Elementary Proof of the Prime Number Theorem

A Brief Sketch

Shuhang Xue, Yilong Song

Dr. Caroline Turnage-Butterbaugh
Introduction to Analytic Number Theory
Carleton College
June 7, 2021

Abstract

Motivated by the bounds on the Chebyshev function $\psi(x)$ and Merten's Theorems, we present an elementary proof of the Prime Number Theorem (PNT) by proving its equivalent statement that $\psi(x)$ is asymptotic to x. In our proof, we rely on the general Mobius inversion formula to derive the Selberg's inequality. Then, to smooth the arithmetic function and amplify the contribution of primes in the Selberg's inequality, we introduce a smoother $S(y)$ and an enhancer $W(x)$. By studying properties of $W(x)$, we prove the PNT in its equivalent form $\alpha = \limsup_{x\to\infty} |W(x)| = 0$.

Contents

1 Introduction and Preliminary Results

The prime number theorem, what formalizes the intuitive understanding that the occurrences of prime numbers become less frequent as one counts the natural numbers toward positive infinity, was conjectured by Gauss in the early 1790s and Legendre toward the end of the same decade independently according to historical record and Gauss' own recollection. Its proof came in 1896 due to Hadamard and de la Vallée Poussin again independently after a century of mathematical inquiry surrounding the topic, to which Chebychef and Riemann, among other great mathematicians, contributed important partial results. This paper provides a version of an elementary proof of said theorem derived from Erdös and Selberg's paper completed in 1949, sharpened by Breusch, Bombieri, Wirsing, and others, and organized by Levinson [1], upon whose paper the following account mainly bases.

In this section, we introduce a number of preliminary ideas and results that are useful for the sections that follow.

Theorem 1.1 (The prime number theorem). *Let $\pi(x)$ denote the number of primes not exceeding x (i.e. $\pi(x) = \sum_{p \le x} 1$), then*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x} = 1.$$

An equivalent version of the above can be written as such:

Theorem 1.2 (Equivalent form of the prime number theorem with Chebychef's ψ-function). *Define $\psi(x) = \sum_{n \le x} \Lambda(n)$ where $\Lambda(x)$ is the von Mangoldt function (see section 2.8 in Apostol* [2] *for detailed definition). Then*

$$\lim_{x\to\infty} \frac{\psi(x)}{x} = 1. \tag{1}$$

A proof for the equivalency of Theorems 1.1 and 1.2 (i.e. why $\psi(x)$ behaves like $\pi(x)\log x$) will be given under Lemma 1.6.

Now we explore more properties of $\psi(x)$. We can start by constructing an arithmetic function $T(x)$ defined as such:

$$T(x) = \sum_{n \le x} \log n.$$

Then by definition of Chebychef's ψ-function and the von Mangolt function, we may observe the following:

$$\begin{aligned}
T(x) &= \sum_{j|x} \Lambda(j) \\
&= \sum_{ij \le x} \Lambda(j) \\
&= \sum_{i \le x} \sum_{j \le x/i} \Lambda(j) \\
&= \sum_{i \le x} \psi(x/i).
\end{aligned} \tag{2}$$

And the result of (2) is Chebychef's identity. We then attempt to write $\psi(x)$ in terms of $T(x)$:

$$\begin{aligned}
T(x) &= \sum_{i \le x} \psi(x/i) \\
&= \psi(x) + \psi(x/2) + \cdots + \psi(x/\lfloor x \rfloor) \\
\psi(x) &= T(x) - \psi(x/2) - \cdots - \psi(x/\lfloor x \rfloor)
\end{aligned} \tag{3}$$

To eliminate the ψ terms on the right hand side, we can construct terms $T(x/2)$, $T(x/3)$, ..., $T(x/\lfloor x \rfloor)$, try subtracting them one by one from $T(x)$ (e.g. $T(x) - T(x/2)$, then $T(x) - T(x/2) - T(x/3)$, etc.), and compare the results to said right hand side. We observe that whether or not $T(x/i)$ ought to be subtracted, not subtracted, or added to make $T(x) - T(x/2) - \cdots \pm T(x/\lfloor x \rfloor) = T(x) - \psi(x/2) - \cdots - \psi(x/\lfloor x \rfloor)$ corresponds to the value of $\mu(x/i)$ (see definition of the Möbius function μ in section 2.2 of Apostol), because if we subtract $T(x/i)$ if $\mu(x/i) = -1$ (so that we can begin to eliminate ψ terms), we need not subtract nor add $T(x/i)$ if $\mu(x/i) = 0$ because since x/i is not square free, all the ψ terms it has has already been eliminated in previous subtractions, and we need to add $T(x/i)$ if $\mu(x/i) = 1$ to account for its ψ terms being subtracted an additional time earlier. Thus we get the relation:

$$\psi(x) = \sum_{k \le x} \mu(k) T\left(\frac{x}{k}\right). \tag{4}$$

Using the same technique and the definition of T from (2), we obtain also:

$$\Lambda(n) = \sum_{k|n} \mu(k) \log n/k, \tag{5}$$

where $n \ge 1$. An additional observation can be made about $T(x)$:

Lemma 1.3.

$$T(x) = x \log x - x + O(\log x).$$

Proof. This can be obtained directly by applying Abel's summation.

With these, we can obtain some more properties involving these two arithmetic functions:

Lemma 1.4.

$$\sum_{n \le x} \Lambda(n)/n = \log x + O(1).$$

.

Proof. By (1), we have:

$$\begin{aligned}
T(x) &= \sum_{j\le x}\Lambda(j)\sum_{i\le x/j} 1 \\
&= \sum_{j\le x}\Lambda(j)\lfloor\frac{x}{j}\rfloor \\
&= x\sum_{j\le x}\frac{\Lambda(j)}{j} - \sum_{j\le x}\Lambda(j)\left\{\frac{x}{j}\right\} \\
&= x\sum_{j\le x}\frac{\Lambda(j)}{j} - O(x).
\end{aligned}$$

By Lemma 1.3, we obtain the desired result. ■

Lemma 1.5.

$$\psi(x) = O(x).$$

Proof. See lecture 19[3]. ■

Lemma 1.6.

$$\psi(x) = \pi(x)\log x + O\left(\frac{x\log\log x}{\log x}\right).$$

Proof. By definition of ψ, we have

$$\psi(x) = \sum_{p\le x}\log p + \sum_{p\le x^{1/2}}\log p + \sum_{p\le x^{1/3}}\log p + \cdots. \tag{6}$$

Where each sum is not zero if and only if $x^{1/j} \ge 2$, which is equivalent to saying $j \le \log x/\log 2$. Thus

$$\begin{aligned}
\psi(x) &\le \sum_{p\le x}\log p + \frac{\log x}{\log 2}\sum_{p\le x^{1/2}}\log p \\
&\le \log x\pi(x) + \frac{\log x}{\log 2}\pi(x^{1/2})\log x^{1/2} \\
&\le \log x\pi(x) + \frac{x^{1/2}\log^2 x}{2\log 2}.
\end{aligned} \tag{7}$$

Again by (6), we see that since $\pi(y) \le y$:

$$\begin{aligned}
\psi(x) &\ge \sum_{x/log^2 x<p\le x}\log p \ge \log\left(\frac{x}{\log^2 x}\right)\sum_{x/\log^2 x<p\le x} 1 \\
&= \log\left(\frac{x}{\log^2 x}\right)\left(\pi(x) - \pi\left(\frac{x}{\log^2 x}\right)\right) \\
\frac{\psi(x)}{\log x - 2\log\log x} &\ge \pi(x) - \frac{x}{\log^2 x} \\
\pi(x)\log x &\le \psi(x)\frac{\log x}{\log x - 2\log\log x} + \frac{x}{\log x} \\
&= \psi(x) + \psi(x)\frac{2\log\log x}{\log x - 2\log\log x} + \frac{x}{\log x}.
\end{aligned}$$

Since $2\log\log x < \log x/4$ and by Lemma 1.5:

$$\pi(x)\log x \le \psi(x) + O\left(\frac{x\log\log x}{\log x}\right).$$

Combined with (7), we have the desired result. ■

Lemma 1.7.

$$\sum_{n\leq x}\frac{1}{n} = \log x + \gamma + O(1/x),$$

where γ *is Euler's constant.*

Proof. See Apostol section 3.4.

2 Selberg's Elementary Inequality

Now we use (4) (Möbius inversion formula) to find how $\psi(x)$ behaves for large x. Notice, first, that the argument presented after (3) and the conclusion drawn at (4) can be rewritten more generally by simply replacing $\psi(x)$ and $T(x)$ with $F(x)$ and $G(x)$. More explicitly, we have:

$$G(x) = \sum_{n\leq x} F(x/n)$$

and

$$F(x) = \sum_{k\leq x}\mu(k)G\left(\frac{x}{k}\right). \tag{8}$$

To simplify the computation for our current endeavor, we try to find a simple $F(x)$ with a transform $G(x)$ that is close to $T(x)$. Then in that case we can subtract (8) from (4):

$$\psi(x) - F(x) = \sum_{k\leq x}\mu(k)(T(\frac{x}{k}) - G(\frac{x}{k})). \tag{9}$$

If the right hand side can be shown to be small, then $\psi(x)$ is also close to $F(x)$. Taking some hint from the intended result (1), we can assume that $\psi(x)$ is close to x for large x, and try setting $F(x) = F_0(x) = x$. Then $G_0(x) = \sum_{n\leq x} F_0(x/n) = x\sum_{n\leq x} n^{-1} = x\log x + \gamma x + O(1)$ by Lemma 1.7, which is not close enough to $T(x)$ as represented in Lemma 1.3. We can therefore refine $F(x)$ such that $F(x) = F_1(x) = x - C$ where C is some constant. Then

$$\begin{aligned} G_1(x) &= x\sum_{n\leq x}\frac{1}{n} - C\sum_{n\leq x} 1 = x\log x + \gamma x + O(1) - C\lfloor x\rfloor \\ &= x\log x - (C-\gamma)x + O(1). \end{aligned}$$

If we set $C = 1+\gamma$,

$$T(x) - G(x) = O(\log x), \tag{10}$$

which is relatively small. Applying this to (9), we get

$$\psi(x) - x + C = \sum_{k\leq x}\mu(k)\left(T\left(\frac{x}{k}\right) - G_1\left(\frac{x}{k}\right)\right). \tag{11}$$

Since the logarithm grows slower than any positive algebraic power, $\log x = O(x^{1/2})$. Then (10) implies:

$$T(x) - G_1(x) = O(x^{1/2}). \tag{12}$$

Using this and the fact that $|\mu(k)| \leq 1$, and by definition of big O, there is a constant K such that the right hand side of (11) can be written as

$$\begin{aligned}\sum_{k\leq x}\mu(k)\left(T\left(\frac{x}{k}\right)-G_1\left(\frac{x}{k}\right)\right) &= Kx^{1/2}\sum_{k\leq x}k^{-1/2}\\ &< Kx^{1/2}\left(1+\sum_{2\leq k\leq x}\int_{k-1}^{k}u^{-1/2}du\right)\\ &\leq Kx^{1/2}\left(1+\int_{1}^{x}u^{-1/2}du\right)\\ &= O(x).\end{aligned} \tag{13}$$

Which suggests that

$$\psi(x) = O(x). \tag{14}$$

This is a weaker result than the prime number theorem.

Now we can investigate further by rewriting (11) as

$$F(x) = \sum_{k\leq x}\mu(k)G\left(\frac{x}{k}\right), \tag{15}$$

where $F = \psi - x + C$ and $G = T - G_1$. We can increase the terms in the sum on the right side without making asymptotic changes. For example, we can replace the right side with the following expression motivated by the general Möbius inversion formula:

$$J(x) = \sum_{k\leq x}\mu(k)\log\frac{x}{k}G\left(\frac{x}{k}\right). \tag{16}$$

Compute in terms of F from the definition of G:

$$\begin{aligned}J(x) &= \sum_{k\leq x}\mu(k)\log\frac{x}{k}\sum_{j\leq x/k}F\left(\frac{x}{jk}\right)\\ &= \sum_{jk\leq x}\mu(k)\log\frac{x}{k}F\left(\frac{x}{jk}\right)\\ &= \sum_{n\leq x}F\left(\frac{x}{n}\right)\sum_{jk=n}\mu(k)\log\frac{x}{k}\\ &= \sum_{n\leq x}F\left(\frac{x}{n}\right)\sum_{k|n}\mu(k)\log\frac{x}{k}\\ &= \sum_{n\leq x}F\left(\frac{x}{k}\right)\log\frac{x}{n}\sum_{k|n}\mu(k)+\sum_{n\leq x}F\left(\frac{x}{n}\right)\sum_{k|n}\mu(k)\log\frac{n}{k}\end{aligned} \tag{17}$$

since $\log(x/k) = \log(x/n) + \log(n/k)$.
By properties of summing μ and (5), combining with (16), we have

$$F(x)\log x + \sum_{n\leq x}F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{k\leq x}\mu(k)\log\frac{x}{k}G\left(\frac{x}{k}\right). \tag{18}$$

On a different note, by (10)

$$\log x(T(x) - G_1(x)) = O(\log^x) = O(x^{1/2}). \tag{19}$$

And by (13)

$$\sum_{k\leq x}\mu(k)\log\frac{x}{k}\left(T\left(\frac{x}{k}\right)-G_1\left(\frac{x}{k}\right)\right) = O(x). \tag{20}$$

Thus combining with (18), with $F(x) = \psi(x) - x + c$ gives

$$(\psi(x) - x)\log x + \sum_{n\le x}\left(\psi\left(\frac{x}{n}\right) - \frac{x}{n}\right)\Lambda(n) = O(x). \tag{21}$$

By Lemma 1.4, (21) can be written as

$$\psi(x)\log x + \sum_{n\le x}\Lambda(n)\psi(x/n) = 2x\log x + O(x). \tag{22}$$

Applying Abel's summation:

$$\sum_{n\le x}\Lambda(n)\log n = \psi(x)\log x + O(x). \tag{23}$$

And

$$\begin{aligned}\sum_{j\le x}\Lambda(j)\psi\left(\frac{x}{j}\right) &= \sum_{j\le x}\Lambda(j)\sum_{k\le x/j}\Lambda(k)\\ &= \sum_{jk\le x}\Lambda(j)\Lambda(k).\end{aligned} \tag{24}$$

Let $\Lambda_2(n) = \Lambda(n)\log n + \sum_{jk=n}\Lambda(j)\Lambda(k)$, then by (23) and (24), (22) becomes:

$$\sum_{n\le x}\Lambda_2(n)\log n = 2x\log x + O(x). \tag{25}$$

By Lemma 1.3,

$$\sum_{n\le x}\log n = x\log x + O(x).$$

Combining (24) and (25)

$$Q(n) = \sum_{k\le n}(\Lambda_2(k) - 2\log k) = O(n), \tag{26}$$

for $n \le 2$ and $Q(1) = 0$.

3 Enhancement on Selberg's Inequality

In this section, we try to advance our understanding of the Selberg's identity towards proving the Prime Number theorem. First, a direct definition of $R(x) = \psi(x) - x, x \ge 2$ can make the Selberg's identity (Eq. **??**) be rephrased as

$$R(x)\log x + \sum\Lambda(n)R(x/n) = O(x). \tag{27}$$

Thus, to prove the PNT in the form of Eq. (1), it is equivalent to prove the limit of $R(x)$ over x goes to zero as x goes to infinity; i.e.

$$\lim_{x\to\infty}\frac{R(x)}{x} = 0. \tag{28}$$

However, even though $R(x/n) = 0$ when $n > x/2$, the term is self-terminated, we cannot relate this information directly to Eq. (28) to obtain a proof. Instead, we need to smooth $R(x)$ to avoid the existence of "bumps" when x is prime and enhance the contribution of primes, while the information encoded in the Selberg's identity is still preserved. Thus, to realize such end, a smoother $R(x)$ and an enhancer $W(x)$ are defined in the following discussion.

3.1 A smoother $S(y)$

First, as an usual technique in the field of analytic number theory, integrals are used to smooth out discontinuous arithmetic functions. In our case, we define $S(y)$ as

$$S(y) = \int_y^x \frac{R(x)}{x} dx, y \leq 2, \tag{29}$$

when $y \geq 2$. From Levinson [1], to prove Eq. (28) (which is equivalent to proving the PNT), it suffices to show the limit

$$\lim_{x\to\infty} \frac{S(x)}{x} = 0. \tag{30}$$

With this smoother $S(y)$, we first observe that $S(y) = O(y)$ and there is a Lipschitz condition associated with $S(y)$ from the following Lemma.

Lemma 3.1. *For an arbitrary $y \geq 2$,*

$$|S(y)| \leq cy, |S(y_2) - S(y1)| \leq c|y_1 - y_2|,$$

where c is a constant (the second inequality is also known as the Lipschitz condition).

Besides, from the Selberg's inequality in $R(x)$, we can derive

$$S(y)\log y + \sum \Lambda(j)S(y/j) = O(y). \tag{31}$$

Proof. First, to prove that $|S(x)|$ is bounded in general, we first recall the result from Apostol that $\psi(x) = O(x)$. Thus, by definition of $R(y)$, we have

$$\limsup_{x\to\infty} \frac{|R(x)|}{x} \leq 1,$$

which implies $|R(x)| \leq cx$, where c is a constant. Thus, by the definition of $S(y)$, we consider the case when $R(x)$ is continuous and discontinuous.

When $R(x)$ is continuous, $y \neq p^j$, where p is prime, because prime values will make the value of $R(x)$ jump. Then, by taking the derivative of $S(y)$, we have the desired inequality

$$|S'(y)| \leq c, y \neq p^j.$$

The case of discontinuous $R(x)$ at $y = p^j$ is slightly more complicated because it relies on first proving the Lipschitz condition.

By the property of integrals, $S(y)$ is continuous. Then, since the magnitude of a sum is less than or equal to the sum of the magnitudes (known as the general triangle inequality), $|S(y_2) - S(y1)| \leq c|y_1 - y_2|$ true guaranteed by the continuity. Therefore, the case of $y = p^j$ is simply a special case of the Lipschitz condition when $y_1 = 2$:

$$||S(y_2)| - |S(2)|| = |S(y_2)| \leq c|y_2 - 2| \Longrightarrow |S(y)| \leq cy.$$

Combined with the continuous case, $|S(y)|$ is bounded by cy, which is equivalent to $S(y) = O(y)$.

To prove Eq. (30), we first divide Eq. (27) by x integrate both sides to get

$$\int_2^y R(x)/x \log x dx + \sum \Lambda(n) \int_2^y \frac{R(x/n)}{x} dx = O(y). \tag{32}$$

Then, integrate the first part in Eq. (32) by parts

$$\int_2^y R(x)/x \log x dx = \log y S(y) - \int_x^y \frac{S(x)}{x} dx = \log y S(y) + O(y)$$

using the bounds on $|S(y)|$ just proved. Finally, by substitute $u = x/n$, Eq. (32) becomes

$$\log yS(y) + O(y) + \sum \Lambda(n) \int_2^{y/n} \frac{R(u)}{u} du = \log yS(y) + O(y) + \sum \Lambda(j)S(y/j) = O(y),$$

which agrees with the desired Eq. (31). ■

Then, from Levinson's Lemma 5.2, 5.3 [1], and 5.4, the Selberg's inequality can be further smoothed so that the prime contribution becomes clearer and clearer. Note that in Levinson's paper, the author first improve the multiples of $S(y)$ inside the sum from $\Lambda(n)$ to $\log n$, which turns the arithmatic function to a continuous one. Then, the sum is further improved to an integral without losing the power of the Selberg's inequality. As the final result, Levinson showed that there exists a constant K such that

$$\log^2 y|S(y)| \leq 2 \int_2^y |S(y/u)| \log u du + Ky \log y. \tag{33}$$

With this result, we are ready to proceed to an enhancer $W(x)$ as the final tool to prove the PNT.

3.2 An enhancer $W(x)$

Indeed, an enhancer $W(x)$ of $S(y)$ is an exponential version that can amplify the contribution of primes in terms of the prime function itself, while the previous sections targeted enlarge the multiples of the prime function. Specifically, the exponential version of $S(y)$ is defined as

$$W(x) = e^{-x} S(e^x). \tag{34}$$

From Levinson [1], proving the PNT is equivalent to showing $\lim_{x\to\infty} |W(x)| = 0$.

However, we can define

$$\alpha = \limsup_{x\to\infty} |W(x)|. \tag{35}$$

Then, we can simplify such goal as showing $\alpha = 0$ (which only involves investigating the lim-sup). To see why, first notice that $|W(x)|$ is a positive function. Thus, the liminf must be larger than 0. Then, since liminf is smaller than or equal to limsup, as long as limsup (α) is restricted to 0, limsup equal to liminf implies the sequence is convergent and the limit is exactly zero.

Before proving α equals to zero, we first observe that α is bounded by 1 and another quantity based Levinson's Lemma 5.5, which is stated below.

Lemma 3.2 (Levinson, [1]). *Define* $\gamma = \limsup_{x\to\infty} \frac{1}{x} \int_0^x |W(u)| du$. *Then,* $\alpha \leq 1$ *and*

$$\alpha \leq \gamma. \tag{36}$$

The proof is omitted here, but the proof technique is similar to the proof of Lemma 3.1.

Although α can be restricted to be smaller than 1 quickly, to prove that α equals to zero exactly requires a considerable amount of efforts. Thus, we first present two lemmas about $|W(x)|$, which will be helpful to showing the final piece of $\alpha = 0$.

Lemma 3.3. *If* $k = 2c$*, then*

$$|W(x_2) - W(x_1)| \leq k|x_2 - x_1|,$$

which is the Lipschitz condition on $W(x)$.

Moreover, $||W(x_2)| - |W(x_1)|| \leq k|x_2 - x_1|$.

Proof. The proof strategy is similar to the one used to prove Lemma 3.1. We consider the cases when $x = j\log p$ and $x \neq j \log p$. In the first case, by the definition of $W(x)$,

$$|W'(x)| \leq e^{-x}|S(e^x)| + |S'(e^x)|$$

by the product rule of taking derivatives. Then, as we have bounded $|S(y)|$ and $|S'(y)|$ bounded by some constant c in the proof of Lemma 3.1, $|W'(x)|$ is bounded by some constant $k = 2c$ when $x \neq j \log p$.

Thus, by apply the triangle inequality on the cases of $x = j \log p$, we can conclude the Lipschitz condition on $W(x)$.

Finally, by the second triangle inequality $||a| - |b|| \leq |a - b|$,

$$||W(x_2)| - |W(x_1)|| \leq |W(x_2) - W(x_1)| \leq k|x_2 - x_1|.$$

■

Lemma 3.4 (Levinson, [1]). *If $W(v) \neq 0$ for $v_1 < v < v_2$, then there exists a number M such that*

$$\int_{v_1}^{v_2} |W(v)|dv \leq M. \tag{37}$$

The lemma is not proved here, but the result follows from studying $\int_2^x \frac{R(t)}{t^2}dt = O(1)$ using the Abel's Summation.

4 Proof of the Prime Number Theorem

Equipped with Lemma 3.3 and 3.4, we are ready to restrict α to exactly zero to prove the PNT. With all the terminology we have developed, the PNT boils down to the following Lemma.

Lemma 4.1. *With $W(x)$ defined in Eq. 34 and α defined in Eq. 35,*

$$\alpha = 0.$$

Proof. First, based on the definition of limsup, for a fixed $\epsilon > 0$, there exists an x_ϵ such that

$$|W(x)| \leq \epsilon, \qquad x \geq x_\epsilon. \tag{38}$$

Then, if $W(x) \neq 0$ for all large x, then by Lemma 3.4, the integral of $|W(x)|$ can be bounded by some number M. Thus, $\gamma = \limsup_{x\to\infty} \frac{1}{x}\int_0^x |W(u)|du \leq \limsup_{x\to\infty} M \cdot \frac{1}{x} \leq 0$. Thus, by Lemma 3.2, $\alpha = 0$.

Thus to prove the result, it suffices to show that $W(x)$ has arbitrarily large zeros. Let a and b be successive zeros of $W(x)$ for $x > x_\epsilon$. Then consider the following three cases regarding the distribution of zeros.

Case I. Assume $b - a \geq 2M/\epsilon$. By Lemma 3.4, since $W(x) \neq 0$ on the interval (a, b),

$$\int_a^b |W(x)|dx \leq M \leq \frac{1}{2}(b-a)\epsilon.$$

Case II. Assume $b - a \geq 2\epsilon/k$, where k is an integer such that $2\epsilon/k < 2M/\epsilon$. Then, by Lemma 3.3, if we consider the graph of $|W(x)|$, any point must lie below the value $k(b-a)/2 \leq \epsilon$ for $x \in (a, b)$. Thus, we have the inequality

$$\int_a^b |W(x)|dx \leq \frac{1}{2}(b-a)\epsilon.$$

Case III. Assume $2\epsilon/k < b - a < 2M/\epsilon$. Then, using the step as in Case II, bound most points a distance ϵ/k from each end by ϵ, and bound other points by Eq. (38). As a result, we have the following inequality

$$\int_a^b |W(x)|dx \leq \frac{\epsilon^2}{k} + (b - a - \frac{2\epsilon}{k})\epsilon = (b-a)\epsilon\left(1 - \frac{\epsilon}{k(b-a)}\right)$$
$$\leq (b-a)\epsilon(1 - \frac{\epsilon^2}{2Mk}) < (b-a)\epsilon(1 - \frac{\alpha^2}{2Mk}).$$

Moreover, observe that since the values $Mk > 1$ and $\alpha \leq 1$, the inequality we derived in Case I and II also implies the inequality above.

Thus, with the inequality and Lemma 3.4 to bound $|W(x)|$, by assuming the first zero of $W(x)$ as x_1 and the largest zero smaller than y as x_f, we have

$$\int_0^y |W(x)|dx \leq \int_0^{x_1} |W(x)|dx + (x_f - x_1)\epsilon(1 - \frac{\alpha^2}{2Mk}) + M,$$

where the first part integrate nonzero $W(x)$ and the second part bounds the integral between the first zero and last zero by the inequality derived above (note that the result of successive zeros is generalized to any two zeros using the triangle inequality).

Next, by dividing both sides by y and use $x_f \leq y$,

$$\frac{1}{y}\int_0^y |W(x)|dx \leq \frac{1}{y}\int_0^{x_1} |W(x)|dx + \epsilon(1 - \frac{\alpha^2}{2Mk}) + M/y.$$

By assuming y goes to infinity, using the definition of limsup, we have $\gamma \leq \epsilon(1 - \alpha^2/2Mk)$. Therefore, we have the final inequality

$$\alpha \leq \epsilon(1 - \frac{\alpha^2}{2Mk}). \tag{39}$$

Thus, since Eq. (39) holds for arbitrarily small ϵ and $\alpha < \epsilon$, it must hold for $\epsilon = \alpha$. Thus, from Eq. (39), $\alpha^3 \leq 0$. Combined the fact that $\alpha \geq 0$, it must be the case that $\alpha = 0$. ■

Finally, with the last piece of the puzzle ($\alpha = 0$), the long waited conjecture about the distribution of primes can finally be called a theorem. In conclusion, we present an elementary proof of the Prime Number Theorem based on the Selberg's inequality.

References

[1] N. Levinson, A Motivated Account of an Elementary Proof of the Prime Number Theorem. The American Mathematical Monthly, 76(3), 225-245, 1969

[2] T. M. Apostol Introduction to Analytic Number TheorySpringer-Verlag New York 978-0-387-90163-31976.

[3] Caroline Turnage-Butterbaugh, 2021, Introduction to Analytic Number Theory, *Lecture 19*, MATH.395, Carleton College, delivered May 13, 2021.