Notes on the prime number theorem

Kenji Kozai

May 2, 2014

1 Statement

We begin with a definition.

Definition 1.1. We say that $f(x)$ and $g(x)$ are asymptotic as $x \to \infty$, written $f \sim g$, if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$.

The prime number theorem tells us about the asymptotic behavior of the number of primes that are less than a given number. Let $\pi(x)$ be the number of primes numbers p such that $p \leq x$. For example, $\pi(2) = 1$, $\pi(3) = 2$, $\pi(4) = 2$, etc. The statement that we will try to prove is as follows.

Theorem 1.2 (Prime Number Theorem, p. 382)**.** *The asymptotic behavior of $\pi(x)$ as $x \to \infty$ is given by*

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

This was originally observed as early as the 1700s by Gauss and others. The first proof was given in 1896 by Hadamard and de la Vallée Poussin. We will give an overview of simplified proof. Most of the material comes from various sections of Gamelin – mostly XIV.1, XIV.3 and XIV.5.

2 The Gamma Function

The gamma function $\Gamma(z)$ is a meromorphic function that extends the factorial $n!$ to arbitrary complex values. For $\operatorname{Re} z > 0$, we can find the gamma function by the integral:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt, \operatorname{Re} z > 0.$$

To see the relation with the factorial, we integrate by parts:

$$\begin{aligned}\Gamma(z+1) &= \int_0^\infty e^{-t}t^z dt \\ &= -t^z e^{-t}|_0^\infty + z\int_0^\infty e^{-t}t^{z-1}dt \\ &= z\Gamma(z).\end{aligned}$$

This holds whenever $\operatorname{Re} z > 0$. We also notice that $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$. By induction, then $\Gamma(2) = 1$, $\Gamma(3) = 2$, $\Gamma(4) = 6$, $\ldots$, $\Gamma(n+1) = n!$. This holds for all non-negative integers n.

Remember that the integral definition for $\Gamma(z)$ only works for $\operatorname{Re} z > 0$. However, using the relation $\Gamma(z+1) = z\Gamma(z)$, we can extend it for negative values of z by defining it to be the function that agrees with the integral definition for $\operatorname{Re} z > 0$, and also satisfies $\Gamma(z+1) = z\Gamma(z)$ for all z. In particular, if we apply this relation m times, we find that

$$\Gamma(z+m) = (z+m-1)\cdots(z+1)z\Gamma(z).$$

We can rewrite this as

$$\Gamma(z) = \frac{\Gamma(z+m)}{(z+m-1)\cdots(z+1)z}.$$

The function $\Gamma(z+m)$ is defined for $\operatorname{Re} z > -m$, so we define $\Gamma(z)$ for $\operatorname{Re} z > -m$ by using the above equation. By doing this for larger and larger value of m, we can progressively define it for more of the complex plane, and in the limit, we get a function defined on all of $\mathbb{C}$.

From this definition, we can see that $\Gamma(z)$ is meromorphic with simple poles at $z = 0, -1, -2, -3, \ldots$.

3 The Zeta Function

Before we give a proof of the theorem, we will need to study the zeta function, which is covered in XIV.3. The zeta function is defined to be

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \operatorname{Re} s > 1.$$

Here, $s = \sigma + it$ is a complex number. Notice that if $\sigma = \operatorname{Re} s > 1$, the series on the right in fact does converge absolutely, as

$$\left|\frac{1}{n^s}\right| = \frac{1}{n^\sigma},$$

with the series $\sum \frac{1}{n^\sigma}$ converging by the p-test when $\sigma > 1$. As $\sigma \to 1$, we obtain the harmonic series $\sum \frac{1}{n}$, which diverges.

The zeta function satisfies a product formula.

Theorem 3.1 ((3.1) on p. 371)**.** *If* $\operatorname{Re} s > 1$*, then*

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right),$$

where the product is taken over all prime numbers p*.*

Proof. The proof uses the geometric series formula

$$\frac{1}{1-p^{-s}} = 1 + p^{-s} + p^{-2s} + p^{-3s} + \dots.$$

Now, if $p_1, p_2, \dots, p_m$ are m different primes, if we multiply the geometric series for each of these primes together, we obtain

$$\frac{1}{(1-p_1^{-s})\cdots(1-p_m^{-s})} = \sum_{k_1,\dots,k_m=0}^{\infty} (p_1^{k_1}\cdots p_m^{k_m})^{-s}.$$

Every integer $n \geq 1$ can be written uniquely as a product of primes, so each number $\frac{1}{n^{-s}}$ such that n is a product of the primes $p_1, \dots, p_m$ appears exactly once. In the limit, when we take all primes, we get all terms $\frac{1}{n^{-s}}$. Hence,

$$\prod_p \frac{1}{1-p^{-s}} = \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s).$$

Taking the reciprocal then yields the formula in the theorem. $\square$

It turns out that $\zeta(s)$ can be extended to be a meromorphic function on $\mathbb{C}$, with only a simple pole at $s = 1$ with residue 1. More details can be found in XIV.3. This extension satisfies the equation

$$\zeta(s) = 2^s \pi^{s-1} \sin\frac{\pi s}{2} \Gamma(1-s)\zeta(1-s).$$

In particular, the right hand side is defined for $\operatorname{Re} s < 0$, so this tells us $\zeta(s)$ for all values of s except for the critical strip $0 \leq \operatorname{Re} s \leq 1$. An integral formula described in XIV.3 is used to define $\zeta(s)$ in this strip.

The product formula for $\zeta(s)$ shows that $\zeta(s)$ has no zeros in the right half-plane $\operatorname{Re} s > 1$. The equation above then implies that the only zeros of $\zeta(s)$ in the left half-plane $\operatorname{Re} s < 0$ are the zeros of $\sin \frac{\pi s}{2}$, which are $s = -2, -4, -6, \dots$. The famous Riemann hypothesis is the conjecture that all other zeros of $\zeta(s)$ lie on the line $\operatorname{Re} s = \frac{1}{2}$.

4 Chebyshev theta function

Instead of comparing the asymptotic behavior of $\pi(x)$ with $\frac{x}{\log x}$ directly, we will consider the Chebyshev theta function,

$$\Theta(x) = \sum_{p \leq x} \log p.$$

We will compare the theta function with the statement from the prime number theorem.

Lemma 4.1 (p. 384)**.** *$\pi(x) \sim \frac{x}{\log x}$ if and only if $\Theta(x) \sim x$.*

Proof. Notice that $\log p \leq \log x$, and there are $\pi(x)$ summands for $\Theta(x)$. Hence, $\Theta(x) \leq \pi(x) \log x$, as long as $x \geq 1$. Dividing by x gives us $\frac{\Theta(x)}{x} \leq \pi(x)\frac{\log x}{x}$.

Let $\epsilon > 0$. Then,

$$\begin{aligned}
\Theta(x) &= \sum_{p \leq x} \log p \\
&\geq \sum_{x^{1-\epsilon} < p \leq x} \log p \\
&\geq \sum_{x^{1-\epsilon} < p \leq x} \log x^{1-\epsilon} \\
&= (1-\epsilon)(\log x)(\pi(x) - \pi(x^{1-\epsilon})) \\
&\geq (1-\epsilon)(\log x)(\pi(x) - x^{1-\epsilon}).
\end{aligned}$$

We can rearrange this to

$$\begin{aligned}
\frac{\Theta(x)}{1-\epsilon} &\geq (\log x)(\pi(x) - x^{1-\epsilon}) \\
\frac{\Theta(x)}{1-\epsilon} + \frac{x \log x}{x^{\epsilon}} &\geq (\log x)\pi(x).
\end{aligned}$$

Then, dividing by x and combining with our first inequality, we can see that

$$\frac{\Theta(x)}{x} \leq \pi(x)\frac{\log x}{x} \leq \frac{1}{1-\epsilon}\frac{\Theta(x)}{x} + \frac{\log x}{x^{\epsilon}}.$$

Since $\lim_{x\to\infty} \frac{\log x}{x^{\epsilon}} = 0$, we see that $\frac{\pi(x)}{x/\log x} \to 1$ if and only if $\frac{\Theta(x)}{x} \to 1$. $\square$

So in order to prove the prime number theorem, it suffices to show $\Theta(x) \sim x$.

5 The Laplace transform

We will begin with a lemma.

Lemma 5.1. *If $f(x)$ is an increasing function of x such that*

$$\lim_{R\to\infty}\int_1^R \left(\frac{f(x)}{x}-1\right)\frac{dx}{x}$$

exists, then $f(x) \sim x$.

Before proving the lemma, we note that $\Theta(x)$ is an increasing function in x, so if we can prove that the limit above exists for $\Theta(x)$, then we would prove the prime number theorem.

Proof. We will prove this by the contrapositive. Since $f(x)$ is increasing, the only way that $f(x) \sim x$ fails is if there exists an $\epsilon > 0$ such that $f(x) > (1+\epsilon)x$ for sufficiently large values of x, or $f(x) < (1-\epsilon)x$ for sufficiently large values of x.

Let us assume the first case, that $f(x) > (1+\epsilon)x$ for large values of x. Suppose that this is true for $x \geq x_0$. Then

$$\begin{aligned}\int_{x_0}^{(1+\epsilon)x_0}\left(\frac{f(x)}{x}-1\right)\frac{dx}{x} &\geq \int_{x_0}^{(1+\epsilon)x_0}\left(\frac{f(x_0)}{x}-1\right)\frac{dx}{x}\\ &\geq \int_{x_0}^{(1+\epsilon)x_0}\left(\frac{(1+\epsilon)x_0}{x}-1\right)\frac{dx}{x}\\ &= \int_1^{1+\epsilon}\left(\frac{(1+\epsilon)}{t}-1\right)\frac{dt}{t}.\end{aligned}$$

The last line is by using the change of coordinates $x = x_0 t$. In particular, the integral in the last line is independent of x_0 and is some non-zero $c > 0$.

We can repeat this for infinitely many disjoint intervals, so

$$\lim_{R\to\infty}\int_1^R \left(\frac{f(x)}{x}-1\right)\frac{dx}{x}$$

diverges to ∞, so it cannot exist. Similarly, we can show that if $f(x) < (1-\epsilon)x$ for sufficiently large x, then the integral diverges to $-\infty$.

Therefore $f(x) \sim x$. □

In order to prove that the limit exists for $\Theta(x)$, we will need to introduce one more tool: the Laplace transform.

Definition 5.2. Let $h(s)$ be a continuous or piecewise function on the positive real axis $s \geq 0$. The Laplace transform of $h(s)$ is the function of z

$$(\mathcal{L}h)(z) = \int_0^\infty e^{-sz} h(s) ds,$$

provided that the integral converges.

We won't need to know too much about the Laplace transform, so we won't go into depth here. But if you want to know more, the book has a section on the Laplace transform in XIV.2. The Laplace transform is somewhat like the Fourier transform.

Notice that the Laplace transform, evaluated at $z = 0$, is

$$(\mathcal{L}h)(0) = \int_0^\infty e^{-s(0)} h(s) ds = \int_0^\infty h(s) ds.$$

Recall that what we are trying to show is that

$$\lim_{R \to \infty} \int_1^R \left(\frac{\Theta(x)}{x} - 1 \right) \frac{dx}{x}$$

exists. If we do a change of variages $x = e^s$, then $dx = e^s ds = x ds$, so the integral becomes

$$\lim_{T \to \infty} \int_0^T (\Theta(e^s) e^{-s} - 1) ds.$$

This limit is exactly the expression for $(\mathcal{L}h)(0)$ above, for $h(s) = \Theta(e^s)e^{-s} - 1$. Thus, out goal will be to show that $(\mathcal{L}h)(0)$ exists. If we can do that, we will have proved the prime number theorem.

We will do this by connecting the Laplace transform with another function whose poles we understand, $\Phi(s)$.

6 Dirichlet Series

Definition 6.1. A *Dirichlet series* is a series of the form

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = a_1 + \frac{a_2}{2^s} + \frac{a_3}{3^s} + \ldots.$$

The zeta function $\zeta(s)$ is a Dirichlet series with $a_1 = a_2 = \cdots = a_n = \cdots = 1$.

The proof will use a particular Dirichlet series defined by

$$\Phi(s) = \sum_p \frac{\log p}{p^s},$$

which converges and is analytic for $\operatorname{Re} s > 1$. Here, p means that we are summing only over the prime numbers.

We will take the product formula

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right),$$

and take the logarithm of both sides and differentiate. On the left hand side, we obtain

$$\frac{d}{ds} \log \frac{1}{\zeta(s)} = \frac{d}{ds} \log \zeta(s)^{-1} = -\frac{d}{ds} \log \zeta(s) = -\frac{\zeta'(s)}{\zeta(s)}.$$

The right hand side becomes

$$\begin{aligned} \frac{d}{ds} \log \prod_p \left(1 - \frac{1}{p^s}\right) &= \frac{d}{ds} \sum_p \log \left(1 - \frac{1}{p^s}\right) \\ &= \frac{d}{ds} \sum_p \log \frac{p^s - 1}{p^s} \\ &= \frac{d}{ds} \sum_p \log(p^s - 1) - s \log p \\ &= \sum_p \frac{p^s \log p}{p^s - 1} - \log p \\ &= \sum_p \frac{\log p}{p^s - 1}. \end{aligned}$$

Noticing that

$$\frac{\log p}{p^s - 1} - \frac{\log p}{p^s} = \frac{\log p}{p^s(p^s - 1)},$$

we conclude that

$$-\frac{\zeta'(s)}{\zeta(s)} = \Phi(s) + \sum_p \frac{\log p}{p^s(p^s - 1)}.$$

The left-hand side is meromorphic on $\mathbb{C}$, and the sum on the right converges to an analytic function for $\operatorname{Re} s > \frac{1}{2}$ using the p-test. Thus, we can extend

$\Phi(s)$ to be meromorphic for $\operatorname{Re} s > \frac{1}{2}$. The poles of $\Phi(s)$ are at the poles and zeros of $\zeta(s)$. In particular, $\Phi(s)$ has a simple pole at $s = 1$ with residue 1 (from the lone simple pole of $\zeta(s)$) and a simple pole at $s = s_0$ with residue $-m$ if $\zeta(s)$ has a zero or order m at s_0.

Theorem 6.2 (p. 383)**.** *The meromorphic function $\Phi(s)$ has no poles on the vertical line $\operatorname{Re} s = 1$ except at $s = 1$. The zeta function $\zeta(s)$ has no zeros on the line $\operatorname{Re} s = 1$.*

Proof. By the observation before the theorem, the two statements are equivalent. Let $t > 0$, and suppose that $\zeta(s)$ has a zero of order q at $1 + it$ and a zero of order q' at $1 + 2it$. Here, if $1 + it$ or $1 + 2it$ are non-zero, then we use the convention that $q = 0$ and $q' = 0$, respectively.

For a prime p, note that

$$0 \leq \left(p^{it/2} + p^{-it/2}\right)^4 = p^{2it} + 4p^{it} + 6 + 4p^{it} + p^{-2it}.$$

Multiply by $\frac{\epsilon(\log p)}{p^{1+\epsilon}}$ to obtain

$$0 \leq \epsilon[\frac{\log p}{p^{1+\epsilon-2it}} + \frac{4\log p}{p^{1+\epsilon-it}} + \frac{6\log p}{p^{1+\epsilon}} + \frac{4\log p}{p^{1+\epsilon+it}} + \frac{\log p}{p^{1+\epsilon+2it}}].$$

Summing over all primes then yields

$$0 \leq \epsilon[\Phi(1+\epsilon+2it)+4\Phi(1+\epsilon+it)+6\Phi(1+\epsilon)+4\Phi(1+\epsilon-it)+\Phi(1+\epsilon-2it)].$$

By Rule 1, if we take $\epsilon \to 0$, we obtain

$$0 \leq \operatorname{Res}[\Phi(z), 1+2it]+4\operatorname{Res}[\Phi(z), 1+it]+6\operatorname{Res}[\Phi(z), 1]+4\operatorname{Res}[\Phi(z), 1-it]+\operatorname{Res}[\Phi(z), 1-2it].$$

The residue at $z = 1$ is 1, and by assumption, the residue at $z = 1 \pm it$ is q and at $z = 1 \pm 2it$ the residue is q'. Hence, the above becomes,

$$0 \leq 6 - 8q - 2q',$$

which can only be satisfied if $q = q' = 0$. This proves the theorem. $\square$

7 Proof of the Prime Number Theorem

Now, we come back to the Laplace transform and connect it to $\Phi(s)$.

Lemma 7.1. *We have that*

$$(\mathcal{L}\Theta(e^t))(s) = \frac{\Phi(s)}{s}$$

for $\operatorname{Re} s > 1$.

Proof. The first part of the statement is actually showing that the Laplace transform exists for $\operatorname{Re} s > 1$. After that, we will show that the formula holds.

We first claim that $\Theta(x) \leq (4\log 2)x$. To see this, we will use the binomial coefficient

$$\binom{2n}{n} < (1+1)^{2n} = 2^{2n}.$$

Every prime number between n and $2n$ divides $\binom{2n}{n}$, hence their product also divides $\binom{2n}{n}$. This means that

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

Taking logarithms yields

$$\sum_{n<p<2n} \log p \leq 2n \log 2.$$

This means that

$$\begin{aligned}\Theta(2^m) &= \sum_{k=1}^{m} \sum_{2^{k-1}<p<2^k} \log p \\ &\leq \sum_{k=1}^{m} (2^k) \log 2 \\ &< 2^{m+1} \log 2.\end{aligned}$$

Now for any x, choose m such that $2^{m-1} < x \leq 2^m$. Then,

$$\Theta(x) \leq \Theta(2^m) \leq 2^{m+1} \log 2 = 4(2^{m-1}) \log 2 < (4\log 2)x.$$

Now, if we look at the Laplace transform,

$$\begin{aligned}
|(\mathcal{L}\Theta(e^t))(s)| &= |\int_0^\infty e^{-ts}\Theta(e^t)dt| \\
&\leq \int_0^\infty |e^{-ts}\Theta(e^t)|dt \\
&\leq \int_0^\infty |e^{-ts}||(4\log 2)e^t|dt \\
&= \int_0^\infty (4\log 2)e^{-t(\operatorname{Re} s-1)}dt.
\end{aligned}$$

This integral converges when $\operatorname{Re} s - 1 > 0$, i.e. when $\operatorname{Re} s > 1$.

To obtain the formula, we enumerate the primes so that p_n is the n-th prime number. Then, $\Theta(e^t)$ is constant for $\log p_n < t < \log p_{n+1}$. Therefore,

$$\begin{aligned}
\int_{\log p_n}^{\log p_{n+1}} e^{-st}\Theta(e^t)dt &= \Theta(p_n)\int_{\log p_n}^{\log p_{n+1}} e^{-st}dt \\
&= \Theta(p_n)\frac{e^{-st}}{-s}\Big|_{\log p_n}^{\log p_{n+1}} \\
&= \frac{1}{s}\Theta(p_n)(p_n^{-s} - p_{n+1}^{-s}).
\end{aligned}$$

Now, we break up the integral $\int_0^\infty e^{-st}\Theta(e^t)dt$ along intervals $[\log p_n, \log p_{n+1}]$ and use the above identity to show

$$\begin{aligned}
\int_0^\infty e^{-st}\Theta(e^t)dt &= \sum_n^\infty \frac{1}{s}\Theta(p_n)(p_n^{-s} - p_{n+1}^{-s}) \\
&= \frac{1}{s}\sum_n^\infty(\Theta(p_n) - \Theta(p_{n-1}))p_n^{-s}.
\end{aligned}$$

Observe that $\Theta(p_n) - \Theta(p_{n-1}) = \log p_n$, and the above becomes

$$\frac{1}{s}\sum_n^\infty \frac{\log p_n}{p_n^s} = \frac{1}{s}\Phi(s).$$

□

Notice that this identity holds for all $\operatorname{Re} s > 1$. We have previously shown that the only pole of $\Phi(s)$ on the line $\operatorname{Re} s = 1$ is the simple pole at $s = 1$. If we can remove this pole, then we can extend $(\mathcal{L}\Theta(e^t))(s)$ past the line $\operatorname{Re} s = 1$.

Now we consider the function $h(t) = \Theta(e^t)e^{-t} - 1$. If we look at the Laplace transform $(\mathcal{L}h)(s)$, the effect of multiplying e^{-t} is to change that e^{-st} in the integrand to $e^{-(s+1)t}$. Hence, the Laplace transform will converge for $\operatorname{Re} s + 1 > 1$, or in other words $\operatorname{Re} s > 0$. The Laplace transform of 1 is

$$\int_0^\infty e^{-st} dt = \frac{e^{-st}}{-s}\Big|_0^\infty = \frac{1}{s}.$$

Hence,

$$(\mathcal{L}h)(s) = \frac{\Phi(s+1)}{s+1} - \frac{1}{s},$$

for $\operatorname{Re} s > 0$. Now, $\Phi(s+1)$ has a simple pole with residue 1 at $s = 0$, but the $-\frac{1}{s}$ part cancels that out. Hence, $(\mathcal{L}h)(s)$ is analytic for $\operatorname{Re} s > 0$, and it has no poles on the line $\operatorname{Re} s = 0$. This means that we can extend the function to be analytic on a domain that contains the line $\operatorname{Re} s = 0$. In particular, $(\mathcal{L}h)(0)$ exists, proving the prime number theorem. This last part of extending the function is not trivial and is shown in the book on p. 385-387, but this gives a general idea of how the proof works.

8 Exercises

1. Exercise XIV.1.2
2. Exercise XIV.2.1
3. Exercise XIV.2.2
4. Exercise XIV.5.5 (this is fairly hard, but see if you can reduce it to the problem of showing that $\lim_{n\to\infty} \frac{\log p_n}{\log n} = 1$)