The prime Number Theorem and Prime Gaps

Hyeon Jun Ahn

June 1, 2022

Abstract: Let there exists $m > 0$ such that $g_n = O((logp_n)^m)$, then

$$\forall k > 0,\ \exists M \in \mathbb{N} \quad s.t.\ \ n \geq M \ \Rightarrow\ g_n := p_{n+1} - p_n < p_n^k$$

where p_n is nth prime number, O is big O notation, log is natural logarithm. This lead to a corollary for Andrica conjecture, Oppermann conjecture.

1. Introduction

By the prime number theorem, primes less than n are asymptotically $\frac{n}{nlogn}$, so the average gap between primes less than n is $logn$. Hence, nth prime is asymptotically $nlogn$.

$$i.e.\ \lim_{n\to\infty} \frac{p_n}{nlogn} = 1$$

This is equivalent to

$$p_n \sim nlogn$$

This means that $nlogn$ approximates p_n in the sense that the relative error of this approximation approaches 0 as n increases without bound. So,

$$p_{n+1} + p_n \sim (n+1)log(n+1) + nlogn$$

because

$$\lim_{n\to\infty} \frac{p_{n+1} + p_n}{(n+1)log(n+1) + nlogn}$$

$$= \lim_{n\to\infty} \left(\frac{1}{(n+1)log(n+1)/p_{n+1} + nlogn/p_{n+1}} + \frac{1}{(n+1)log(n+1)/p_n + nlogn/p_n}\right)$$

$$= \lim_{n\to\infty} \left(\frac{1}{1+1} + \frac{1}{1+1}\right) = 1$$

This result shows it is possible to add $p_n \sim nlogn$ and $p_{n+1} \sim (n+1)log(n+1)$. But

$$p_n - p_{n+1} \nsim (n+1)log(n+1) - nlogn \tag{1}$$

Rather,

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{(n+1)log(n+1) - nlogn} = \infty \ \ and \ \ \liminf_{n\to\infty} \frac{p_{n+1} - p_n}{(n+1)log(n+1) - nlogn} = 0$$

proof. Note that

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{logp_n} = \infty \ and \ \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{logp_n} = 0 \tag{2}$$

E. Westzynthius proved the former in 1931 [1,2], Daniel Goldston, János Pintz and Cem Yıldırım proved the latter in 2005 [3]. And note that the formula

$$\lim_{n\to\infty} \frac{log(nlogn)}{logp_n} = 1 \tag{3}$$

holds. Because, for every $3 \le n \in \mathbb{N}$,

$$\frac{log(nlogn)}{logp_n} = log_{p_n}(nlogn) = k(n) \in \mathbb{R}$$

Then,

$$p_n = (nlogn)^{k(n)} \Rightarrow \frac{p_n}{(nlogn)^{k(n)}} = 1$$

Since $p_n \sim nlogn$,

$$\lim_{n\to\infty} \frac{nlogn}{(nlogn)^{k(n)}} = \lim_{n\to\infty} \frac{p_n}{(nlogn)^{k(n)}} \frac{nlogn}{p_n} = 1 \times 1 = 1$$

$$\therefore \lim_{n\to\infty} k(n) = 1$$

And note that the formula

$$\lim_{n\to\infty} \frac{log(nlogn)}{(n+1)log(n+1) - nlogn} = 1 \tag{4}$$

holds. Because, by L'ospital's rule,

$$\lim_{n\to\infty} \frac{log(nlogn)}{(n+1)log(n+1) - nlogn} = \lim_{n\to\infty} \frac{logn + log(logn)}{(n+1)log(n+1) - nlogn}$$

$$\overset{\mathrm{L'H}}{=} \lim_{n\to\infty} \frac{1/n + 1/nlogn}{log(n+1) - logn} = \lim_{n\to\infty} \frac{logn + 1}{nlogn(log(n+1) - logn)}$$

$$= \lim_{n\to\infty} \frac{logn + 1}{lognlog(1+1/n)^n} = 1 \ \blacksquare$$

Now, let $F(n) = \dfrac{logp_n}{log(nlogn)} \dfrac{log(nlogn)}{(n+1)log(n+1) - nlogn}$, then, due to (3),(4),

$$\lim_{n\to\infty} F(n) = 1$$

$$i.e. \ \exists M \in \mathbb{N} \quad s.t \ \ n \ge M \ \Rightarrow \ \frac{1}{2} < F(n) < \frac{3}{2}$$

$$\Rightarrow \frac{1}{2} \frac{p_{n+1}-p_n}{logp_n} < \frac{p_{n+1}-p_n}{logp_n} F(n) < \frac{3}{2} \frac{p_{n+1}-p_n}{logp_n}$$

$$\therefore \limsup_{n\to\infty} \frac{p_{n+1}-p_n}{logp_n} F(n) = \infty \ and \ \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{logp_n} F(n) = 0$$

Since $\frac{p_{n+1}-p_n}{logp_n} F(n) = \frac{p_{n+1}-p_n}{(n+1)log(n+1) - nlogn}$,

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{(n+1)log(n+1) - nlogn} = \infty \ and \ \liminf_{n\to\infty} \frac{p_{n+1}-p_n}{(n+1)log(n+1) - nlogn} = 0$$

Therefore we need another method to find the approximate expression of $p_{n+1} - p_n$. $(n+1)log(n+1) - nlogn$ is not appropriate though $p_n \sim nlogn$.

Remark: Cramer conjecture is a conjecture regerding the gaps between prime numbers. The conjecture state that

$$g_n := p_{n+1} - p_n = O((logp_n)^2)$$

holds where O is big O notation. And sometimes the following formulation is called Cramer's conjecture

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{(logp_n)^2} = 1$$

which is stronger than former. But Maier's theorem shows that the Cramér random model does not adequately describe the distribution of primes on short intervals, and a refinement of Cramér's model taking into account divisibility by small primes suggests that

$$\limsup_{n\to\infty} \frac{p_{n+1}-p_n}{(logp_n)^2} \geq 2\exp(-\gamma) \approx 1.1229\cdots$$

These conjecture say that the limit sup of $\frac{g_n}{(logp_n)^2}$ converges. János Pintz suggested that the limit sup may diverges,[4] but It is supported that there exists m such that $\frac{g_n}{(logp_n)^m}$ converge by the preceding several heuristics. So, Let μ be The smallest m that satisfies the following conditions:

$$m \in \mathbb{N}, \ \lim_{n\to\infty} \frac{g_n}{(logp_n)^m} = 0 \tag{5}$$

This paper develops on a assumption that there exists μ.

2. Prime gap

Remark 1. For every $k > 0$,

$$\lim_{n\to\infty} \frac{p_n^k}{(nlogn)^k} = \lim_{n\to\infty} (\frac{p_n}{nlogn})^k = 1 \tag{6}$$

Lemma 1. For every $k > 0$

$$\lim_{n\to\infty} \frac{(log(nlogn))^\mu}{(nlogn)^k} = 0 \tag{7}$$

proof. Let $x = nlogn$, By L'ospital's rule,

$$\lim_{x\to\infty} \frac{(logx)^\mu}{x^k} \overset{\text{L'H}}{=} \lim_{x\to\infty} \frac{\mu(logx)^{\mu-1}}{kx^k} \overset{\text{L'H}}{=} \cdots \overset{\text{L'H}}{=} \lim_{x\to\infty} \frac{\mu!}{k^\mu x^k} = 0 \blacksquare$$

Due to (6), for every $k > 0$

$$\lim_{n\to\infty} \frac{p_{n+1} - p_n}{p_n^k} \frac{(nlogn)^k}{(log(nlogn))^\mu} = \lim_{n\to\infty} \frac{p_{n+1} - p_n}{(log(nlogn))^\mu} \frac{(nlogn)^k}{p_n^k} = 0 \times 1 = 0 \quad (8)$$

Hence,

$$\lim_{n\to\infty} \frac{p_{n+1} - p_n}{p_n^k} = (7) \times (8) = 0 \quad (9)$$

$$\Leftrightarrow \lim_{n\to\infty} \frac{p_n^k}{p_{n+1} - p_n} = \infty$$

By epsilon-delta argument,

$$\begin{aligned} \forall k > 0,\ \exists N \in \mathbb{N} \quad s.t. \quad n \geq N \ &\Rightarrow\ g_n := p_{n+1} - p_n < p_n^k \\ &\Rightarrow\ p_n < p_{n+1} < p_n + p_n^k \end{aligned} \quad (10)$$

3. About Andrica conjeutre

Andrica conjecture is a conjecture regarding the gaps between prime numbers. The conjecture states that the inequality

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

holds for all $n \in \mathbb{N}$. And a strong version of Andrica conjecture is as follows: Excert for $p_n \in \{3, 7, 13, 23, 31, 113\}$, that is $n \in \{2, 4, 6, 9, 11, 30\}$, one has

$$\sqrt{p_{n+1}} - \sqrt{p_n} < \frac{1}{2}; \qquad \textit{equivalently} \qquad g_n := p_{n+1} - p_n < p_n^{1/2} + \frac{1}{4}$$

And This paper proves that

$$\lim_{n\to\infty} (\sqrt{p_{n+1}} - \sqrt{p_n}) = 0$$

proof. Let $\epsilon \in (0, \frac{1}{2})$, $k \in (0, \frac{1}{2})$, Then

$$\lim_{n\to\infty} \frac{p_n^k}{(\sqrt{p_n} + \epsilon)^2 - p_n} = \lim_{n\to\infty} \frac{p_n^k}{2\epsilon\sqrt{p_n} + \epsilon^2} = 0$$

Thus,

$$\begin{aligned} \exists N_1 \in \mathbb{N} \quad s.t. \quad n > N_1 \ &\Rightarrow\ p_n^k < (\sqrt{p_n} + \epsilon)^2 - p_n \\ &\Rightarrow\ p_n + p_n^k < (\sqrt{p_n} + \epsilon)^2 \end{aligned}$$

Meanwhile,

$$\exists N_2 \in \mathbb{N} \quad s.t. \;\; n > N_2 \;\Rightarrow\; p_{n+1} < p_n + p_n^k \qquad (\because (10))$$

Let N=max(N_1, N_2), Then

$$n > N \;\Rightarrow\; p_{n+1} < (\sqrt{p_n} + \epsilon)^2 \;\Rightarrow\; \sqrt{p_{n+1}} - \sqrt{p_n} < \epsilon$$

By epsilon-delta argument,

$$\lim_{n\to\infty} (\sqrt{p_{n+1}} - \sqrt{p_n}) = 0 \;\blacksquare \tag{11}$$

Furthermore, let $y > 1,\; x < \frac{y-1}{y}$. Then, since $\forall\epsilon > 0,\; \exists M \in \mathbb{N} \quad s.t.\; n > M$ $\Rightarrow |p_n^{1/y}| > |\epsilon|$, by generalized binomial theorem,

$$\lim_{n\to\infty} \frac{p_n^x}{(p_n^{1/y} + \epsilon)^y - p_n}$$

$$= \lim_{n\to\infty} \frac{p_n^x}{(p_n + \binom{y}{1} p_n^{(y-1)/y}\epsilon + \binom{y}{2} p_n^{(y-2)/y}\epsilon^2 + \cdots) - p_n}$$

$$= \lim_{n\to\infty} \frac{p_n^x}{(\binom{y}{1} p_n^{(y-1)/y}\epsilon + \binom{y}{2} p_n^{(y-2)/y}\epsilon^2 + \cdots)} = 0 \; (\because x < \frac{y-1}{y})$$

In the same method as the proof of (11),

$$\forall y > 1, \quad \lim_{n\to\infty} (p_{n+1}^{1/y} - p_n^{1/y}) = 0$$

3-1. The arithmetic mean, the geometric mean and harmonic mean of primes

The relation between the arithmetic mean and the geometric mean of nth prime and $(n+1)$th prime is as follows:

$$\frac{p_{n+1} + p_n}{2} \sim \sqrt{p_{n+1} p_n}$$

proof.

$$\begin{aligned} &\lim_{n\to\infty} (\sqrt{p_{n+1}} - \sqrt{p_n}) = 0 \\ &\Rightarrow \lim_{n\to\infty} (\sqrt{p_{n+1}} - \sqrt{p_n})^2 = 0 \\ &\Rightarrow \lim_{n\to\infty} (p_{n+1} + p_n - 2\sqrt{p_{n+1} p_n}) = 0 \end{aligned} \tag{12}$$

Thus,

$$\lim_{n\to\infty} \frac{p_{n+1} + p_n}{2\sqrt{p_{n+1} p_n}} = \lim_{n\to\infty} (\frac{p_{n+1} + p_n - 2\sqrt{p_{n+1} p_n}}{2\sqrt{p_{n+1} p_n}} + 1) = 1 \;\blacksquare \tag{13}$$

Furthermore,

$$\lim_{n\to\infty}(\frac{p_{n+1}+p_n}{2}-\sqrt{p_{n+1}p_n})=0$$

trivially holds by (12). And similarly, the relation between the arithmetic mean and the harmonic mean of nth prime and $(n+1)$th prime is as follows:

$$\frac{p_{n+1}+p_n}{2}\sim\frac{2p_{n+1}p_n}{p_{n+1}+p_n}$$

proof. By (13)

$$\lim_{n\to\infty}\frac{2p_{n+1}p_n}{p_{n+1}+p_n}\frac{2}{p_{n+1}+p_n}=\lim_{n\to\infty}(\frac{2\sqrt{p_{n+1}p_n}}{p_{n+1}+p_n})^2=1\ \blacksquare$$

Furthermore,

$$\lim_{n\to\infty}(\frac{p_{n+1}+p_n}{2}-\frac{2p_{n+1}p_n}{p_{n+1}+p_n})=0$$

proof.

$$\lim_{n\to\infty}(\frac{p_{n+1}+p_n}{2}-\frac{2p_{n+1}p_n}{p_{n+1}+p_n})$$

$$=\lim_{n\to\infty}\frac{(p_{n+1}+p_n)^2-4p_{n+1}p_n}{2(p_{n+1}+p_n)}=\lim_{n\to\infty}\frac{(p_{n+1}-p_n)^2}{2(p_{n+1}+p_n)}$$

$$\leq\lim_{n\to\infty}\frac{(p_{n+1}-p_n)^2}{4p_n}=\lim_{n\to\infty}(\frac{p_{n+1}-p_n}{2\sqrt{p_n}})^2=0\ (\because (9))$$

By the relation between the geometric mean and the harmonic mean,

$$\lim_{n\to\infty}(\frac{p_{n+1}+p_n}{2}-\frac{2p_{n+1}p_n}{p_{n+1}p_n})=0\ \blacksquare$$

Hence,

$$\frac{p_{n+1}+p_n}{2}\sim\sqrt{p_{n+1}p_n}\sim\frac{2p_{n+1}p_n}{p_{n+1}p_n}$$

Furthermore, the arithmetic mean, the geometric mean, and the harmonic mean of nth prime and $(n+1)$th prime become asymptotically the same as n increases without bound.

4. About Oppermann conjecture

Oppermann conjecture is a conjecture regarding the distribution of prime numbers. It is closely related to but stronger than Legendre conjecture, Andrica conjecture, and Brocard conjecture. The conjecture states that for every integer $n\geq 1$,

$$\pi(n^2-n)<\pi(n^2)<\pi(n^2+n)$$

Definition 1. Let $\hat{p}(x)$ is the nearest prime less than x, $\hat{P}(x)$ is the nearest prime more than x.

$$e.g.\ \ \hat{p}(10)=7,\ \hat{P}(10)=11$$

Lemma 2. Let $f\colon \mathbb{R} \to \mathbb{R}$ is an increasing function and m is constant, then

$$p_n < p_{n+1} < f(p_n) \quad \Rightarrow \quad \exists p \in \mathbb{P} \quad with \quad x < p < f(x)$$

proof (by contradiction). Let $\exists x \in \mathbb{R}$ such that $\nexists p \in \mathbb{P}$ in $(x, f(x))$, then $\hat{P}(x) > f(x)$. And By definition, $\hat{p}(x) \leq x$ and $\hat{P}(x)$ is the next prime of $\hat{p}(x)$. thus,

$$\hat{p}(x) < \hat{P}(x) < f(\hat{p}(x))$$

But, because f is an increasing function, $\hat{p}(x) \leq x \ \Rightarrow \ f(\hat{p}(x)) \leq f(x) < \hat{P}(x)$. It's contradiction. ■

Lemma 3. By **Lemma 2**, (10) is equivalent to

$$\forall k > 0,\ \exists M_1 \in \mathbb{R}, \quad s.t.\ \ \exists p \in \mathbb{P} \ \ with \ \ x < p < x + x^k \ \ for\ x \geq M_1 \tag{14}$$

Lemma 4.

$$\forall k > 0,\ \exists M_2 \in \mathbb{R}, \quad s.t.\ \ \exists p \in \mathbb{P} \ \ with \ \ x - x^k < p < x \ \ for\ x \geq M_2 \tag{15}$$

proof. In **Lemma 3**, let $x = m + m^k$, then there is a prime in the open interval (m, x). Since $x > m \ \Rightarrow \ x^k > m^k$, $(m, x) \subset (x - x^k, x)$. Hence, there is a prime in the open interval $(x - x^k, x)$. ($c.f.$ $M_1 < M_2$) ■

This psaper proves that for every $k > 0$, there exists $M \in \mathbb{R}$ such that

$$x \geq M \ \Rightarrow \ \pi(x^k - x) < \pi(x) < \pi(x^k + x) \tag{16}$$

proof. By (14),(15),

$$\forall k > 0,\ \exists M_2 \in \mathbb{R}, \quad s.t.\ \exists p, q \in \mathbb{P}\ with\ x - x^k < p < x < q < x + x^k\ \ for\ x \geq M_2$$

Let $x = t^m$ where $m = \dfrac{1}{k}$, then

$$\forall m > 0,\ \exists M' \in \mathbb{R}, \quad s.t.\ \exists p, q \in \mathbb{P}\ with\ t^m - t < p < t^m < q < t^m + t\ \ for\ t \geq M'$$

($c.f.$ $x = t^m \Rightarrow M' = M_2^k$) This fomula implies that

$$\forall m > 0,\ \exists M' \in \mathbb{R} \quad s.t.\ \ t \geq M' \ \Rightarrow \ \pi(t^m - t) < \pi(t^m) < \pi(t^m + t) \ ■$$

Furthermore, how many primes exist in $(x^k, x^k + x)$? In other word, what is the result of $\lim_{x\to\infty} (\pi(x^k + x) - \pi(x^k))$?

Remark 2.

$$f_1 \sim g_1 \ \wedge \ f_2 \sim g_2 \ \rightarrow \ f_1 - f_2 \sim g_1 - g_2$$

doesn't always hold. (1) is a counterexample. Due to this,

$$\lim_{x\to\infty} \frac{\pi(x^m + x) - \pi(x^m)}{(x^m + x)/log(x^m + x) - x^m/log(x^m)} = 1$$

may not hold. We need other method.

Lemma 5. for function f and g such that $\forall x \in \mathbb{R},\ g(x) > f(x) > 0$, if $\lim_{x\to\infty}(g(x)-f(x)) = \infty$ and there exists $k \in (0,1)$ such that $g(x)^k < g(x)-f(x)$ for sufficiently large x, then

$$\lim_{x\to\infty}(\pi(g(x)) - \pi(f(x))) = \infty$$

proof. Because of (15),

$$\begin{aligned}\forall j \in (0,k),\ \exists N \in \mathbb{R} \quad s.t.\ \ x \geq N \ &\Rightarrow\ \exists p \in \mathbb{P} \ \ with \ \ g(x) - g(x)^j < p < g(x) \\ &\Rightarrow\ \exists p \in \mathbb{P} \ \ with \ \ f(x) < p < g(x)\end{aligned}$$

Let $a_1 = g(x),\ a_{n+1} = a_n - a_n^j$, then there exists a prime in the open interval $(a_n - a_n^j, a_n) = (a_{n+1}, a_n)$ and for every $n \in \mathbb{N},\ a_1 \geq a_n$.
Let $f(x) < a_m,\ f(x) > a_{m+1}$, then $\pi(g(x)) - \pi(f(x)) \geq m-1$. Therefore, for sufficiently large x,

$$g(x) - f(x) < \sum_{n=1}^{m}(a_n - a_{n+1}) = \sum_{n=1}^{m} a_n^j < \sum_{n=1}^{m} a_1^j = m a_1^j$$

$$\Rightarrow\ m > \frac{g(x)-f(x)}{a_1^j} = \frac{g(x)-f(x)}{g(x)^j} > \frac{g(x)^k}{g(x)^j}$$

Note that

$$\lim_{x\to\infty}\frac{g(x)^k}{g(x)^j} = \infty\ (\because j \in (0,k))$$

Hence,

$$\lim_{x\to\infty}(\pi(g(x)) - \pi(f(x))) = \infty\ \blacksquare$$

Since $\forall x \in \mathbb{R},\ (x + x^m) > x^m > 0$ and for sufficiently large x, every $m > 0$, there exists $k \in (0,1)$ such that $(x^m + x)^k < (x^m + x) - x^m = x$,

$$\forall m > 0\ \lim_{x\to\infty}(\pi(x^m + x) - \pi(x^m)) = \infty$$

References

[1] E. Westzynthius. Über die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind. Commentationes Helsingfors 5, Nr. 25, 37 S. (1931)., 1931.

[2] Erik Westzynthius. On the distribution of numbers that are not divisible by the first n primes. *Commentat. Phys.-Math.*, 5(25):1–37, 1931.

[3] DA Goldston and J Pintz. Y. yıldırım, c.: Primes in tuples i. *Ann. Math., AIM Preprint*, 19, 2005.

[4] János Pintz. Very large gaps between consecutive primes. *Journal of Number Theory*, 63(2):286–301, 1997.

Keyword: Prime; Prime gap; Andrica conjecture; Cramer conjecture; Oppermann conjecture; Arithmetic mean of primes; Geometric mean of primes; Harmonic mean of primes

E-mail: fghwer1051@gmail.com