Selberg's elementary proof of the Prime Number Theorem

S. Das[1]

[1]B.Sc. 3rd year,
Dept. of Mathematics,
Jadavpur University.
Advised by Prof. Satadal Ganguly, ISI Kolkata.

April 2026

Table of Contents

Table of Contents

History

The Prime Number Theorem (PNT) is the pinnacle of classical analytic number theory, and a fundamental result about the distribution of primes.

History

The Prime Number Theorem (PNT) is the pinnacle of classical analytic number theory, and a fundamental result about the distribution of primes.

The original proof of the PNT, due to Hadamard and de la Valée Poussin, uses complex analytic techniques and properties of Riemann's ζ-function.

History

The Prime Number Theorem (PNT) is the pinnacle of classical analytic number theory, and a fundamental result about the distribution of primes.

The original proof of the PNT, due to Hadamard and de la Valée Poussin, uses complex analytic techniques and properties of Riemann's ζ-function.

It was not known whether an elementary proof was possible (Hardy even thought it to be impossible), until Selberg's proof [Sel49] in 1949; Erdős also published a proof around the same time.

Notation

For $f : \mathbb{R} \to \mathbb{C}$, $g : \mathbb{R} \to [0, \infty)$, $f(x) = \mathrm{O}\big(g(x)\big)$ means $|f| \leqslant Cg$ for some constant C,

Notation

For $f : \mathbb{R} \to \mathbb{C}$, $g : \mathbb{R} \to [0, \infty)$, $f(x) = \mathrm{O}\big(g(x)\big)$ means $|f| \leqslant Cg$ for some constant C,
$f(x) = \mathrm{o}\big(g(x)\big)$ means $|f/g| \to 0$ as $x \to \infty$,

Notation

For $f : \mathbb{R} \to \mathbb{C}$, $g : \mathbb{R} \to [0, \infty)$, $f(x) = \mathrm{O}\big(g(x)\big)$ means $|f| \leqslant Cg$ for some constant C,
$f(x) = \mathrm{o}\big(g(x)\big)$ means $|f/g| \to 0$ as $x \to \infty$,
and $f(x) \sim g(x)$ means $|f/g| \to 1$ as $x \to \infty$.

Notation

For $f : \mathbb{R} \to \mathbb{C}$, $g : \mathbb{R} \to [0, \infty)$, $f(x) = \mathrm{O}\big(g(x)\big)$ means $|f| \leqslant Cg$ for some constant C,
$f(x) = \mathrm{o}\big(g(x)\big)$ means $|f/g| \to 0$ as $x \to \infty$,
and $f(x) \sim g(x)$ means $|f/g| \to 1$ as $x \to \infty$.
The notation $\lfloor x \rfloor$ refers to the floor function, the largest integer n such that $n \leqslant x$.

Preliminaries

The PNT is usually stated in terms of the function $\pi(x)$,

Preliminaries

The PNT is usually stated in terms of the function $\pi(x)$,

Definition

The prime counting function $\pi(x) : (0, \infty) \to \mathbb{C}$ is given by

$$\pi(x) = \# \{p : p \leqslant x\}.$$

Preliminaries

The PNT is usually stated in terms of the function $\pi(x)$,

Definition

The prime counting function $\pi(x) : (0, \infty) \to \mathbb{C}$ is given by

$$\pi(x) = \#\{p : p \leqslant x\}.$$

Prime Number Theorem

$$\pi(x) \sim \frac{x}{\log x}.$$

This is the main result that we will prove.

Definition

The Riemann ζ-function is defined by

$$\zeta(s) = \begin{cases} \displaystyle\sum_{n=1}^{\infty} \frac{1}{n^s}, & \text{if } s > 1 \\ \displaystyle\lim_{x \to \infty} \left(\sum_{n \leqslant x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s} \right), & \text{if } 0 < s < 1. \end{cases}$$

Definition

The Riemann ζ-function is defined by

$$\zeta(s) = \begin{cases} \displaystyle\sum_{n=1}^{\infty} \frac{1}{n^s}, & \text{if } s > 1 \\ \displaystyle\lim_{x\to\infty} \left(\sum_{n\leqslant x} \frac{1}{n^s} - \frac{x^{1-s}}{1-s} \right), & \text{if } 0 < s < 1. \end{cases}$$

Definition

The von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{C}$ is given by

$$\Lambda(n) = \begin{cases} \log(p), & \text{if } n = p^k, \text{ where } k \geqslant 1 \text{ is an integer} \\ 0, & \text{otherwise.} \end{cases}$$

Definition

The Chebyshev ϑ-function $\vartheta : (0, \infty) \to \mathbb{C}$ is given by

$$\vartheta(x) = \sum_{p \leqslant x} \log(p).$$

Definition

The Chebyshev ϑ-function $\vartheta : (0, \infty) \to \mathbb{C}$ is given by

$$\vartheta(x) = \sum_{p \leqslant x} \log(p).$$

Definition

The Chebyshev ψ-function $\psi : (0, \infty) \to \mathbb{C}$ is given by

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n).$$

We prove the equivalent Chebyshev form of PNT using Selberg's method.

PNT in Chebyshev form

$$\pi(x) \sim \frac{x}{\log x} \iff \vartheta(x) \sim x \iff \psi(x) \sim x.$$

We prove the equivalent Chebyshev form of PNT using Selberg's method.

PNT in Chebyshev form

$$\pi(x) \sim \frac{x}{\log x} \iff \vartheta(x) \sim x \iff \psi(x) \sim x.$$

Abel summation

For any $a : \mathbb{N} \to \mathbb{C}$ let $A(x) = \sum_{n \leqslant x} a(n)$, where $A(x) = 0$ for all $x < 1$. If f cont. diff. on $[y, x]$, where $0 < y < x$, then

$$\sum_{y < n \leqslant x} a(n) f(n) = A(x) f(x) - A(y) f(y) - \int_y^x A(t) f'(t) \mathrm{d}t.$$

Some useful asymptotic estimates

Theorem

For all $x \geqslant 1$ we have

$$\sum_{n \leqslant x} \psi\left(\frac{x}{n}\right) = x \log x - x + O(\log x).$$

Some useful asymptotic estimates

Theorem

For all $x \geqslant 1$ we have

$$\sum_{n \leqslant x} \psi\left(\frac{x}{n}\right) = x \log x - x + \mathrm{O}(\log x).$$

Using the above estimate and that $\log x < x$, we get

Theorem

$$\psi(x) = \mathrm{O}(x).$$

Tatuzawa and Iseki's theorem

The following theorem due to Tatuzawa and Iseki [TI51] will be used to prove Selberg's asymptotic formula $\psi(x)\log x + \sum_{n\leqslant x}\Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x)$.

Tatuzawa and Iseki's theorem

The following theorem due to Tatuzawa and Iseki [TI51] will be used to prove Selberg's asymptotic formula $\psi(x)\log x + \sum_{n\leqslant x}\Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x)$.

Theorem

Let $F:(0,\infty)\to\mathbb{C}$ and $G(x)=\log x\sum\limits_{n\leqslant x}F\left(\frac{x}{n}\right)$. Then

$$F(x)\log x + \sum_{n\leqslant x}F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{d\leqslant x}\mu(d)G\left(\frac{x}{d}\right).$$

Proof of Tatuzawa and Iseki's theorem

First write $F(x)\log x$ as a sum

$$F(x)\log x = \sum_{n\leqslant x}\left[\frac{1}{n}\right] F\left(\frac{x}{n}\right)\log\frac{x}{n} = \sum_{n\leqslant x} F\left(\frac{x}{n}\right)\log\frac{x}{n}\sum_{d|n}\mu(d)$$

Proof of Tatuzawa and Iseki's theorem

First write $F(x)\log x$ as a sum

$$F(x)\log x = \sum_{n\leqslant x}\left[\frac{1}{n}\right]F\left(\frac{x}{n}\right)\log\frac{x}{n} = \sum_{n\leqslant x}F\left(\frac{x}{n}\right)\log\frac{x}{n}\sum_{d|n}\mu(d)$$

and using the identity $\Lambda(n) = \sum_{d|n}\mu(d)\log\frac{n}{d}$ we can write

$$\sum_{n\leqslant x}F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{n\leqslant x}F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\log\frac{n}{d}.$$

Now adding these two equations we get

$$\begin{aligned} F(x)\log x + \sum_{n\leqslant x} F\left(\frac{x}{n}\right)\Lambda(n) &= \sum_{n\leqslant x} F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\left\{\log\frac{x}{n} + \log\frac{n}{d}\right\} \\ &= \sum_{n\leqslant x} F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\log\frac{x}{d} \\ &= \sum_{n\leqslant x}\sum_{d|n} F\left(\frac{x}{n}\right)\mu(d)\log\frac{x}{d} \text{ write } n = qd \\ &= \sum_{q\leqslant x/d}\sum_{d\leqslant x} F\left(\frac{x}{qd}\right)\mu(d)\log\frac{x}{d} \\ &= \sum_{d\leqslant x}\mu(d)\log\frac{x}{d}\sum_{q\leqslant x/d} F\left(\frac{x}{qd}\right) \\ &= \sum_{d\leqslant x}\mu(d)G\left(\frac{x}{d}\right). \end{aligned}$$

Table of Contents

Overview of the proof

The key lemma is Selberg's asymptotic formula

$$\psi(x)\log x + \sum_{n \leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x).$$

Overview of the proof

The key lemma is Selberg's asymptotic formula

$$\psi(x)\log x + \sum_{n\leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x).$$

It is natural to define a function

$$\sigma(x) = e^{-x}\psi(e^{x}) - 1,$$

Overview of the proof

The key lemma is Selberg's asymptotic formula

$$\psi(x)\log x + \sum_{n \leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x).$$

It is natural to define a function

$$\sigma(x) = e^{-x}\psi(e^x) - 1,$$

then Selberg's formula implies the inequality

$$x^2\left|\sigma(x)\right| \leqslant 2\int_0^x \int_0^y \left|\sigma(u)\right| \mathrm{d}u \, \mathrm{d}y + \mathrm{O}(x). \tag{1}$$

The PNT is then equivalent to the statement: $\sigma(x) \to 0$ as $x \to \infty$. Hence, if we let

$$C = \limsup_{x\to\infty} \left|\sigma(x)\right|, \; K = \limsup_{x\to\infty} \frac{1}{x}\int_0^x \left|\sigma(u)\right| \mathrm{d}u$$

The PNT is then equivalent to the statement: $\sigma(x) \to 0$ as $x \to \infty$. Hence, if we let

$$C = \limsup_{x\to\infty} \left|\sigma(x)\right|, \ K = \limsup_{x\to\infty} \frac{1}{x} \int_0^x \left|\sigma(u)\right| \mathrm{d}u$$

then the PNT is equivalent to showing that $C = 0$. This is proved as follows by assuming towards a contradiction that $C > 0$. From the definition of C and K,

$$\left|\sigma(x)\right| \leqslant C + \mathrm{o}(1), \ \left|\sigma(x)\right| \leqslant K + \mathrm{o}(1) \tag{2}$$

The PNT is then equivalent to the statement: $\sigma(x) \to 0$ as $x \to \infty$. Hence, if we let

$$C = \limsup_{x\to\infty} \left|\sigma(x)\right|, \ K = \limsup_{x\to\infty} \frac{1}{x}\int_0^x \left|\sigma(u)\right| \mathrm{d}u$$

then the PNT is equivalent to showing that $C = 0$. This is proved as follows by assuming towards a contradiction that $C > 0$. From the definition of C and K,

$$\left|\sigma(x)\right| \leqslant C + \mathrm{o}(1), \ \left|\sigma(x)\right| \leqslant K + \mathrm{o}(1) \tag{2}$$

with $C \leqslant K$. If $C > 0$, then this inequality along with (1) yields $K < C$, which is absurd. So $C = 0$.

Lemma (Selberg)

For $x > 0$ we have

$$\psi(x)\log x + \sum_{n\leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + O(x)$$

and

$$\sum_{n\leqslant x} \Lambda(n)\log n + \sum_{mn\leqslant x} \Lambda(m)\Lambda(n) = 2x\log x + O(x).$$

Lemma (Selberg)

For $x > 0$ we have

$$\psi(x)\log x + \sum_{n \leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + \mathrm{O}(x)$$

and

$$\sum_{n \leqslant x} \Lambda(n)\log n + \sum_{mn \leqslant x} \Lambda(m)\Lambda(n) = 2x\log x + \mathrm{O}(x).$$

The two statements above are equivalent as

$$\sum_{n \leqslant x} \Lambda(n)\psi\left(\frac{x}{n}\right) = \sum_{n \leqslant x} \Lambda(n) \sum_{m \leqslant x/n} \Lambda(m) = \sum_{mn \leqslant x} \Lambda(m)\Lambda(n)$$

and using Abel summmation with $a(n) = \Lambda(n)$ and $f(t) = \log t$ we get

$$\sum_{n \leqslant x} \Lambda(n) \log n = \psi(x) \log x - \int_2^x \frac{\psi(t)}{t} \mathrm{d}t = \psi(x) \log x + \mathrm{O}(x).$$

and using Abel summmation with $a(n) = \Lambda(n)$ and $f(t) = \log t$ we get

$$\sum_{n \leqslant x} \Lambda(n) \log n = \psi(x) \log x - \int_2^x \frac{\psi(t)}{t} \mathrm{d}t = \psi(x) \log x + \mathrm{O}(x).$$

Now, we apply Tatuzawa and Iseki's theorem with to the functions $F_1 = \psi(x)$ as well as $F_2(x) = x - \gamma - 1$, where γ is the Euler-Mascheroni constant.

and using Abel summmation with $a(n) = \Lambda(n)$ and $f(t) = \log t$ we get

$$\sum_{n \leqslant x} \Lambda(n) \log n = \psi(x) \log x - \int_2^x \frac{\psi(t)}{t} \mathrm{d}t = \psi(x) \log x + \mathrm{O}(x).$$

Now, we apply Tatuzawa and Iseki's theorem with to the functions $F_1 = \psi(x)$ as well as $F_2(x) = x - \gamma - 1$, where γ is the Euler-Mascheroni constant. For F_1, we have

$$G_1(x) = \log x \sum_{n \leqslant x} \psi\left(\frac{x}{n}\right) = x \log^2 x - x \log x + \mathrm{O}(\log^2 x).$$

For F_2, we have

$$\begin{aligned} G_2(x) &= \log x \sum_{n \leqslant x} \left(\frac{x}{n} - \gamma - 1 \right) \\ &= x \log x \sum_{n \leqslant x} \frac{1}{n} - (\gamma + 1) \log x \sum_{n \leqslant x} 1 \\ &= x \log x \left(\log x + \gamma + \mathrm{O}\left(\frac{1}{x} \right) \right) - (\gamma + 1)\big(x + \mathrm{O}(1)\big) \log x \\ &= x \log^2 x - x \log x + \mathrm{O}(\log x). \end{aligned}$$

For F_2, we have

$$\begin{aligned}
G_2(x) &= \log x \sum_{n \leqslant x} \left(\frac{x}{n} - \gamma - 1 \right) \\
&= x \log x \sum_{n \leqslant x} \frac{1}{n} - (\gamma + 1) \log x \sum_{n \leqslant x} 1 \\
&= x \log x \left(\log x + \gamma + \mathrm{O}\left(\frac{1}{x} \right) \right) - (\gamma + 1) \big(x + \mathrm{O}(1) \big) \log x \\
&= x \log^2 x - x \log x + \mathrm{O}(\log x).
\end{aligned}$$

Hence, $G_1(x) - G_2(x) = \mathrm{O}(\log^2 x)$. Only the weaker estimate $G_1(x) - G_2(x) = \mathrm{O}(\sqrt{x})$ is needed in fact.

Now applying Tatuzawa and Iseki's theorem to F_1 and F_2 we get

$$\{F_1(x) - F_2(x)\} \log x + \sum_{n \leqslant x} \left\{ F_1\left(\frac{x}{n}\right) - F_2\left(\frac{x}{n}\right) \right\} \Lambda(n)$$

$$= \sum_{d \leqslant x} \mu(d) \left\{ G_1\left(\frac{x}{d}\right) - G_2\left(\frac{x}{d}\right) \right\} = \mathrm{O}\left(\sum_{d \leqslant x} \sqrt{\frac{x}{d}} \right).$$

Now applying Tatuzawa and Iseki's theorem to F_1 and F_2 we get

$$\{F_1(x) - F_2(x)\}\log x + \sum_{n\leqslant x}\left\{F_1\left(\frac{x}{n}\right) - F_2\left(\frac{x}{n}\right)\right\}\Lambda(n)$$

$$= \sum_{d\leqslant x}\mu(d)\left\{G_1\left(\frac{x}{d}\right) - G_2\left(\frac{x}{d}\right)\right\} = \mathrm{O}\left(\sum_{d\leqslant x}\sqrt{\frac{x}{d}}\right).$$

Then

$$\{\psi(x) - (x-\gamma-1)\}\log x + \sum_{n\leqslant x}\left\{\psi\left(\frac{x}{n}\right) - \left(\frac{x}{n}-\gamma-1\right)\right\}\Lambda(n)$$

$$= \mathrm{O}\left(\sqrt{x}\sum_{d\leqslant x}\frac{1}{\sqrt{d}}\right) = \mathrm{O}(x).$$

Hence, we get

$$\begin{aligned}
&\psi(x)\log x + \sum_{n\leqslant x}\Lambda(n)\psi\left(\frac{x}{n}\right)\\
&= (x-\gamma-1)\log x + \sum_{n\leqslant x}\left(\frac{x}{n}-\gamma-1\right)\Lambda(n) + \mathrm{O}(x)\\
&= x\log x + x\sum_{n\leqslant x}\frac{\Lambda(n)}{n} - (\gamma+1)\left\{\log x + \sum_{n\leqslant x}\Lambda(n)\right\} + \mathrm{O}(x)\\
&= 2x\log x + \mathrm{O}(1) - 2(\gamma+1)\log x + \mathrm{O}(x)\\
&= 2x\log x + \mathrm{O}(x)
\end{aligned}$$

Hence, we get

$$
\begin{aligned}
&\psi(x)\log x + \sum_{n\leqslant x}\Lambda(n)\psi\left(\frac{x}{n}\right)\\
&= (x-\gamma-1)\log x + \sum_{n\leqslant x}\left(\frac{x}{n}-\gamma-1\right)\Lambda(n) + \mathrm{O}(x)\\
&= x\log x + x\sum_{n\leqslant x}\frac{\Lambda(n)}{n} - (\gamma+1)\left\{\log x + \sum_{n\leqslant x}\Lambda(n)\right\} + \mathrm{O}(x)\\
&= 2x\log x + \mathrm{O}(1) - 2(\gamma+1)\log x + \mathrm{O}(x)\\
&= 2x\log x + \mathrm{O}(x)
\end{aligned}
$$

where the last step is due to the fact that $\log x < x$.

Proof of the PNT in Chebyshev form

Set $\psi(x) = x + R(x)$, then our goal is to show that $R(x) = \mathrm{o}(x)$.

Proof of the PNT in Chebyshev form

Set $\psi(x) = x + R(x)$, then our goal is to show that $R(x) = \mathrm{o}(x)$. Using Selberg's asymptotic formula we get

$$R(x)\log x + \sum_{n \leqslant x} \Lambda(n) R\left(\frac{x}{n}\right) = \mathrm{O}(x).$$

Proof of the PNT in Chebyshev form

Set $\psi(x) = x + R(x)$, then our goal is to show that $R(x) = \mathrm{o}(x)$. Using Selberg's asymptotic formula we get

$$R(x)\log x + \sum_{n \leqslant x} \Lambda(n) R\left(\frac{x}{n}\right) = \mathrm{O}(x).$$

Replacing n by m, x by x/n

$$R\left(\frac{x}{n}\right)\log\left(\frac{x}{n}\right) + \sum_{m \leqslant x/n} \Lambda(m) R\left(\frac{x}{mn}\right) = \mathrm{O}\left(\frac{x}{n}\right).$$

After some manipulations we get

$$|R(x)| \log^2 x \leqslant \sum_{n \leqslant x} a_n \left| R\left(\frac{x}{n}\right) \right| + \mathrm{O}(x \log x) \tag{3}$$

After some manipulations we get

$$|R(x)| \log^2 x \leqslant \sum_{n \leqslant x} a_n \left| R\left(\frac{x}{n}\right) \right| + \mathrm{O}(x \log x) \tag{3}$$

where $a_n = \Lambda(n) \log n + \sum_{hk=n} \Lambda(h)\Lambda(k)$ and $\sum_{n \leqslant x} a_n = 2x \log x + \mathrm{O}(x)$.

After some manipulations we get

$$|R(x)| \log^2 x \leqslant \sum_{n \leqslant x} a_n \left| R\left(\frac{x}{n}\right) \right| + \mathrm{O}(x \log x) \tag{3}$$

where $a_n = \Lambda(n) \log n + \sum_{hk=n} \Lambda(h)\Lambda(k)$ and $\sum_{n \leqslant x} a_n = 2x \log x + \mathrm{O}(x)$. Now we replace the sum with an integral.

$$|R(x)| \log^2 x \leqslant 2 \int_1^x \left| R\left(\frac{x}{t}\right) \right| \log t \mathrm{d}t + \mathrm{O}(x \log x). \tag{4}$$

Hence, we can rewrite (3) as

$$\log^2(z)\,|R(z)| \leqslant 2\int_1^z \left|R\left(\frac{z}{t}\right)\right| \log t\,\mathrm{d}t + \mathrm{O}(z\log z). \tag{5}$$

Hence, we can rewrite (3) as

$$\log^2(z)\,|R(z)| \leqslant 2\int_1^z \left|R\left(\frac{z}{t}\right)\right| \log t\,\mathrm{d}t + \mathrm{O}(z\log z). \qquad (5)$$

Showing $R(x) = \mathrm{o}(x)$ directly using (5) is hard because the behaviour of the von Mangoldt function $\Lambda(n)$ depends on the location of primes which is exactly what we're trying to find. Hence it is natural to define a smoother function $\sigma(x) = e^{-x}R(e^x) = e^{-x}\psi(e^x) - 1$.

Substitute $z = e^x,\ t = ze^{-u} \implies \mathrm{d}t = -ze^{-u}\mathrm{d}u.$

Substitute $z = e^x,\ t = ze^{-u} \implies \mathrm{d}t = -ze^{-u}\mathrm{d}u$. Then (5) becomes the simpler

$$x^2|\sigma(x)| \leqslant 2\int_0^x \int_0^y |\sigma(u)|\mathrm{d}u\mathrm{d}y + \mathrm{O}(x). \tag{6}$$

Substitute $z = e^x,\ t = ze^{-u} \implies \mathrm{d}t = -ze^{-u}\mathrm{d}u$. Then (5) becomes the simpler

$$x^2|\sigma(x)| \leqslant 2\int_0^x \int_0^y |\sigma(u)|\mathrm{d}u\mathrm{d}y + \mathrm{O}(x). \tag{6}$$

But $\psi(x) = \mathrm{O}(x)$, so $\sigma(x)$ is bounded for large x.

Substitute $z = e^x$, $t = ze^{-u} \implies \mathrm{d}t = -ze^{-u}\mathrm{d}u$. Then (5) becomes the simpler

$$x^2|\sigma(x)| \leqslant 2\int_0^x \int_0^y |\sigma(u)|\mathrm{d}u\mathrm{d}y + \mathrm{O}(x). \tag{6}$$

But $\psi(x) = \mathrm{O}(x)$, so $\sigma(x)$ is bounded for large x. So the upper limits

$$C = \limsup_{x\to\infty} |\sigma(x)| \quad \text{and} \quad K = \limsup_{x\to\infty} \frac{1}{x}\int_0^x |\sigma(u)|\mathrm{d}u \tag{7}$$

exist.

Then

$$\sigma(x) \leqslant C + \mathrm{o}(1) \quad \text{and} \quad \int_0^x |\sigma(u)| \mathrm{d}u \leqslant Kx + \mathrm{o}(x) \qquad (8)$$

Then

$$\sigma(x) \leqslant C + \mathrm{o}(1) \quad \text{and} \quad \int_0^x |\sigma(u)|\mathrm{d}u \leqslant Kx + \mathrm{o}(x) \qquad (8)$$

so using (6) we get

$$\sigma(x) \leqslant K + \mathrm{o}(1). \qquad (9)$$

Then

$$\sigma(x) \leqslant C + \mathrm{o}(1) \quad \text{and} \quad \int_0^x |\sigma(u)| \mathrm{d}u \leqslant Kx + \mathrm{o}(x) \tag{8}$$

so using (6) we get

$$\sigma(x) \leqslant K + \mathrm{o}(1). \tag{9}$$

Hence $C \leqslant K$.

Then

$$\sigma(x) \leqslant C + \mathrm{o}(1) \quad \text{and} \quad \int_0^x |\sigma(u)| \mathrm{d}u \leqslant Kx + \mathrm{o}(x) \tag{8}$$

so using (6) we get

$$\sigma(x) \leqslant K + \mathrm{o}(1). \tag{9}$$

Hence $C \leqslant K$. Now $R(x) = \mathrm{o}(x) \iff \sigma(x) = \mathrm{o}(1)$ so our aim is to show $C = 0$.

Then

$$\sigma(x) \leqslant C + \mathrm{o}(1) \quad \text{and} \quad \int_0^x |\sigma(u)| \mathrm{d}u \leqslant Kx + \mathrm{o}(x) \tag{8}$$

so using (6) we get

$$\sigma(x) \leqslant K + \mathrm{o}(1). \tag{9}$$

Hence $C \leqslant K$. Now $R(x) = \mathrm{o}(x) \iff \sigma(x) = \mathrm{o}(1)$ so our aim is to show $C = 0$. So we assume for contradiction that $C > 0$, then show that $K < C$ which is absurd.

Lemma 1

There is a fixed $A_1 > 0$ such that for all $x_1, x_2 > 0$ we have $\left|\int_{x_1}^{x_2} \sigma(u)\mathrm{d}u\right| < A_1$.

Lemma 2

If $\sigma(u_0) = 0$ for some $u_0 > 0$ then $\int_0^C |\sigma(u_0 + t)|\,\mathrm{d}t \leqslant \frac{C^2}{2} + \mathrm{O}(u_0^{-1})$.

Lemma 1

There is a fixed $A_1 > 0$ such that for all $x_1, x_2 > 0$ we have $\left|\int_{x_1}^{x_2} \sigma(u)\mathrm{d}u\right| < A_1$.

Lemma 2

If $\sigma(u_0) = 0$ for some $u_0 > 0$ then $\int_0^C |\sigma(u_0 + t)| \,\mathrm{d}t \leqslant \frac{C^2}{2} + \mathrm{O}(u_0^{-1})$.

(cf. [HW08] for proofs)

Now let $\delta = \dfrac{3C^2 + 4A_1}{2C} > C > 0$ and let $y > 0$ be arbitrary.

Now let $\delta = \dfrac{3C^2 + 4A_1}{2C} > C > 0$ and let $y > 0$ be arbitrary. We study the behaviour of $\sigma(u)$ on the interval $[y, y + \delta - C]$.

Now let $\delta = \dfrac{3C^2 + 4A_1}{2C} > C > 0$ and let $y > 0$ be arbitrary. We study the behaviour of $\sigma(u)$ on the interval $[y, y + \delta - C]$. By definition, $\sigma(u) = e^{-u}\psi(e^u) - 1$ is monotone increasing only at the jump discontinuities $u = \log p^k$ where it increases by $\log p$

Now let $\delta = \dfrac{3C^2 + 4A_1}{2C} > C > 0$ and let $y > 0$ be arbitrary. We study the behaviour of $\sigma(u)$ on the interval $[y, y + \delta - C]$. By definition, $\sigma(u) = e^{-u}\psi(e^u) - 1$ is monotone increasing only at the jump discontinuities $u = \log p^k$ where it increases by $\log p$ and between any two jump discontinuities $\sigma(u)$ decreases monotonically as $\psi(e^u)$ remains constant whilst e^{-u} decreases.

Now let $\delta = \dfrac{3C^2 + 4A_1}{2C} > C > 0$ and let $y > 0$ be arbitrary. We study the behaviour of $\sigma(u)$ on the interval $[y, y + \delta - C]$. By definition, $\sigma(u) = e^{-u}\psi(e^u) - 1$ is monotone increasing only at the jump discontinuities $u = \log p^k$ where it increases by $\log p$ and between any two jump discontinuities $\sigma(u)$ decreases monotonically as $\psi(e^u)$ remains constant whilst e^{-u} decreases. This means that either $\sigma(u)$ vanishes at some point $u = u_0$ or $\sigma(u)$ changes sign at most once.

<u>CASE I:</u> As $\sigma(u_0) = 0$ for some $u_0 \in [y, y + \delta - C]$, we use (8) and Lemma (27) to obtain

CASE I: As $\sigma(u_0) = 0$ for some $u_0 \in [y, y + \delta - C]$, we use (8) and Lemma (27) to obtain

$$\begin{aligned}\int_y^{y+\delta} |\sigma(u)| \mathrm{d}u &= \int_y^{u_0} + \int_{u_0}^{u_0+C} + \int_{u_0+C}^{y+\delta} |\sigma(u)| \mathrm{d}u \\ &\leqslant C(u_0 - y) + \frac{C^2}{2} + C\big(y + \delta - u_0 - C\big) + \mathrm{o}(1) \\ &= C\left(\delta - \frac{C}{2}\right) = \mathrm{o}(1) = C'\delta + \mathrm{o}(1)\end{aligned}$$

CASE I: As $\sigma(u_0) = 0$ for some $u_0 \in [y, y + \delta - C]$, we use (8) and Lemma (27) to obtain

$$
\begin{aligned}
\int_y^{y+\delta} |\sigma(u)| \mathrm{d}u &= \int_y^{u_0} + \int_{u_0}^{u_0+C} + \int_{u_0+C}^{y+\delta} |\sigma(u)| \mathrm{d}u \\
&\leqslant C(u_0 - y) + \frac{C^2}{2} + C\big(y + \delta - u_0 - C\big) + \mathrm{o}(1) \\
&= C\left(\delta - \frac{C}{2}\right) = \mathrm{o}(1) = C'\delta + \mathrm{o}(1)
\end{aligned}
$$

for all y sufficiently large, where we took $C' = C\left(1 - \frac{C}{2\delta}\right) < C$.

<u>CASE II:</u> If $\sigma(u)$ changes sign exactly once at some point $u = u_1 \in [y, y + \delta - C]$, then by Lemma (27)

<u>CASE II:</u> If $\sigma(u)$ changes sign exactly once at some point $u = u_1 \in [y, y+\delta-C]$, then by Lemma (27)

$$\int_y^{y+\delta-C} |\sigma(u)| \mathrm{d}u = \left| \int_y^{u_1} \sigma(u) \mathrm{d}u \right| + \left| \int_{u_1}^{y+\delta-C} \sigma(u) \mathrm{d}u \right| < 2A_1.$$

<u>Case II:</u> If $\sigma(u)$ changes sign exactly once at some point $u = u_1 \in [y, y + \delta - C]$, then by Lemma (27)

$$\int_y^{y+\delta-C} |\sigma(u)| \mathrm{d}u = \left| \int_y^{u_1} \sigma(u) \mathrm{d}u \right| + \left| \int_{u_1}^{y+\delta-C} \sigma(u) \mathrm{d}u \right| < 2A_1.$$

If $\sigma(u)$ does not change sign at all in the interval, then by Lemma (27) again

$$\int_y^{y+\delta-C} |\sigma(u)| \mathrm{d}u = \left| \int_y^{y+\delta-C} \sigma(u) \mathrm{d}u \right| < A_1 < 2A_1.$$

<u>Case II:</u> If $\sigma(u)$ changes sign exactly once at some point $u = u_1 \in [y, y+\delta - C]$, then by Lemma (27)

$$\int_y^{y+\delta-C} |\sigma(u)|\mathrm{d}u = \left|\int_y^{u_1} \sigma(u)\mathrm{d}u\right| + \left|\int_{u_1}^{y+\delta-C} \sigma(u)\mathrm{d}u\right| < 2A_1.$$

If $\sigma(u)$ does not change sign at all in the interval, then by Lemma (27) again

$$\int_y^{y+\delta-C} |\sigma(u)|\mathrm{d}u = \left|\int_y^{y+\delta-C} \sigma(u)\mathrm{d}u\right| < A_1 < 2A_1.$$

$$\begin{aligned}\text{Hence, } \int_y^{y+\delta} |\sigma(u)|\mathrm{d}u &= \int_y^{y+\delta-C} + \int_{y+\delta-C}^{y+\delta} |\sigma(u)|\mathrm{d}u \\ &< 2A_1 + \int_{y+\delta-C}^{y+\delta} |C + \mathrm{o}(1)|\mathrm{d}u \\ &= 2A_1 + C^2 + \mathrm{o}(1) = C''\delta + \mathrm{o}(1)\end{aligned}$$

where we took

$$C'' = \frac{2A_1 + C^2}{\delta} = C\left(\frac{4A_1 + 2C^2}{4A_1 + 3C^2}\right) = C\left(1 - \frac{C}{2\delta}\right) = C'.$$

where we took
$C'' = \frac{2A_1 + C^2}{\delta} = C\left(\frac{4A_1 + 2C^2}{4A_1 + 3C^2}\right) = C\left(1 - \frac{C}{2\delta}\right) = C'$. In both
cases we always have

$$\int_y^{y+\delta} |\sigma(u)| \mathrm{d}u \leqslant C'\delta + \mathrm{o}(1).$$

where we took
$C'' = \dfrac{2A_1 + C^2}{\delta} = C\left(\dfrac{4A_1 + 2C^2}{4A_1 + 3C^2}\right) = C\left(1 - \dfrac{C}{2\delta}\right) = C'$. In both cases we always have

$$\int_y^{y+\delta} |\sigma(u)| \mathrm{d}u \leqslant C'\delta + \mathrm{o}(1).$$

If $M = [x/\delta]$, then

$$\begin{aligned} \int_0^x |\sigma(u)| \mathrm{d}u &= \sum_{m=0}^{M-1} \int_{m\delta}^{(m+1)\delta} |\sigma(u)| \mathrm{d}u + \int_{M\delta}^x |\sigma(u)| \mathrm{d}u \\ &\leqslant C'M\delta + \mathrm{o}(M) + \mathrm{O}(1) = C'x + \mathrm{o}(x). \end{aligned}$$

where we took
$C'' = \dfrac{2A_1 + C^2}{\delta} = C\left(\dfrac{4A_1 + 2C^2}{4A_1 + 3C^2}\right) = C\left(1 - \dfrac{C}{2\delta}\right) = C'$. In both cases we always have

$$\int_y^{y+\delta} |\sigma(u)| \mathrm{d}u \leqslant C'\delta + \mathrm{o}(1).$$

If $M = [x/\delta]$, then

$$\begin{aligned}\int_0^x |\sigma(u)| \mathrm{d}u &= \sum_{m=0}^{M-1} \int_{m\delta}^{(m+1)\delta} |\sigma(u)| \mathrm{d}u + \int_{M\delta}^x |\sigma(u)| \mathrm{d}u \\ &\leqslant C'M\delta + \mathrm{o}(M) + \mathrm{O}(1) = C'x + \mathrm{o}(x).\end{aligned}$$

Hence,

$$K = \limsup_{x\to\infty} \frac{1}{x} \int_0^x |\sigma(u)| \mathrm{d}u \leqslant C' < C.$$

An absurdity. Hence, we must have $C = 0$. ■

Table of Contents

Landau's Prime Ideal Theorem

Let K be a number field and $\pi_K(X)$ denote the number of prime ideals in K of norm at most X,

$$\pi_K(X) = \#\{\mathfrak{p} : \mathrm{N}_K(\mathfrak{p}) \leqslant X\}.$$

Landau's Prime Ideal Theorem

Let K be a number field and $\pi_K(X)$ denote the number of prime ideals in K of norm at most X,

$$\pi_K(X) = \#\{\mathfrak{p} : \mathrm{N}_K(\mathfrak{p}) \leqslant X\}.$$

In 1903, Landau [Lan03] proved the following asymptotic estimate analogous to the PNT

Prime Ideal Theorem

$$\pi_K(X) \sim \frac{X}{\log X}.$$

Landau's Prime Ideal Theorem

Let K be a number field and $\pi_K(X)$ denote the number of prime ideals in K of norm at most X,

$$\pi_K(X) = \#\{\mathfrak{p} : \mathrm{N}_K(\mathfrak{p}) \leqslant X\}.$$

In 1903, Landau [Lan03] proved the following asymptotic estimate analogous to the PNT

Prime Ideal Theorem

$$\pi_K(X) \sim \frac{X}{\log X}.$$

An elementary proof similar in spirit to Selberg's was given by Shapiro [Sha49] in 1949.

[HW08] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers*. Oxford Mathematics. Oxford University Press, 2008.

[Lan03] Edmund Landau. "Neuer Beweis des Primzahlsatzes und Beweis des Primidealsatzes". In: *Mathematische Annalen* 56.4 (1903).

[Sel49] Atle Selberg. "An Elementary Proof of the Prime-Number Theorem". In: *Annals of Mathematics* 50.2 (1949).

[Sha49] Harold Shapiro. "An elementary proof of the prime ideal theorem". In: *Communications on Pure and Applied Mathematics* 2.4 (1949).

[TI51] Tikao Tatuzawa and Kanesiroo Iseki. "On Selberg's elementary proof of the prime-number theorem". In: *Proceedings of the Japan Academy, Series A, Mathematical Sciences* 27.7 (1951).