac{1}{\pi(x, x^\beta)} \left( 1 + O \left( \frac{1}{\pi(x, x^\beta) \log^2 x} \right) \right)
\]

Now to determine the main term, let \( z = \left( \frac{\beta^2}{x^{\beta^2}} \right) \frac{1}{\pi(x, x^\beta)} \). Taking logarithm on both sides, we have

\[
\log z = \frac{2 \log \beta - (\beta - 1) \log x}{\pi(x, x^\beta)}
\]

\[
= \frac{2 \beta \log \beta \log x - \beta (\beta - 1) \log^2 x}{x^\beta} \left( 1 + O \left( \frac{1}{\log x} \right) \right).
\]

At this stage, we take \( \beta = 1 + \frac{c}{\log x} \), for a fixed \( c > 0 \). Accordingly, we have \( \beta \log \beta = \frac{c}{\log x} + O(\frac{1}{\log x}) \), \( \beta (\beta - 1) = \frac{c}{\log x} + O(\frac{1}{\log x}) \) and \( x^\beta = x(x^\beta)^{\frac{c}{\log x}} = x(1 + \frac{c}{\log x} + O(\frac{1}{\log x})) \). We now get from Equation 4.2

\[
\log z = \left( \frac{2c}{x^\beta \log x} - \frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log^2 x} \right) \right) \left( 1 + O \left( \frac{1}{\log x} \right) \right)
\]

\[
= -\frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log x} \right).
\]

This gives us \( z = 1 - \frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log x} \right) \). Plugging this value of \( z \) in Equation 4.1 we get \( M_0(x, x^\beta) = \left( 1 - \frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log x} \right) \right) \left( 1 + O \left( \frac{1}{x^\beta \log x} \right) \right) = 1 - \frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log x} \right) \).

Finally we get \( M_\infty(x, x^\beta) > M_0(x, x^\beta) = 1 - \frac{c}{x^\beta} + O \left( \frac{1}{x^\beta \log x} \right) \). \( \square \)
References

1. Jing-run Chen, On the representation of a larger even integer as the sum of a prime and the product of at most two primes, *Scientia Sinica* **XVI** (1973), 157-176.
2. Y. Zhang, Bounded gaps between primes, *Ann. of Math.* **179** (2014), 1121–1174.
3. DHJ Polymath, Variants of the Selberg sieve, and bounded intervals containing many primes, *Res. Math. Sci.* **1** (2014), Art. 12, 1-83.
4. J. Maynard, Small gaps between primes, *Ann. of Math.* **181** (2015), 383-413.
5. H. Cramér, On the order of magnitude of the difference between consecutive prime numbers, *Acta Arith.* **2** (1936), 396-403.
6. A. Granville, Harald Cramér and the distribution of prime numbers, *Scand. Actuar. J.* (1995), 12-28.
7. R. C. Baker, G. Harman, and J. Pintz, The difference between consecutive primes. II, *Proc. Lond. Math. Soc.* **83** (2001), 532-562.
8. J. M. Steele, The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities, Cambridge University Press, New York, 2004.
9. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen, Teubner, Leipzig, 1909. Reprinted: Chelsea, New York, 1953.

Indian Institute of Technology, Tirupati 517506, India

Email address: shaon@iittp.ac.in