ON THE PRIME NUMBER THEOREM

ZHENGJUN LIANG (JASPER)

Abstract. The following report is to give an exposition of the prime number theorem, which is a famous theorem in analytic number theory that gives an asymptotic of the number of primes. The proof uses tools of complex analysis.

1. Introduction

If one traces the pattern of the primes, one may discover that the distribution of primes seems to get increasingly sparser as the primes become larger. The prime number theorem, initially proven by Jacques Hadamard and Charles Jean de la Vallée Poussin in 1896, gives a more precise characterization of the intuition by giving an asymptotic estimate of the number of primes. The theorem is precisely stated as the following:

Theorem 1.1 (Prime Number Theorem)**.** *For any $x \in \mathbb{R}$, let $\pi(x)$ denote the number of primes less than or equal to x. Then we have the asymptotics*

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

which we can write as

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

using asymptotics notations.

Heuristically, if we use $f(x) := x/\log x$ as an approximation of $\pi(x)$ for very large x, we can indeed see that the distribution of the primes gets sparser. This is because for x sufficiently large we have

$$f''(x) = \frac{2 - 2\log x}{x(\log x)^3} < 0$$

showing that the growth of f gradually slows down. The rest of the article is dedicated to proving Theorem 1.1 using an approach that I am familiar with, which uses the method of complex analysis. A prerequisite to understand the proof is a fair amount of knowledge in complex analysis (which can be found on [7]), and basic familiarity with elementary number theory will be helpful.

It is worth mentioning that our proof of the prime number theorem is far from unique. For example, the reader is advised to see [5] for a proof using techniques in dynamical systems, and see [6] for a proof using a different complex analysis argument. The introduction of [5] also gives a rather detailed review on the history and proofs of the prime number theorem.

1.1. **Notations.** Before we start the proof, we fix some notations. As in Theorem 1.1, we denote

$$\pi(x) := \#\{p \text{ prime} : p \leq x\}$$

We also define

$$\theta(x) := \sum_{p \leq x} \log(p)$$

As in the convention of number theory, we have the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$, and the von Mangoldt function $\Lambda : \mathbb{N} \to \mathbb{R}$ defined by

$$\Lambda(k) = \begin{cases} \log(p) & \text{if } k = p^n \\ 0 & \text{otherwise} \end{cases}$$

We adopt the asymptotic notations in harmonic analysis, where we use $X \lesssim Y$, $Y \gtrsim X$, or $X = O(Y)$ to denote $|X| \leq CY$ for some absolute constant C. As above, we also use $f(x) \sim g(x)$ to denote

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1$$

Eventually, the notation p always means a prime in this article, and the notations $\sum_p$, $\sum_{p \leq x}$ are summations over primes.

1.2. **Acknowledgement.** The report serves as the final project to MATH 205A at University of California, Los Angeles, a graduate class on analytic number theory taught by professor William Duke during Fall 2020. I appreciate professor Duke for giving elegant lectures and answering my inquiries after classes. Moreover, many inspirations of this report come from a class MATH 246B taught by professor Rowan Killip during Winter 2020.

2. Outlining the Proof of the Prime Number Theorem

At the beginning of the proof, we provide a crude bound of $\theta(x)/x$, following an argument of Chebyshev.

Proposition 2.1 (Crude Bound). *We have the estimate*

$$\theta(x) \leq 4\log(2)x$$

which in particular implies that $\theta(x)/x$ is bounded.

Proof. We first need to show $\binom{2n}{n} \leq 2^{2n}$. For this we notice that

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} = \frac{(2n)...(n+1)n...1}{n...1 \cdot n...1} \leq \frac{(2n)(2n)(2n-2)(2n-2)...(2)(2)}{n...1 \cdot n...1} = 2^{2n}$$

Moreover, we claim that every prime number in $(n, 2n]$ divides $\binom{2n}{n}$ exactly once. To prove the claim, we first notice that every $p \in (n, 2n]$ divides $\binom{2n}{n}$ at most once since there is only one copy of p among $n+1, ..., 2n$. Moreover, $p \mid \binom{2n}{n} n!$ and $p \nmid n!$, so $p \mid \binom{2n}{n}$. This claim then gives

$$\prod_{n < p \leq 2n} p \leq \binom{2n}{n} \leq 2^{2n}$$

Thus, suppose 2^{k+1} is the first dyadic number exceeding x, we have

$$\prod_{p\leq x} p \leq \prod_{p\leq 2^{k+1}} p \leq \binom{2}{1}\binom{4}{2}\cdots\binom{2^{k+1}}{2^k} \leq 2^{2+4+\ldots+2^{k+1}} = 2^{2^{k+1}-1} \leq 2^{2x}$$

and thus

$$\theta(x) = \log\left(\prod_{p\leq x} p\right) \leq \log(2^{2x}) = 4\log(2)x$$

as desired. □

The constant $4\log(2)$ is not precisely what we want, but this does not really matter. What we want is "some bound", telling us that

$$\theta(x) = O(x) \tag{2.2}$$

which shall be useful in our argument later.

The next step of the proof is to give an equivalent characterization of Theorem 1.1, transforming it to a problem involving $\theta(x)$.

Lemma 2.3. *As $x \to \infty$, we have*

$$\frac{\theta(x)}{x} \to 1 \tag{2.4}$$

if and only if

$$\frac{\pi(x)\log(x)}{x} \to 1$$

The proof of this lemma will be given in the next section. With this alternative characterization available, our task of proving Theorem 1.1 is reduced to showing (2.4), for which we provide the following functional equation:

Lemma 2.5. *For $s > 1$, we have*

$$\int_1^\infty \theta(x)\frac{dx}{x^{s+1}} = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{F(s)}{s}$$

where F is holomorphic in the half plane $\mathrm{Re}(s) > \frac{1}{2}$.

Again, the proof of this lemma is postponed to the next section. The last major ingredient to finishing the proof of Theorem 1.1 is a Tauberian argument by Newman, stated as the following:

Theorem 2.6 (Newman)**.** *Suppose $\varphi : [1,\infty) \to \mathbb{R}$ is $O(x)$ and*

$$\Phi(s) := \int_1^\infty \varphi(x)\frac{dx}{x^{s+1}}$$

is absolutely convergent and thus holomorphic on $\mathrm{Re}(s) > 1$*. If $\Phi(s)$ admits a holomorphic extension to an open neighborhood of of* $\mathrm{Re}(s) \geq 1$*, then*

$$\lim_{x\to\infty}\int_1^x \varphi(x)\frac{dx}{x^2} = \Phi(1)$$

A later section will be dedicated to the proof of this theorem. For now, we notice that $\varphi(x) = \theta(x) - x$ satisfies the hypothesis of Theorem 2.6 with the help of Proposition 2.1 and Lemma 2.5. In particular, Proposition 2.1 tells us that $\theta(x) - x = O(x)$, and by Lemma 2.5 we have

$$\Phi(s) := \int_1^\infty \frac{\theta(x) - x}{x^{s+1}}\, dx = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{F(s)}{s} - \frac{1}{s-1}$$

holomorphic on $\mathrm{Re}(s) > 1$. Moreover, by our knowledge of $\zeta(s)$, we can deduce that $\zeta'(s)/\zeta(s)$ has a simple pole at $s = 1$ with residue -1, so

$$-\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1}$$

can be holomorphically extended to a neighborhood of $\mathrm{Re}(s) \geq 1$, and thus so is $\Phi(s)$. Therefore, applying Theorem 2.6 we know that

$$\int_1^\infty \frac{\theta(x) - x}{x^2}\, dx = c \tag{2.7}$$

for some $c \in \mathbb{R}$.

We now claim that (2.4) must hold, and thus finish the proof. To prove the claim, we split to two cases. First we assume towards contradiction that

$$\limsup_{x\to\infty} \frac{\theta(x)}{x} > 1$$

Then there is some $\varepsilon > 0$ and $x_n \to \infty$ such that

$$\frac{\theta(x_n)}{x_n} \geq 1 + \varepsilon \quad \text{for all } n$$

Since $\theta(x)$ is non-decreasing in x, we can see that $\theta(x) \geq (1+\varepsilon)x_n$ for any $x \geq x_n$. Thus direct computation gives

$$\int_{x_n}^{(1+\varepsilon)x_n} \frac{\theta(x) - x}{x^2}\, dx \geq \int_{x_n}^{(1+\varepsilon)x_n} \frac{(1+\varepsilon)x_n - x}{x^2}\, dx \geq \varepsilon - \log(1+\varepsilon) > 0$$

uniformly in n. This contradicts the convergence displayed in (2.7). In the second case we assume towards contradiction that

$$\liminf_{x\to\infty} \frac{\theta(x)}{x} < 1$$

Then there is some $\varepsilon > 0$ and a sequence $x_n \to \infty$ such that $\theta(x_n) \leq (1-\varepsilon)x_n$. Using the monotonicity of θ once again, we can deduce that

$$\int_{(1-\varepsilon)x_n}^{x_n} \frac{\theta(x) - x}{x^2}\, dx \leq \int_{(1-\varepsilon)x_n}^{x_n} \frac{(1-\varepsilon)x_n - x}{x^2}\, dx \leq \varepsilon - \log\left(\frac{1}{1-\varepsilon}\right) < 0$$

which is again inconsistent with the convergence in (2.7). Thus our claim is proven and the proof is finished.

3. Proof of the Lemmas

In this section, we give detailed proofs of the lemmas mentioned in the previous section. At the beginning, we introduce a summation technique that is extensively used in analytic number theory.

Theorem 3.1 (Abel Summation)**.** *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real or complex numbers. Define the partial sum function A by*

$$A(t) := \sum_{0\le n\le t} a_n$$

for any real number t. Fix real numbers $x < y$, and let ϕ be a continuously differentiable function on $[x, y]$. Then

$$\sum_{x<n\le y} a_n\phi(n) = A(y)\phi(y) - A(x)\phi(x) - \int_x^y A(u)\phi'(u)\ du$$

The proof follows from integration by parts. For details, the reader can consult any standard text in analytic number theory such as [1].

Proof of Lemma 2.3. Suppose $\theta(x) \sim x$,

$$\begin{aligned}\pi(x) &= \sum_{p\le x} 1 = \sum_{2\le p\le x} \frac{\log p}{\log p} \\ &\le \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t(\log t)^2} dt \\ &\sim \frac{x}{\log x} + \int_2^x \frac{dt}{\log^2(t)}\end{aligned}$$

where the second line follows from Abel summations for integrals. Notice that

$$\frac{d}{dx}\left(\int_2^x \frac{dt}{\log t} - \frac{x}{\log x}\right) = \frac{1}{\log x} - \frac{\log x - 1}{\log^2 x} = \frac{1}{\log^2 x}$$

So

$$\frac{x}{\log x} + \int_2^x \frac{dt}{\log^2(t)} = \frac{x}{\log x} + \int_2^x \frac{dt}{\log t} - \frac{x}{\log x} + \frac{2}{\log 2} = \int_2^x \frac{dt}{\log t} + \frac{2}{\log 2}$$

Note that by L'Hospital's rule

$$\lim_{x\to\infty} \frac{\int_2^x \frac{dt}{\log t}}{\frac{x}{\log x}} = \lim_{x\to\infty} \frac{1/\log x}{(\log x - 1)/\log^2 x} = \lim_{x\to\infty} \frac{\log x}{\log x - 1} = 1 \tag{3.2}$$

so actually

$$\frac{x}{\log x} + \int_2^x \frac{dt}{\log^2(t)} \sim \frac{x}{\log x} + \frac{2}{\log 2} \sim \frac{x}{\log x}$$

which means

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

as desired. Conversely, suppose $\pi(x) \sim \frac{x}{\log x}$. Then

$$\begin{aligned}\theta(x) &= \sum_{p\le x} \log p = \sum_{2<n\le x} [\pi(n) - \pi(n-1)] \log n \\ &= \pi(x)\log x - \int_2^x \frac{\pi(t)}{t} dt \\ &\sim x - \int_2^x \frac{dt}{\log t} \sim x - \frac{x}{\log x}\end{aligned}$$

where the second line is again by Abel summations, and the last step follows from (3.2). Now

$$\lim_{x\to\infty} \frac{\theta(x)}{x} = 1 - \lim_{x\to\infty} \frac{1}{\log x} = 1$$

and this finishes the proof. □

Proof of Lemma 2.5. Notice that

$$\int_1^\infty \frac{\theta(x)}{x^{s+1}} = \sum_p \log p \int_p^\infty x^{-s-1} dx = \frac{1}{s} \sum_p \log p \cdot p^{-s}$$

while we have the functional equation (see for example [1] or [3])

$$-\frac{\zeta'(s)}{s\zeta(s)} = \frac{1}{s}\sum_{k=1}^\infty \Lambda(k)$$

so it suffices to prove that

$$F(s) = \sum_{k=1}^\infty \Lambda(k) - \sum_p \log p \cdot p^{-s} \tag{3.3}$$

is a holomorphic function on $\mathrm{Re}s > \frac{1}{2}$. For this, we notice that

$$(3.3) = \sum_p \sum_{n=2}^\infty \log(p)(p^{-n})^{-s} = \sum_p \log p \frac{p^{-2s}}{1-p^{-s}} = \sum_p \frac{\log p}{p^s(p^s-1)}$$

To show it is holomorphic on $\mathrm{Re}(s) > \frac{1}{2}$, we need to show that it converges locally uniformly on $\mathrm{Re}(s) > \frac{1}{2}$. Let $\varepsilon > 0$. If $\mathrm{Re}(s) > \frac{1}{2} + \varepsilon$, then $\log n \lesssim n^\varepsilon$ since $\log n / n^\varepsilon \to 0$ as $n \to \infty$, and thus

$$\sum_p \frac{\log p}{p^s(p^s-1)} \le \sum_n \frac{\log n}{n^{\mathrm{Re}(s)}(n^{\mathrm{Re}(s)}-1)} \lesssim \sum_n \frac{1}{n^{2\mathrm{Re}(s)-1}} \le \sum_n \frac{1}{n^{1+\varepsilon}} < \infty$$

so it converges locally uniformly on $\mathrm{Re}(s) > \frac{1}{2} + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we have proven the claim that $F(s)$ is holomorphic on $\mathrm{Re}(s) > \frac{1}{2}$. Thus we have proven the lemma. □

4. Proof of Theorem 2.6

In this section, we fill the one last gap left behind in the previous outline by proving the Newman Tauberian theorem. We follow the approach of Killip in his lecture notes [4].

We define

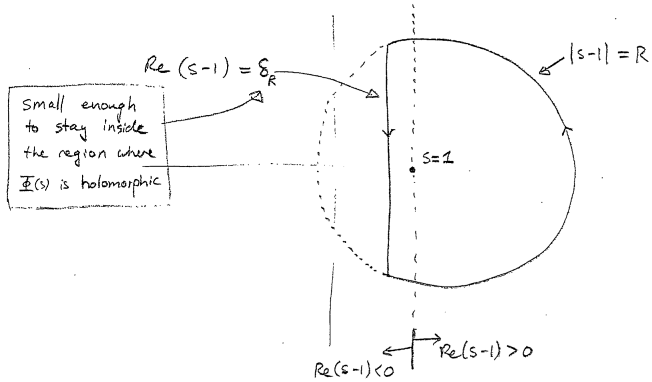
$$\Phi_x(s) := \int_1^x \varphi(x) \frac{dx}{x^{s+1}}$$

which is an entire function of s. We wish to show that $\Phi_x(1) \to \Phi(1)$ as $x \to \infty$.

For any $R > 0$, by our knowledge of contour integration we have

$$\Phi(1) - \Phi_x(1) = \frac{1}{2\pi i} \int_{\gamma_R} [\Phi(s) - \Phi_x(s)] x^{s-1} \left(1 + \frac{(s-1)^2}{R^2}\right) \frac{ds}{s-1}$$

where γ_R is the contour (the figure is taken from [4])



Notice that $s - 1 = Re^{i\theta}$, so

$$\left| \frac{1}{s-1} \left(1 + \frac{(s-1)^2}{R^2}\right) \right| = \frac{1}{R} |e^{-i\theta} + e^{i\theta}| = \frac{2|\cos\theta|}{R} \lesssim \frac{|\mathrm{Re}(s-1)|}{R^2}$$

Now, let $\gamma_R^+ := \gamma_R \cap \{\mathrm{Re}(s-1) > 0\}$. On this part of the contour, we have

$$|x^{s-1}(\Phi(s) - \Phi(s))| \leq x^{\mathrm{Re}(s-1)} \int_x^\infty \left| \frac{\varphi(x)}{x} \right| \frac{dx}{x^{\mathrm{Re}(s)}} \lesssim \frac{1}{\mathrm{Re}(s-1)}$$

where the last step follows from our assumption that $\varphi(x) = O(x)$. It follows that

$$\int_{\gamma_R^+} [\Phi(s) - \Phi_x(s)] x^{s-1} \left(1 + \frac{(s-1)^2}{R^2}\right) \frac{ds}{s-1} \lesssim \int_{\gamma_R^+} \frac{1}{|\mathrm{Re}(s-1)|} \frac{\mathrm{Re}(s-1)}{R^2} \frac{|ds|}{2\pi}$$
$$\lesssim \frac{1}{R}$$

We next consider $\gamma_R^- := \gamma_R \cap \{\mathrm{Re}(s-1) < 0\}$, for which we treat $\Phi(s)$ and $\Phi_x(s)$ separately. In the former case, we notice that

$$\left| \Phi(s) \left(1 + \frac{(s-1)^2}{R^2}\right) \frac{1}{s-1} \right| \lesssim C(R)$$

where the implicit constant is independent of s. Note that C depends on R through δ_R and also the size of Φ on γ_R^-. Nevertheless, since $|x^{s-1}| = x^{\mathrm{Re}(s-1)} \to 0$ as $x \to \infty$ pointwise on γ_R^-, we deduce that

$$\limsup_{x\to\infty} \left| \int_{\gamma_R^-} \Phi(s) x^{s-1} \left(1 + \frac{(s-1)^2}{R^2}\right) \frac{ds}{2\pi i (s-1)} \right| = 0$$

It remains to study the $\Phi_x(s)$, for which we deform the contour to

$$\tilde{\gamma}_R^- = \{|s-1| = R,\ \mathrm{Re}(s-1) < 0\}$$

and use

$$\left| x^{s-1} \int_1^x \frac{\varphi(x)}{x} \frac{dx}{x^s} \right| \lesssim x^{\mathrm{Re}(s-1)} \int_1^x \frac{dx}{x^{\mathrm{Re}(s)}} \lesssim \frac{1}{\mathrm{Re}(s-1)}$$

to deduce that

$$\left| \int_{\tilde{\gamma}_R^-} \Phi_x(s) x^{s-1} \left(1 + \frac{(s-1)^2}{R^2}\right) \frac{ds}{s-1} \right| \lesssim \int_{\tilde{\gamma}_R^-} \frac{1}{|\mathrm{Re}(s-1)|} \cdot \frac{|\mathrm{Re}(s-1)|}{R^2}\ |ds| \lesssim \frac{1}{R}$$

much as in the case of γ_R^+. Putting all the pieces together we obtain

$$\limsup_{x\to\infty} |\Phi(1) - \Phi_x(1)| \lesssim R^{-1}$$

where the dependence is uniform in R. Since $R > 0$ is arbitrary, we get the theorem as desired.

References

[1] Tom Apostol, *Introduction to Analytic Number Theory.* 1976 Springer-Verlag New York, Inc.

[2] Harold Davenport, *Multiplicative Number Theory.* Third Edition. 2000 Springer-Verlag New York, Inc.

[3] Kenneth Ireland, Michael Rosen. *A Classical Introduction to Modern Number Theory.* Second Edition. 1972, 1982, 1990 Springer-Verlag New York, Inc.

[4] Rowan Killip. *MATH 246B Lecture Notes.* Available at `https://www.math.ucla.edu/~killip/246b/Last_two_lectures.pdf`

[5] Redmond McNamara. *A Dynamical Proof of the Prime Number Theorem.* Available at `https://arxiv.org/pdf/2002.04007.pdf`

[6] Walter Rudin. *Functional Analysis.* 2006 McGraw Hill (India) Education.

[7] Elias Stein and Rami Shakarchi. *Complex Analysis.* Princeton Lectures in Analysis II. 2003 Princeton University Press.