The Prime Number Theorem

Eric Nordenstam Dan Sjöstrand Tomas Torstensson

November 24, 2000

Abstract

The Prime Number Theorem states that the n:th prime number is approximately $n \log n$. An important part of the proof is the "Hadamard-de la Vallée-Poussin" theorem, which states that the Riemann ζ-function has no zeros in the half-plane $\Re(z) \geq 1$ and only a simple pole at point 1.

1 Introduction

The Riemann ζ-function is defined by

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} = \prod_p \left(1 - \frac{1}{p^z}\right)^{-1} \tag{1}$$

and we also introduce the function

$$\Phi(z) = \sum_p \frac{\log p}{p^z} \tag{2}$$

We also need the following functions:

$$\varphi(x) = \sum_p \log p \tag{3}$$

and

$$\pi(x) = \sum_p 1 \tag{4}$$

The proof can be divided into the following steps:

1. We show that both $\zeta(z)$ and $\Phi(z)$ converge uniformly and absolutely when $\Re(z) \geq 1 + \varepsilon$.
2. The function $\zeta(z) - \frac{1}{z-1}$ extends to a regular function for $\Re(z) > 0$. It has the symmetry $\overline{\zeta(z)} = \zeta(\bar{z})$.
3. We show that $-\dfrac{\zeta'(z)}{\zeta(z)} = \Phi(z) + \sum_p h_p(z)$ for $h_p = \dfrac{\log p}{p^{2z} - p^z}$.
4. Show that the function $\Phi(z)$ is meromorphic, i.e. it is analytic with isolated poles for $\Re(z) > \frac{1}{2}$ and that it has poles only at point 1 and at zeros of $\zeta(z)$.
5. Show that $\zeta(z)$ has no zeros within $\Re(z) \geq 1$ and hence $\Phi(z)$ has only a simple pole at point 1 in this area.
6. We show that $\varphi(x) < 2x$.
7. Show that the integral $\displaystyle\int_1^{\infty} \frac{\varphi(x) - x}{x^2} dx$ converges.
8. Show that $\dfrac{\varphi(x)}{x} \to 1$, $x \to \infty$ and $\dfrac{\pi(x)\log x}{x} \to 1$, $x \to \infty$.

Step 1

Show that $\zeta(z)$ converges absolutely and uniformly. We use the integral test.

$$\int_1^{\infty} \left|\frac{1}{n^z}\right| dn = \int_1^{\infty} \left|\frac{1}{n}\right|^{\Re(z)} dn \tag{5}$$

$$= \int_1^{\infty} \frac{1}{n^{\Re(z)}} dn = \tag{6}$$

$$= \left[\frac{1}{(1 - \Re(z)) n^{\Re(z)-1}}\right]_1^{\infty} = \tag{7}$$

$$= \{\Re(z) > 1\} = -\frac{1}{1 - \Re(z)} \tag{8}$$

Thus it converges absolutely.

Now let us show that it converges uniformly. Choose a domain $D = \{z : 1 + \varepsilon < \Re(z) < N\}$ for some large N. Let $\zeta_n(z) = \sum_{k=1}^{n} \frac{1}{k^z}$. We use Weierstrass M-test:

$$|\zeta_n(z)| \leq \sum_{k=1}^{n} \left|\frac{1}{k^z}\right| \leq \sum_{k=1}^{n} \left|\frac{1}{k}\right|^{\Re(z)} \leq \sum_{k=1}^{n} \left|\frac{1}{k}\right|^{1+\varepsilon} = \sum_{k=1}^{n} M_n \tag{9}$$

We need to show that $\sum M_n$ converges. But this converges because of the proof for the absolute convergence of $\zeta(z)$.

Now we must show that $\Phi(z) = \sum_p \frac{\log p}{p^z}$ converges. In stead of summing over all prime numbers, we can sum over all natural numbers. If this new sum converges, then $\Phi(z)$ will converge. Again we use the integral test.

$$\int_1^\infty \left|\frac{\log n}{n^z}\right| dn = \int_1^\infty \frac{\log n}{n^{\Re(z)}} dn = \tag{10}$$

$$= \left[\frac{1}{-\Re(z)-1}\frac{1}{n^{\Re(z)-1}}\left(\log n - \frac{1}{-\Re(z)+1}\right)\right]_1^\infty \tag{11}$$

The resulting expression is bounded.

Does it also converge uniformly? Again we use the Weierstrass M-test, and again we use the trick of summing over all natural numbers instead of over all primes.

$$\sum_n \left|\frac{\log n}{n^z}\right| = \sum_n \frac{\log n}{n^{\Re(z)}} \leq \sum_n \frac{\log n}{n^{1+\varepsilon}} \tag{12}$$

This converges by proof above.

Step 2

We now wish to show that $\zeta(z) - \frac{1}{z-1}$ converges in $\Re(z) > 0$.

$$\zeta(z) - \frac{1}{z-1} = \sum_{n=1}^\infty \frac{1}{n^z} - \int_1^\infty \frac{1}{x^z} dx = \sum_{n=1}^\infty \int_n^{n+1} \left(\frac{1}{n^z} - \frac{1}{x^z}\right) dx \tag{13}$$

By the ML-inequality we have

$$\sum_{n=1}^\infty \int_n^{n+1} \left|\frac{1}{n^z} - \frac{1}{x^z}\right| dx \leq \sum_{n=1}^\infty 1 \cdot \max_{x\in[n,n+1]} \left|\frac{1}{n^z} - \frac{1}{x^z}\right| \tag{14}$$

$$\leq \sum_{n=1}^\infty 1 \cdot 1 \cdot \max_{x\in[n,n+1]} \frac{d}{dx}\left|\frac{1}{n^z} - \frac{1}{x^z}\right| \tag{15}$$

$$\leq \sum_{n=1}^\infty \frac{|z|}{n^{\Re(z)+1}} \tag{16}$$

$$\tag{17}$$

Which converges according to step 1, where we showed that $\sum 1/n^z$ converges for $\Re(z) > 1 + \varepsilon$.

The function $\zeta(z)$ has the following symmetry: $\overline{\zeta(z)} = \zeta(\bar{z})$. This is obvious by

$$\overline{\sum \frac{1}{n^z}} = \sum \overline{\left(\frac{1}{n^z}\right)} = \sum \frac{1}{n^{\bar{z}}} \tag{18}$$

Also, ζ' has this same symmetry. For any analytic f such that $f(\bar{z}) = \overline{f(z)}$, $f(a+h) = f(a) + f'(a)(h) + o(h)$, which gives $f(\bar{a}+\bar{h}) = f(\bar{a}) + f'(a)(\bar{h}) + o(h)$. But we know that $f(\bar{a}+\bar{h}) = f(\bar{a}) + \overline{f'(a)}(\bar{h}) + o(h)$ which implies $\overline{f'(a)} = f'(\bar{a})$.

Step 3

By logarithmating the product definition of $\zeta(z)$ we know that

$$\log \zeta(z) = -\sum_p log\left(1 - \frac{1}{p^z}\right) \tag{19}$$

By differentiating this formula we get:

$$\frac{\zeta'(z)}{\zeta(z)} = -\sum_p \frac{1}{1-\frac{1}{p^z}} \frac{\log p}{p^z} = \tag{20}$$

$$= -\sum_p \frac{1-\frac{1}{p^z}+\frac{1}{p^z}}{1-\frac{1}{p^z}} \frac{\log p}{p^z} \tag{21}$$

$$= -\sum_p \left(1+\frac{1}{p^z-1}\right) \frac{\log p}{p^z} = \tag{22}$$

$$= -\sum_p \frac{\log p}{p^z} - \sum_p \frac{\log p}{p^z(p^z-1)} \tag{23}$$

So we now know that

$$-\frac{\zeta'(z)}{\zeta(z)} = \Phi(z) + \sum_p \frac{\log p}{p^z(p^z-1)} \tag{24}$$

For future convenience we introduce $h_p(z) = \frac{\log p}{p^z(p^z-1)}$.

Step 4

We want to prove that $\Phi(z)$ is meromorphic (analytic with isolated poles) for $\Re(z) > \frac{1}{2}+\varepsilon$. Show that it has poles only at point 1 and zeros of $\zeta(z)$.

From step 3 we know that

$$\Phi(z) = -\frac{\zeta'(z)}{\zeta(z)} - \sum_p h_p(z) \tag{25}$$

and from step 2 we know that

$$\zeta(z) = \frac{1}{z-1} + \text{ a nice function} \tag{26}$$

$$\Rightarrow \qquad \zeta'(z) = -\frac{1}{(z-1)^2} + \text{ nice} \tag{27}$$

$$\Rightarrow \qquad \frac{\zeta'(z)}{\zeta(z)} = -\frac{1}{z-1} + \text{ something} \tag{28}$$

This shows that $\Phi(z)$ has a pole at 1.

If $\zeta(z)$ has a zero then we can write $\zeta(z) = g(z)(z-z_0)^m$ where $g(z)$ is an analytic function.

$$\zeta(z) = g(z)(z-z_0)^m \tag{29}$$

$$\Rightarrow \qquad \zeta'(z) = g'(z)(z-z_0)^m + g(z)m(z-z_0)^{m-1} \tag{30}$$

$$\Rightarrow \qquad \frac{\zeta'(z)}{\zeta(z)} = \frac{g'(z)(z-z_0)^m + g(z)m(z-z_0)^{m-1}}{g(z)(z-z_0)^m} = \frac{g'(z)}{g(z)} + m\frac{1}{z-z_0} \tag{31}$$

So $\Phi(z)$ will have poles of order 1 at zeros of $\zeta(z)$.

The error terms converge absolutely for $\Re(z) > \frac{1}{2}+\varepsilon$.

$$\sum_{\text{primes } p} \left|\frac{\log p}{p^z(p^z-1)}\right| \leq \sum_{n=0}^{\infty} \left|\frac{\log n}{n^{2z}-n^z}\right| \tag{32}$$

We use the comparison test, comparing this series to

$$\sum_{n=0}^{\infty} \left|\frac{\log n}{n^{2z}}\right| \tag{33}$$

which converges for $\Re(z) > \frac{1}{2}+\varepsilon$ according to step 1.

The comparison turns out like this:

$$\lim_{n\to\infty} \frac{\left|\frac{\log n}{n^{2z}-n^z}\right|}{\left|\frac{\log n}{n^{2z}}\right|} = \lim_{n\to\infty} \left|\frac{n^{2z}}{n^{2z}-n^z}\right| = 1 \tag{34}$$

So the error terms converge absolutely. To show that they converge uniformly, we use the Weierstrass M-test.

$$\left|\frac{\log n}{n^{2z}-n^z}\right| \leq \left|\frac{\log n}{n^{1+2\varepsilon}-n^{1/2+\varepsilon}}\right| \tag{35}$$

The infinite sum of which converges by proof above. Thereby we know that the error terms converge absolutely and uniformly, so no poles can come from them. Therefore $\Phi(z)$ has a pole at 1 and at zeros of $\zeta(z)$.

Step 5

We wish to show that $\zeta(z)$ has no zeros for $\Re(z) \geq 1$ and hence $\Phi(z)$ has only a simple pole at point 1.

For $\Re(z) > 1$ this follows from the product definition of $\zeta(z)$. We only need to check $\Re(z) = 1$.

If analytic function $f(z)$ has a zero of order $-m$ at point a then $f'(z)/f(z)$ has a simple pole of residue m at that point. This is easy to check: We can write $f(z) = g(z)(z-a)^{-m}$ where $g(z)$ analytic and $g(a) \neq 0$. This gives $f'(z) = g'(z)(z-a)^{-m} + g(z)(-m)(z-a)^{-m-1}$ which implies

$$f'(z)/f(z) = \frac{g'(z)(z-a)^{-m} + g(z)(-m)(z-a)^{-m-1}}{g(z)(z-a)^{-m}} = -m\frac{1}{z-a} + \frac{g'(z)}{g(z)} \tag{36}$$

which has residue $-m$ at a.

To prove that $\zeta(z)$ has no zeros on $\Re(z) = 1$ we assume that it does have zeros there and try to derive a contradiction. Assume $\zeta(z)$ has a zero of order $m \geq 1$ at point $z = 1 + ib$ with $b \neq 0$. Then, since $\overline{\zeta(z)} = \zeta(\bar{z})$ we know that has a zero of the same order at $z = 1 - ib$. Also, let n be the order of the zero at $z = 1 + 2ib$ or 0 if there is no zero there.

We need to evaluate $\mathrm{Res}[\Phi(z), 1]$. We know that

$$\Phi(z) = -\frac{\zeta'(z)}{\zeta(z)} - \sum h_p(z) \tag{37}$$

The error terms converged uniformly so they have no pole and will not give any residue at 1. We also know that

$$\zeta(z) = \frac{1}{z-1} + \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^z} - \frac{1}{x^z}\right) dx \tag{38}$$

The sum in the above expression is nice: it has no pole at 1. So

$$\zeta'(z) = -\frac{1}{(z-1)^2} + \text{analytic with no pole at } 1 \tag{39}$$

which gives

$$\frac{\zeta'(z)}{\zeta(z)} = -\frac{1}{z-1} + \text{analytic with no pole at } 1 \tag{40}$$

So we see that $\mathrm{Res}[\Phi(z), 1] = 1$. The error terms have no poles at all at $\Re(z) = 1$. Also, $\overline{\zeta'(z)} = \zeta'(\bar{z})$, so $\mathrm{Res}[\Phi(z), i \pm ib] = \mathrm{Res}[\zeta'(z)/\zeta(z), 1 \pm ib] = -m$ and $\mathrm{Res}[\Phi(z), i \pm 2ib] = \mathrm{Res}[\zeta'(z)/\zeta(z), 1 \pm 2ib] = -n$.

Now, we are supplied with a trick. We are given the following identity

$$\sum_p \frac{\log p}{p^z}(p^{ib/2} + p^{-ib/2}) = \Phi(z - 2ib) + 4\Phi(z - ib) + 6\Phi(z) + 4\Phi(z + ib) + \Phi(z + 2ib) \tag{41}$$

It is easy to verify this equality:

$$\sum_p \frac{\log p}{p^z}(p^{ib/2}+p^{-ib/2}) = \sum_p \frac{\log p}{p^z}(p^{2ib}+4p^{ib}+6p^0+4p^{-ib}+p^{-2ib}) \tag{42}$$

$$= \sum_p \frac{\log p}{p^{z-2ib}} + 4\sum_p \frac{\log p}{p^{z-ib}} + 6\sum_p \frac{\log p}{p^z} + 4\sum_p \frac{\log p}{p^{z+ib}} + \sum_p \frac{\log p}{p^{z+2ib}} \tag{43}$$

$$= \Phi(z-2ib)+4\Phi(z-ib)+6\Phi(z)+4\Phi(z+ib)+\Phi(z+2ib) \tag{44}$$

The left hand side of this equality is positive for $z = 1+\varepsilon$ since $\log p$ is positive, p^z is positive and $(p^{ib/2}+p^{-ib/2})^4 = (e^{ib(\log p)/2}+e^{-ib(\log p)/2})^4 = (2\cos(b(\log p)/2))^4 > 0$. By adding residues of each side of the equality we get $6-8m-2n \geq 0$ which gives $m = 0$, but this is a contradiction since we assumed $m > 0$. Therefore $\zeta(z)$ has no zeros on the line $\Re(z) = 1$.

Step 6

We want to show that $\varphi(x) < 2x$. To do this, we must first prove that

$$\prod_{\text{primes } p<n} p \leq 4^n \tag{45}$$

$$2^n = \sum_{j=0}^{n}\binom{n}{j} \geq \binom{n}{n/2} = \frac{n(n-1)(n-2)\dots(n/2+1)}{(n/2)!} = \frac{(n/2)!(\prod_{n/2<p<n} p)f(n)}{(n/2)!} \geq \prod_{n/2<p<n} p \tag{46}$$

$\prod_{p\leq n/2} p$ should be smaller than $\prod_{n/2<p\leq n} p$, so we get

$$\prod_{p<n} p \leq 4^n \tag{47}$$

Lets have a look at $\exp(\varphi(x))$.

$$\exp(\varphi(x)) = \exp(\log p_1 + \log p_2 + \dots) = \exp(\log(p_1p_2p_3\dots)) = \prod_{p\leq x} p \leq 4^x = \exp(x\log 4) \leq \exp(2x) \tag{48}$$

So we see that $\varphi(x) < 2x$.

Step 7

Now we want to show that the following integral converges.

$$\int_1^\infty \frac{\varphi(x)-x}{x^2}dx \tag{49}$$

First let us check that

$$\Phi(z) = z\int_1^\infty \frac{\varphi(x)}{x^{z+1}}dx \tag{50}$$

by rewriting the integral as a sum of integrals such that $\varphi(x)$ is constant over each little integral.

$$z\int_1^\infty \frac{\varphi(x)}{x^{z+1}}dx = z\varphi(0)\int_1^{p_1}\frac{1}{x^{z+1}}dx + z\varphi(p_1)\int_{p_1}^{p_2}\frac{1}{x^{z+1}}dx + z\varphi(p_2)\int_{p_2}^{p_3}\frac{1}{x^{z+1}}dx + \dots \tag{51}$$

$$= z\varphi(1)\left[-\frac{1}{zx^z}\right]_1^{p_1} + z\varphi(p_1)\left[-\frac{1}{zx^z}\right]_{p_1}^{p_2} + z\varphi(p_2)\left[-\frac{1}{zx^z}\right]_{p_2}^{p_3} + \dots \tag{52}$$

$$= \log(p_1)\left(\frac{1}{p_1^z}-\frac{1}{p_2^z}\right) + \log(p_1p_2)\left(\frac{1}{p_2^z}-\frac{1}{p_3^z}\right) + \cdots \tag{53}$$

$$= \frac{\log p_1}{p_1^z} + \frac{\log p_2}{p_2^z} + \cdots = \Phi(z) \tag{54}$$

Next we want to show that

$$\frac{\Phi(z+1)}{z+1} - \frac{1}{z} = \mathcal{L}\{f\} \tag{55}$$

where

$$f(u) = \frac{\varphi(e^u) - e^u}{e^u} \tag{56}$$

which gives

$$\frac{\Phi(z+1)}{z+1} - \frac{1}{z} = \int_1^\infty \frac{\varphi(x)}{x^{z+2}} dx - \frac{1}{z} \tag{57}$$

Now we will calculate $\mathcal{L}\{f\}$.

$$\mathcal{L}\{f\} = \mathcal{L}\left\{\frac{\varphi(e^u) - e^u}{e^u}\right\} = \mathcal{L}\left\{\frac{\varphi(e^u)}{e^u}\right\} + \mathcal{L}\{1\} = G(z) + \frac{1}{z} \tag{58}$$

What is $G(z)$? The definition of Laplace transform gives us

$$G(z) = \int_0^\infty e^{-\omega z} \frac{\varphi(e^\omega)}{e^\omega} d\omega = \int_0^\infty \frac{\varphi(e^\omega)}{e^{\omega(z+1)}} d\omega = \left\{\begin{matrix} t = e^\omega \\ \frac{1}{t}dt = d\omega \end{matrix}\right\} = \int_1^\infty \frac{\varphi(t)}{t^{z+2}} dt = \frac{\Phi(z+1)}{z+1} \tag{59}$$

Thus the integral at the start of this step converges because

$$\int_1^\infty \frac{\varphi(x) - x}{x^2} dx = \left\{\begin{matrix} x = e^\omega \\ dx = e^\omega d\omega \end{matrix}\right\} = \int_0^\infty \frac{\varphi(e^\omega) - e^\omega}{e^\omega} d\omega = \mathcal{L}\{f\}(0) \tag{60}$$

Step 8

Now we want to prove that $\frac{\varphi(x)}{x} \to 1$, $x \to \infty$. We assume that that is not so and try to prove a contradiction. Take $\varepsilon > 0$. Suppose there exists n such that $\forall x > n : \frac{\varphi(x)}{x} > 1 + \varepsilon$. The the integral from the previous step can be adjusted like this:

$$\int_1^\infty \frac{\varphi(x) - x}{x^2} dx \geq \int_1^n \frac{\varphi(x) - x}{x^2} dx + \int_n^\infty \frac{(1+\varepsilon)x - x}{x^2} dx \to +\infty \tag{61}$$

But we know that the integral converges, therefore we have a contradiction. In the same way, suppose there exists n such that $\forall x > n : \frac{\varphi(x)}{x} < 1 - \varepsilon$. Then

$$\int_1^\infty \frac{\varphi(x) - x}{x^2} dx \leq \int_1^n \frac{\varphi(x) - x}{x^2} dx + \int_n^\infty \frac{(1-\varepsilon)x - x}{x^2} dx \to -\infty \tag{62}$$

but we knew that the integral converged, therefore we have a contradiction. Thereby we know that for every $\varepsilon > 0$ there exists an n such that $\forall x > n : 1 - \epsilon < \frac{\varphi(x)}{x} < 1 + \varepsilon$. We then get

$$\frac{\varphi(x)}{x} \to 1, \quad x \to \infty \tag{63}$$

Now we wish to prove that $\frac{\pi(x)\log(x)}{x} \to 1$, $x \to \infty$.

$$\varphi(x) = \sum_{p_k < x} \log p_k = \sum_{k \leq \pi(x)} \log p_k \leq \pi(x)\log(x) \tag{64}$$

$$\varphi(x) \geq \sum_{x^{1-\varepsilon} \leq p_k \leq x} \log p_k \geq (1-\varepsilon)(\log x)(\pi(x) - \pi(x^{1-\varepsilon})) \geq \{\pi(n) \leq n\} \geq (1-\varepsilon)(\log x)(\pi(x) - x^{1-\varepsilon}) \tag{65}$$

This gives us

$$(1-\varepsilon)(\log x)(\pi(x)-x^{1-\varepsilon}) \leq \varphi(x) \leq \pi(x)\log(x) \tag{66}$$

$$\Rightarrow \quad \frac{\varphi(x)}{x} \leq \frac{\pi(x)\log(x)}{x} \leq \frac{1}{1-\varepsilon}\frac{\varphi(x)}{x} + \frac{\log x}{x^{\varepsilon}} \tag{67}$$

$$\Rightarrow \quad 1 \leq \lim \frac{\pi(x)\log(x)}{x} \leq \frac{1}{1-\varepsilon} \tag{68}$$

$$\tag{69}$$

So we deduce that

$$\frac{\pi(x)\log(x)}{x} \to 1, \quad x \to \infty \tag{70}$$

Last but not least we wish to prove that the n:th prime number is approximately $n\log n$, i.e. that

$$\frac{\pi(n\log n)}{n} \to 1, \quad n \to \infty \tag{71}$$

This is just a matter of manipulating our last limit.

$$1 = \lim_{n\to\infty} \frac{\pi(n\log n)\log(n\log n)}{n\log n} \tag{72}$$

$$= \lim_{n\to\infty} \frac{\pi(n\log n)}{n\log n}(\log n + \log\log n) \tag{73}$$

$$= \lim_{n\to\infty} \left(\frac{\pi(n\log n)}{n} + \frac{\pi(n\log n)}{n\log n}\log\log n)\right) \tag{74}$$

$$\tag{75}$$

But the last term in the last limit tends to zero, because, by L'Hôpital

$$\lim_{n\to\infty} \frac{\log\log n}{\log n} = \lim_{n\to\infty} \frac{1}{\log n} = 0 \tag{76}$$

and

$$\lim_{n\to\infty} \frac{\pi(n\log n)}{n} = 1 \tag{77}$$

Conclusion: $\pi(n\log n) \approx n$.