The Prime Number Theorem

Charles Alley

1 Introduction

In analytic number theory, it is all too often the case that the details of proofs are left as exercises for the reader (usually branded with some off putting adjective such as *easy* or *obvious*). In these notes, I will attempt to give an honest to goodness proof of the Prime Number Theorem, with all the details written out (to the extent that this is possible). My reference in writing this is the foundational text *Multiplicative Number Theory, 3rd ed.* by H. Davenport as revised by H.L. Montgomery.

Of course, we must assume some things. For example, I will freely use certain analytic techniques common to number theory such as partial summation (also known as Abel's formula) and a fair amount of complex analysis, without justification. However, when it comes to details related to the theorem itself, I will do my best to be complete. These notes are mostly for my own reference. That said, to any readers who may come across this, I welcome your comments, especially regarding typos (both mathematical and typographical).

Let's begin with some definitions and a clear statement of what we'll prove. In what follows, p will always denote a prime number, $\mathbb{P}$ the set of primes, and $\mathbb{1}_X$ the characteristic function of a given set X. Given a complex valued function $f(x)$ and a positive, real valued function $g(x)$ of a real variable x, the notation $f(x) = O(g(x))$, or equivalently $f(x) \ll g(x)$, means that for some real number x_0 and some constant $M > 0$, which does not depend on x, we have $|f(x)| < Mg(x)$ for all $x > x_0$. Writing $f(x) = o(g(x))$ means that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0,$$

while the notation $f(x) \sim g(x)$ means that

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

Definition 1. *Given a real number $x \geq 2$, we define the prime counting function $\pi(x)$ as the number of primes less than x, i.e.*

$$\pi(x) = \sum_{n\leq x} \mathbb{1}_{\mathbb{P}}(n) = \sum_{p\leq x} 1.$$

Theorem 1 (Prime Number Theorem)**.** *There exists an absolute, positive constant* $c < 1$ *such that for* $x \geq 3$ *we have*

$$\pi(x) = \mathrm{li}(x) + O(xe^{-c\sqrt{\log x}}). \tag{1}$$

In particular, this gives

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

Here, log denotes the natural logarithm and $\mathrm{li}(x)$ denotes the logarithmic integral:

$$\mathrm{li}(x) = \int_2^x \frac{dt}{\log t}.$$

We can prove that $\mathrm{li}(x) \sim \frac{x}{\log x}$ using integration by parts. We have

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

and so we are left with showing that $\int_2^x 1/(\log^2 t)\, dt = o(\frac{x}{\log x})$. To see this, write

$$\begin{aligned}\int_2^x \frac{dt}{\log^2 t} &= \int_2^e \frac{dt}{\log^2 t} + \int_e^{x^{1/2}} \frac{dt}{\log^2 t} + \int_{x^{1/2}}^x \frac{dt}{\log^2 t} \\ &= \int_e^{x^{1/2}} \frac{dt}{\log^2 t} + \int_{x^{1/2}}^x \frac{dt}{\log^2 t} + O(1).\end{aligned}$$

On the interval $e \leq t \leq x^{1/2}$ we have $\log t \geq 1$ and on the interval $x^{1/2} \leq t \leq x$ we have $\log t \geq \frac{1}{2}\log x$. Using these estimates in the above expression gives

$$\int_2^x \frac{dt}{\log^2 t} \leq \int_e^{x^{1/2}} 1\, dt + \int_{x^{1/2}}^x \frac{4}{\log^2 x}\, dt + O(1) \leq x^{1/2} + \frac{4x}{\log^2 x} + O(1).$$

If we now multiply both sides by $\frac{\log x}{x}$ and take the limit as x goes to infinity, we have

$$\lim_{x\to\infty} \frac{\log x}{x} \int_2^x \frac{dt}{\log^2 t} \leq \lim_{x\to\infty} \left(\frac{\log x}{x^{1/2}} + \frac{4}{\log x} + O\left(\frac{\log x}{x}\right)\right) = 0.$$

It now follows from (2) that $\mathrm{li}(x) \sim \frac{x}{\log x}$. (Note: one can prove that $\int_2^x 1/(\log^2 t)\, dt = o(\frac{x}{\log x})$ using L'H'ôpital's rule, though this proof seems somehow less satisfying).

Next, observe that

$$\lim_{x\to\infty} \frac{(\log x)^n}{e^{c\sqrt{\log x}}} = 0; \tag{3}$$

which can be seen by setting $u = \log x$ and recalling that $u^n/e^{cu} \to 0$ as $u \to \infty$. In particular,

$$\frac{xe^{-c\sqrt{\log x}}}{\mathrm{li}(x)} \sim \frac{xe^{-c\sqrt{\log x}}}{x/\log x} \to 0 \text{ as } x \to \infty$$

proving the last assertion of Theorem 1.

Now, our goal is to develop an analytic proof of (1). But the function $\pi(x)$ is a poor function from an analytic standpoint: it is a step function with discontinuities at the primes. To remedy this, we introduce a more analytically appealing function, still a step function but one which we can relate to the (smooth) Riemann zeta function.

Definition 2 (Tchebychev's function)**.** *Consider the function*

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^k,\ k \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

and define

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

Lemma 1. *The Prime Number Theorem is implied by the asymptotic expression*

$$\psi(x) = x + O(xe^{-c\sqrt{\log x}}).$$

Here, it is important to note that $c < 1$.

Proof. Consider the sum

$$\pi_1(x) = \sum_{n \leq x} \frac{\Lambda(n)}{\log n} = \sum_{p^m \leq x} \frac{\log p}{m \log p}.$$

Applying partial summation and the assumed asymptotic for $\psi(x)$ we have

$$\begin{aligned} \sum_{n \leq x} \frac{\Lambda(n)}{\log n} &= \frac{1}{\log x} \sum_{n \leq x} \Lambda(n) + \int_2^x \left(\frac{1}{t \log^2 t} \right) \sum_{n \leq t} \Lambda(n)\, dt \\ &= \frac{\psi(x)}{\log x} + \int_2^x \frac{\psi(t)\, dt}{t \log^2 t} \\ &= \frac{x}{\log x} + \int_2^x \frac{t\, dt}{t \log^2 t} + O\left(\frac{xe^{-c\sqrt{\log x}}}{\log x} + \int_2^x \frac{e^{-c\sqrt{\log t}}}{\log^2 t} dt \right). \end{aligned} \tag{4}$$

Integration by parts gives

$$\begin{aligned} \int_2^x \frac{t\, dt}{t \log^2 t} &= \int_2^x t \frac{d}{dt} \left(\frac{-1}{\log t} \right) dt = \left. \frac{-t}{\log t} \right|_2^x + \int_2^x \frac{dt}{\log t} \\ &= \frac{2}{\log 2} - \frac{x}{\log x} + \mathrm{li}(x). \end{aligned}$$

Inserting this into (4) gives a main term of $\mathrm{li}(x)$. For the error term we have

$$\begin{aligned}\int_2^x \frac{e^{-c\sqrt{\log t}}}{\log^2 t}dt &\le \frac{1}{\log^2 2}\int_2^x e^{-c\sqrt{\log t}}dt \\ &= \frac{1}{\log^2 2}\left(\int_2^{x^{1/4}} e^{-c\sqrt{\log t}}dt + \int_{x^{1/4}}^x e^{-c\sqrt{\log t}}dt\right). \qquad (5)\end{aligned}$$

Now, since $(\log t)^{1/2} \ge \frac{1}{2}(\log x)^{1/2}$ for $t \ge x^{1/4}$, the second summand in (5) is less than

$$(x - x^{1/4})e^{-\frac{c}{2}\sqrt{\log x}} = O\left(xe^{-\frac{c}{2}\sqrt{\log x}}\right).$$

The first summand in (5) is $\le x^{1/4}$ simply because $e^{-c\sqrt{\log t}} \le 1$ for $t \ge 2$. We then observe, using the fact that $c < 1$, that $-\frac{c}{2}\sqrt{\log x} > -\frac{1}{2}\sqrt{\log x} > -\frac{1}{2}\log x$ and so

$$xe^{-\frac{c}{2}\sqrt{\log x}} > xe^{-\frac{1}{2}\log x} = x^{1/2} > x^{1/4}. \qquad (6)$$

Also,

$$\frac{xe^{-c\sqrt{\log x}}}{\log x} < xe^{-c\sqrt{\log x}} < xe^{-\frac{c}{2}\sqrt{\log x}}.$$

Putting these observations together we have

$$\pi_1(x) = \mathrm{li}(x) + O\left(xe^{-\frac{c}{2}\sqrt{\log x}}\right). \qquad (7)$$

Finally, we note that if $x \ge p^m \ge 2^m$ then $1 \le m \le \log_2 x$. This gives

$$\pi_1(x) = \sum_{p^m \le x} \frac{\log p}{m \log p} = \sum_{1 \le m \le \log_2 x} \sum_{p \le x^{1/m}} \frac{1}{m} = \pi(x) + \frac{1}{2}\pi(x^{1/2}) + \frac{1}{3}\pi(x^{1/3}) + \dots$$

Then, since $\pi(x^{1/2}) < x^{1/2}$, $\pi(x^{1/3}) < x^{1/3}$, $\dots$ we have

$$\begin{aligned}\sum_{2 \le m \le \log_2 x} \frac{1}{m}\pi(x^{1/m}) &\le \frac{x^{1/2}}{2} + \sum_{3 \le m \le \log_2 x} \frac{x^{1/m}}{m} \\ &< x^{1/2} + x^{1/3}\sum_{m \le x}\frac{1}{m}. \qquad (8)\end{aligned}$$

Here, we apply partial summation to get

$$\sum_{m \le x}\frac{1}{m} = \frac{1}{x}\sum_{m \le x} 1 + \int_1^x \frac{1}{t^2}\left(\sum_{m \le t} 1\right) dt = \log x + 1$$

and note that $\log x = o(x^{\epsilon})$ for any $\epsilon > 0$. With this we've shown that (8) is $< x^{1/2} + O(x^{1/3+\epsilon}) = O(x^{1/2})$ and so

$$\pi_1(x) = \pi(x) + O\left(x^{1/2}\right).$$

As was already observed in (6), $x^{1/2} = O\left(xe^{-\frac{c}{2}\sqrt{\log x}}\right)$ and so it follows from (7) that

$$\pi(x) = \mathrm{li}(x) + O\left(xe^{-\frac{c}{2}\sqrt{\log x}}\right).$$

□

It now remains to prove the asymptotic stated in Lemma 1.

2 The Asymptotic for $\psi(x)$

We wish to prove that

$$\psi(x) = x + O(xe^{-c\sqrt{\log x}}) \tag{9}$$

where $c < 1$. This is, it turns out, a rather deep fact and will require some careful analysis of the Riemann zeta function, $\zeta(s)$, the definition of which we now recall along with some fun facts to be proven later.

Definition 3 (Riemann Zeta Function)**.** *Let* $s \in \mathbb{C}$ *with* $\mathrm{Re}(s) > 1$ *and define*

$$\zeta(s) = \sum_{n \geq 1} n^{-s}.$$

Note that the series defining $\zeta(s)$ converges absolutely for $\mathrm{Re}(s) > 1$ by the integral test. For if we let $s = \sigma + it$, then

$$\int_1^{\infty} |x^{-s}|\,dx = \int_1^{\infty} x^{-\sigma}\,dx = \left.\frac{x^{1-\sigma}}{1-\sigma}\right|_1^{\infty} = \frac{1}{\sigma - 1} < \infty.$$

Fun Fact 1. *We have*

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

This product converges for $\mathrm{Re}(s) > 1$ *and in particular is non-zero.*

Fun Fact 2. $\zeta(s)$ *extends to a meromorphic function on* $\mathbb{C}$ *with its only pole at* $s = 1$ *of residue 1*

Fun Fact 3. *We have the functional equation*

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

where

$$\Gamma(s)=\int_0^\infty e^{-t}t^{s-1}\,dt.$$

The gamma function, $\Gamma(s)$, itself extends to all of $\mathbb{C}$ with poles at 0 and at the negative integers. We then see from the functional equation that $\zeta(s)=0$ whenever $s=-2m$ for m a positive integer. These are called the 'trivial' zeros of $\zeta(s)$. Any other zeros of $\zeta(s)$ must lie in the so called 'critical strip' $0<\mathrm{Re}(s)<1$. Notice the strict inequality on the right. This is important in the proof of the Prime Number Theorem. In particular we have:

Fun Fact 4 (Zero Free Region for $\zeta(s)$)**.** *There exists some absolute constant $c>0$ such that $\zeta(\sigma+it)$ has no zero in the region*

$$\sigma>1-\tfrac{c}{\log t},\quad t\geq 2.$$

Fun Fact 5. *The zeros of $\zeta(s)$ are symmetric about the real axis and the 'critical line'* $\mathrm{Re}(s)=1/2$.

Fun Fact 6. *Let $N(T)$ be the number of zeros of $\zeta(\sigma+it)$ in the critical strip with $0<t\leq T$. Then*

$$N(T)=\frac{T}{2\pi}\log\frac{T}{2\pi}-\frac{T}{2\pi}+O(\log T).$$

We list the next fact as a theorem since it is from this that we will deduce the asymptotic for $\psi(x)$ thereby proving the Prime Number Theorem.

Theorem 2 (Explicit Formula for $\psi(x)$)**.** *Given $x>2$ and not a prime power, Tchebychev's function satisfies the formula*

$$\psi(x)=x-\sum_{\rho}\frac{x^\rho}{\rho}-\frac{\zeta'(0)}{\zeta(0)}-\frac{1}{2}\log(1-x^{-2}),$$

where the sum extends over the 'nontrivial' zeros of $\zeta(s)$. More precisely, we write $\rho=\beta+i\gamma$ for ρ a zero of $\zeta(s)$ in the critical strip and we have

$$\psi(x)=x-\sum_{|\gamma|<T}\frac{x^\rho}{\rho}-\frac{\zeta'(0)}{\zeta(0)}-\frac{1}{2}\log(1-x^{-2})+R(x,T)$$

where, if we let $\langle x\rangle$ be the distance from x to the nearest prime power, we have

$$R(x,T)=O\left(\frac{x\log^2(xT)}{T}+(\log x)\min\left(1,\frac{x}{T\langle x\rangle}\right)\right).$$

To deduce (9) from this we will take T to be a function of x, going to infinity with x but more slowly and such that $R(x,T) = O(xe^{-c\sqrt{\log x}})$. We can also note that $\log(1 - x^{-2}) \to 0$ as $x \to \infty$. Thus, we may write

$$\psi(x) = x - \sum_{|\gamma|<T} \frac{x^\rho}{\rho} + R(x,T) + O(1),$$

making it clear that the main question now in proving the Prime Number Theorem is that of estimating the sum over the nontrivial zeros of $\zeta(s)$. Taking absolute values in this sum gives

$$\sum_{|\gamma|<T} \left| \frac{x^\rho}{\rho} \right| \leq \sum_{|\gamma|<T} \frac{x^\beta}{|\gamma|}.$$

But we know that if $|\gamma| < T$ and $T > 2$ then $\beta < 1 - c/\log T$. This gives

$$\begin{aligned} |x^\rho| = x^\beta &= e^{\beta \log x} \\ &< e^{(1-c/\log T)\log x} \\ &= xe^{-c\log x/\log T}. \end{aligned}$$

Next, by partial summation and since $N(T) = O(T\log T)$, we have

$$\sum_{0<\gamma<T} \frac{1}{\gamma} = \frac{N(T)}{T} + \int_0^T \frac{N(t)\,dt}{t^2} = O(\log^2 T).$$

Putting these together, we have

$$\sum_{|\gamma|<T} \frac{x^\beta}{|\gamma|} = O\left(x(\log^2 T)e^{-c\log x/\log T}\right).$$

We now take T to be a function of x by letting $(\log T)^2 = \log x$ and we get

$$\sum_{|\gamma|<T} \frac{x^\beta}{|\gamma|} = O\left(x(\log x)e^{-c\sqrt{(\log x)}}\right).$$

Finally, we note that for $c' < c$ we have

$$(\log x)e^{-c\sqrt{\log x}} = o\left(e^{-c'\sqrt{\log x}}\right).$$

To see this observe that, if we let $u = \log x$, then

$$\frac{(\log x)e^{-c\sqrt{\log x}}}{e^{-c'\sqrt{\log x}}} = \frac{\log x}{e^{(c-c')\sqrt{\log x}}} = \frac{u}{e^{(c-c')\sqrt{u}}} \to 0.$$

This gives us

$$\sum_{|\gamma|<T} \frac{x^{\beta}}{|\gamma|} = O\left(xe^{-c'\sqrt{(\log x)}}\right).$$

Next, we can assume without loss of generality that x is an integer so that $\langle x \rangle \geq 1$. With this choice we have

$$R(x,T) = O\left(\frac{x \log^2(xT)}{T}\right).$$

Recalling our choice of T we have $T^{-1} = e^{-(\log x)^{1/2}}$ and

$$\log^2(xT) = (\log x)^2 + 2(\log x)^{3/2} + \log x = O(\log^2 x).$$

This gives us

$$R(x,T) = O\left(x(\log^2 x)e^{-(\log x)^{1/2}}\right).$$

Now, if we take $c'' < \min(1, c')$ then we have $(\log^2 x)e^{-(\log x)^{1/2}} = o\left(e^{-c''\sqrt{\log x}}\right)$. In particular, we now have

$$\begin{aligned} \psi(x) &= x + O\left(xe^{-c'\sqrt{(\log x)}} + x(\log^2 x)e^{-(\log x)^{1/2}}\right) \\ &= x + O\left(xe^{-c''\sqrt{\log x}}\right) \end{aligned}$$

with $c'' < 1$, which proves (9) and the Prime Number Theorem follows.

Remark 1. *The Riemann Hypothesis says that all the nontrivial zeros of $\zeta(s)$ lie on the critical line; that is, we may take $\beta = 1/2$. This would give us*

$$\psi(x) = x + O\left(x^{1/2} \log x\right).$$

Then carrying out the proof of Lemma 1 with this new error term, we note that for $x > 3$ we have

$$\int_2^x \frac{t^{1/2} \log t}{t \log^2 t}\, dt \ll \int_2^x \frac{\log t}{t^{1/2}}\, dt \leq (\log x) \int_2^x t^{-1/2}\, dt = O\left(x^{1/2} \log x\right).$$

Thus, assuming the Riemann Hypothesis, we have

$$\pi(x) = \mathrm{li}(x) + O\left(x^{1/2} \log x\right).$$

Remark 2. *In the statement of Theorem 2 we require that x is not a prime power. This restriction can be disregarded and Theorem 2 remains valid if we agree to alter $\psi(x)$ by subtracting $\frac{1}{2}\Lambda(x)$ when x is a prime power. However, there are plenty of large integers which are not prime powers and so we have chosen to avoid the added complication. See Davenport's book for the full discussion.*

3 Proof of the Explicit Formula for $\psi(x)$

We begin the proof of Theorem 2 by observing that

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n\geq 1} \frac{\Lambda(n)}{n^s}. \tag{10}$$

To prove this, recall that we can write

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

for $\mathrm{Re}(s) > 1$. Then, using the Taylor expansion for $\log(1-z)$ when $|z| < 1$, we have

$$\log \zeta(s) = -\sum_p \log(1 - p^{-s}) = \sum_p \sum_{m\geq 1} \frac{p^{-ms}}{m}.$$

The convergence of the Euler product guarantees the absolute convergence of the above series. In particular, this implies that the series converges uniformly to an analytic function on compact subsets of the half plane $\mathrm{Re}(s) > 1$ by Weierstass' theorem. The same theorem tells us that we can therefore take the term by term derivative to get

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \sum_{m\geq 1} \frac{m(\log p)p^{-ms}}{m} = \sum_p \sum_{m\geq 1} \frac{\log p}{p^{ms}} = \sum_{n\geq 1} \frac{\Lambda(n)}{n^s}.$$

Next, we make use of *Perron's Formula*, which states that for $c > 0$ [1] we have

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{y^s}{s}\, ds = \begin{cases} 0 & \text{if } 0 < y < 1, \\ \frac{1}{2} & \text{if } y = 1, \\ 1 & \text{if } y > 1. \end{cases} \tag{11}$$

This formula follows from the following lemma, which can be found on page 105 in Davenport's book.

Lemma 2. *Let $\delta(y)$ denote the function of y on the right of* (11) *and let*

$$I(y,T) = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{y^s}{s}\, ds.$$

Then, for $y > 0, c > 0, T > 0$, we have

$$|I(y,T) - \delta(y)| < \begin{cases} y^c \min\left(1, \frac{1}{T|\log y|}\right) & \text{if } y \neq 1, \\ c/T & \text{if } y = 1. \end{cases} \tag{12}$$

---

[1] Note that this is *not* the same constant c that appears in the error term given by the Prime Number Theorem.

Let us accept this Lemma for now and proceed with the proof of Theorem 2. As $T \to \infty$, the right hand side of (12) goes to 0, which proves Perron's formula. We fix $x > 2$, not a prime power, and let $y = x/n$; then using Perron's formula and the fact that

$$\sum_{n\geq 1} \frac{\Lambda(n)}{n^s}$$

is uniformly convergent on compact subsets of the half plane $\mathrm{Re}(s) > 1$, we have

$$\begin{aligned} \psi(x) &= \sum_{n\leq x} \Lambda(n) = \frac{1}{2\pi i} \sum_{n\geq 1} \Lambda(n) \int_{c-i\infty}^{c+i\infty} \frac{(x/n)^s}{s}\, ds \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \left(\sum_{n\geq 1} \frac{\Lambda(n)}{n^s} \right) \frac{x^s}{s}\, ds \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds. \end{aligned}$$

This is the key observation that allows us to prove the Prime Number Theorem using complex analysis. It is also a trick worth remembering; we've used Perron's formula to write a function of arithmetic interest as a contour integral, thereby turning an arithmetic problem into an analytic one.

The next step is to express this line integral in terms of a simple closed contour which will allow us to apply the residue theorem. We choose $c = 1 + (\log x)^{-1}$ (so that in particular $c > 1$ and $x^c = ex$); U we choose to be a large odd integer and T so that the line segments from $c + iT$ to $-U + iT$ or from $c - iT$ to $-U - iT$ do not pass through any zeros of $\zeta(s)$ in the critical strip (we can do this because the zeros are well spaces as will be made explicit below in Fun Fact 7 and symmetric about the real axis). Let $\mathcal{C}$ be the rectangle with corners at $c - iT$, $c + iT$, $-U + iT$, and $-U - iT$. Then by the residue theorem,

$$\frac{1}{2\pi i} \int_{\mathcal{C}} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds = x - \sum_{|\gamma|<T} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \sum_{0<2m<U} \frac{x^{-2m}}{-2m}. \tag{13}$$

To see this, recall first that the logarithmic derivative of a meromorphic function $f(z)$ has a simple pole at each of the poles and zeros of $f(z)$ with residue equal to the order of the pole or zero. Thus, the x term is from the pole of $\zeta(s)$ at $s = 1$, the first sum gives the residues at the nontrivial zeros of $\zeta(s)$ (taken with multiplicity), the term $\zeta'(0)/\zeta(0)$ from the pole of x^s/s at $s = 0$ and the last sum from the trivial zeros of $\zeta(s)$.

We now wish to relate the integral in (13) to $\psi(x)$. We do this by estimating the error term

$$R(x, T) = \psi(x) - \frac{1}{2\pi i} \int_{\mathcal{C}} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, ds.$$

Define

$$J(x,T)=\frac{1}{2\pi i}\int_{c-iT}^{c+iT}-\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds=\frac{1}{2\pi i}\sum_{n\geq 1}\Lambda(n)\int_{c-iT}^{c+iT}\frac{(x/n)^s}{s}\,ds.$$

Then by Lemma 2 we have

$$\begin{aligned}
|\psi(x)-J(x,T)| &= \left|\sum_{n\geq 1}\Lambda(n)\left(\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{(x/n)^s}{s}\,ds-\frac{1}{2\pi i}\int_{c-iT}^{c+iT}\frac{(x/n)^s}{s}\,ds\right)\right| \\
&= \left(\sum_{n\geq 1}\Lambda(n)\right)\left|\delta(x/n)-I\left(\frac{x}{n},T\right)\right| \\
&< \sum_{n\geq 1}\Lambda(n)(x/n)^c\min\left(1,\frac{1}{T|\log x/n|}\right). \qquad (14)
\end{aligned}$$

We claim that this is

$$O\left(\frac{x(\log x)^2}{T}+(\log x)\min\left(1,\frac{x}{T\langle x\rangle}\right)\right),$$

where $\langle x\rangle$ is the distance from x to the nearest prime power.

To see this, recall that $c=1+(\log x)^{-1}$ and therefore, if we consider the Laurent expansion of $\zeta'(s)/\zeta(s)$, we have

$$-\frac{\zeta(c)}{\zeta(c)}=\frac{1}{c-1}+O(1)=O(\log x).$$

Now, consider those n for which $n\leq\frac{3}{4}x$ or $n\geq\frac{5}{4}x$. We have

$$\left|\log\tfrac{x}{n}\right|\geq\log\tfrac{4}{3}>0\quad\text{or}\quad\left|\log\tfrac{x}{n}\right|\geq\log\tfrac{5}{4}>0$$

and so for these n we can bound $|\log x/n|^{-1}$ from above (say by $\max\left(\left(\log\frac{4}{3}\right)^{-1},\left(\log\frac{5}{4}\right)^{-1}\right)$). The corresponding contribution to (14) is thus

$$\ll\frac{x}{T}\sum_{n\geq 1}\frac{\Lambda(n)}{n^c}=\frac{x}{T}\left(-\frac{\zeta'(c)}{\zeta(c)}\right)\ll\frac{x(\log x)}{T}.$$

Next, we consider the range $\frac{3}{4}x<n<x$. Let x_1 be that largest prime power less than x which we can assume is within the given range since otherwise these terms vanish. For the term $n=x_1$ we have

$$\log\frac{x}{n}=-\log\left(1-\frac{x-x_1}{x}\right)\geq\frac{x-x_1}{x}$$

so that the contribution to (14) is less than or equal to

$$\Lambda(x_1)(x/x_1)^c \min\left(1, \frac{x}{T(x-x_1)}\right) \ll (\log x)\min\left(1, \frac{x}{T(x-x_1)}\right).$$

Here we've used that $\Lambda(x_1) < \log x$ and that $(x/x_1)^c < (4/3)^2$ (provided that $x > e \Rightarrow c < 2$). The terms for which $x_1 < n < x$ contribute nothing, since $\Lambda(n) = 0$ for these. For $\frac{3}{4}x < n < x_1$, write $n = x_1 - v$ where $0 < v < \frac{1}{4}x$ and we have

$$\log\frac{x}{n} > \log\frac{x_1}{n} = -\log\left(1 - \frac{v}{x_1}\right) \geq \frac{v}{x_1}.$$

The corresponding contribution to (14) is then less than

$$\left(\frac{4}{3}\right)^c \sum_{0<v<\frac{x}{4}} \Lambda(x_1 - v)T^{-1}(x_1/v) \ll \frac{x\log x}{T}\sum_{0<v<\frac{x}{4}}\frac{1}{v} \ll \frac{x(\log x)^2}{T}.$$

Here, we've used the fact that

$$\sum_{n\leq x}\frac{1}{n} = \frac{1}{x}\sum_{n\leq x}1 + \int_1^x \frac{1}{t^2}\sum_{n\leq t}1\,dt = \log x + O(1).$$

For the range $x < n < \frac{5}{4}x$ we let x_2 be the least prime power greater than x. Then the terms for which $x < n < x_2$ do not contribute to (14), since as before $\Lambda(n) = 0$ for these. For $n = x_2$ we have

$$\left|\log\frac{x}{n}\right| = -\log\frac{x}{x_2} = -\log\left(1 - \frac{x_2 - x}{x_2}\right) \geq \frac{x_2 - x}{x_2} > \frac{4}{5}\left(\frac{x_2 - x}{x}\right).$$

Thus, for this term we have a contribution which is

$$O\left((\log x)\min\left(1, \frac{x}{T(x_2 - x)}\right)\right).$$

Finally, if $x_2 < n < \frac{5}{4}x$, write $n = x_2 + w$ where $0 < w < \frac{1}{4}x$ and we have

$$\begin{aligned}\left|\log\frac{x}{n}\right| &= -\log\frac{x}{x_2 + w} = -\log\left(1 - \frac{x_2 + w - x}{x_2 + w}\right)\\ &\geq \frac{x_2 + w - x}{x_2 + w} > \frac{w}{x_2 + w} > \frac{4}{5}\left(\frac{w}{x}\right).\end{aligned}$$

Then, in (14) the contribution is less than

$$\sum_{0<w<\frac{x}{4}} \Lambda(x_2 + w)T^{-1}(x/w) \ll \frac{x\log x}{T}\sum_{0<w<\frac{x}{4}}\frac{1}{w} \ll \frac{x(\log x)^2}{T}.$$

Putting all this together, we've proven our claim that

$$|\psi(x) - J(x,T)| = O\left(\frac{x(\log x)^2}{T} + (\log x)\min\left(1, \frac{x}{T\langle x\rangle}\right)\right),$$

where $\langle x\rangle$ is the distance from x to the nearest prime power.

In estimating $R(x,T)$, it remains to bound the difference

$$\frac{1}{2\pi i}\int_{\mathcal{C}} -\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds - J(x,T).$$

That is we must bound the horizontal components and the remaining vertical component. For this, define

$$\overline{H}(x,T) = \frac{1}{2\pi i}\int_{c+iT}^{-U+iT} -\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds = \frac{1}{2\pi i}\int_{-U}^{c} \frac{\zeta'(\sigma+iT)}{\zeta(\sigma+iT)}\frac{x^{\sigma+iT}}{\sigma+iT}\,d\sigma,$$

and similarly

$$\underline{H}(x,T) = \frac{1}{2\pi i}\int_{-U-iT}^{c-iT} -\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds = \frac{1}{2\pi i}\int_{-U}^{c} -\frac{\zeta'(\sigma-iT)}{\zeta(\sigma-iT)}\frac{x^{\sigma-iT}}{\sigma-iT}\,d\sigma.$$

Also, let

$$K(x,T) = \frac{1}{2\pi i}\int_{-U+iT}^{-U-iT} \frac{-\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds = \frac{1}{2\pi i}\int_{-T}^{T} \frac{\zeta'(-U+it)}{\zeta(-U+it)}\frac{x^{-U+it}}{-U+it}\,dt.$$

To accomplish our task we must invoke some properties of $\zeta(s)$ which we've not yet stated.

Fun Fact 7. *Let $T > 2$ and write $\rho = \beta + i\gamma$ for a zero of $\zeta(s)$ in the critical strip. The number of zeros ρ for which $|\gamma - T| < 1$ is $O(\log T)$ and among these zeros there is a gap of length $\gg (\log T)^{-1}$.*

Fun Fact 8. *Let $s = \sigma + it$ and suppose $-1 \le \sigma \le 2$. Then we have*

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{|\gamma-T|<1}\frac{1}{s-\rho} + O(\log T).$$

Moreover, for $\sigma < -1$ and provided that s is not too near to one of the trivial zeros of $\zeta(s)$ (say outside of a circle of radius $1/2$ about a zero), we have

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| = O(\log 2|s|).$$

With these facts, we can complete the proof of Theorem 2. First, by Fun Fact 7, we can choose T so that $|\gamma - T| \gg (\log T)^{-1}$ (we vary our initial choice T_0 by some amount < 1 so that we cut right through the middle of the largest gap in the range $|\gamma - T_0| < 1$). Then applying this fact along with Fun Fact 8, we have

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{|\gamma - T| < 1} \frac{1}{s-\rho} + O(\log T) = O(\log^2 T)$$

whenever $-1 \le \sigma \le 2$. For each term in the above sum contributes $\ll \log T$ (since $|s-\rho| = |(\sigma-\beta) + i(T-\gamma)| \ge |T-\gamma| \gg (\log T)^{-1}$) and there are $\ll \log T$ terms. Thus, in this range we have

$$\begin{aligned}
\left| \int_{-1}^{c} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, d\sigma \right| &\ll \log^2 T \int_{-1}^{c} \left| \frac{x^s}{s} \right| d\sigma = \log^2 T \int_{-1}^{c} \frac{x^\sigma}{\sqrt{\sigma^2 + T^2}}\, d\sigma \\
&\le \frac{\log^2 T}{T} \int_{-\infty}^{c} x^\sigma\, d\sigma = \frac{\log^2 T}{T} \left[\frac{x^\sigma}{\log x} \right]_{-\infty}^{c}. \\
&\ll \frac{x \log^2 T}{T \log x}. \qquad (15)
\end{aligned}$$

When, $-U \le \sigma < -1$, we apply Fun Fact 8 to find

$$\left| \frac{1}{2\pi i} \int_{-U}^{-1} -\frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s}\, d\sigma \right| \ll \frac{\log 2T}{T} \int_{-U}^{-1} x^\sigma\, d\sigma \ll \frac{\log T}{Tx \log x}.$$

(Here we've observed that $|s| = \sqrt{\sigma^2 + T^2} = T\sqrt{(\sigma/T)^2 + 1} \le T\sqrt{(U/T)^2 + 1} < T\sqrt{U^2+1} \ll T$.)This bound is negligible compared with the bound in (15). Putting these together gives us

$$\overline{H}(x,T) = O\left(\frac{x \log^2 T}{T \log x} \right),$$

and similarly for $\underline{H}(x,T)$. Finally, applying Fun Fact 8 to $K(x,T)$ gives us

$$K(x,T) \ll \log 2U \int_{-T}^{T} \frac{x^{-U}}{\sqrt{U^2 + t^2}}\, dt \le \frac{\log 2U}{Ux^U} \int_{-T}^{T} dt \ll \frac{T \log U}{Ux^U}.$$

This error tends to 0 as $U \to \infty$ and so in particular we have

$$\begin{aligned}
R(x,T) &= O\left(\frac{x \log^2 T}{T \log x} + \frac{x(\log x)^2}{T} + (\log x) \min\left(1, \frac{x}{T\langle x \rangle}\right) \right) \\
&= O\left(\frac{x(\log xT)^2}{T} + (\log x) \min\left(1, \frac{x}{T\langle x \rangle}\right) \right).
\end{aligned}$$

This completes the proof of Theorem 2, minus the following technicality.

Proof of Lemma 2. First, let $0 < y < 1$ and note that here $|I(y,T) - \delta(y)| = |I(y,T)|$. We consider the rectangle R with vertices $c - iT$, $c + iT$, $X + iT$, and $X - iT$ where $X > c$ and observe that

$$-I(y,T) = \frac{1}{2\pi i}\left(\int_R \frac{y^s}{s}\,ds - \int_{c-iT}^{X-iT} \frac{y^s}{s}\,ds - \int_{X-iT}^{X+iT} \frac{y^s}{s}\,ds - \int_{X+iT}^{c+iT} \frac{y^s}{s}\,ds\right).$$

But $\int_R y^s/s\,ds = 0$ by the residue theorem and

$$\left|\frac{1}{2\pi i}\int_{X-iT}^{X+iT} \frac{y^s}{s}\,ds\right| \le \frac{Ty^X}{\pi X} \to 0 \text{ as } X \to \infty$$

since here $|s| = |X + iT| > X$. This gives

$$I(y,T) = \frac{1}{2\pi i}\left(\int_{c-iT}^{\infty-iT} \frac{y^s}{s}\,ds - \int_{c+iT}^{\infty+iT} \frac{y^s}{s}\,ds\right)$$

Then, since $|s| = |\sigma + iT| > T$ we have

$$\left|\int_{c-iT}^{\infty-iT} \frac{y^s}{s}\,ds\right| \le \int_{c-iT}^{\infty-iT} \left|\frac{y^s}{s}\right|\,ds < \frac{1}{T}\int_c^\infty y^\sigma\,d\sigma = \frac{y^c}{T|\log y|}.$$

The same computation gives

$$\left|\int_{c+iT}^{\infty+iT} \frac{y^s}{s}\,ds\right| < \frac{y^c}{T|\log y|}$$

and so

$$|I(y,T)| \le \frac{1}{2\pi}\left(\left|\int_{c-iT}^{\infty-iT} \frac{y^s}{s}\,ds\right| + \left|\int_{c+iT}^{\infty+iT} \frac{y^s}{s}\,ds\right|\right) < \frac{1}{\pi}\frac{y^c}{T|\log y|} < \frac{y^c}{T|\log y|}.$$

This gives one inequality in the case $0 < y < 1$. For the other, consider the semicircular arc

$$\mathcal{C} = \{z \in \mathbb{C} : |z| = \sqrt{c^2 + T^2}, \mathrm{Re}(z) \ge c\}$$

taken in the counter clockwise direction. By the residue theorem we have

$$\left|I(y,T) - \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{y^s}{s}\,ds\right| = 0.$$

Letting $R = \sqrt{c^2 + T^2}$, this gives

$$|I(y,T) = \frac{1}{2\pi}\left|\int_{\mathcal{C}} \frac{y^s}{s}\,ds\right| \le \frac{1}{2\pi R}\int_{\mathcal{C}} |y^s|\,ds \le \frac{\pi R y^c}{2\pi R} < y^c,$$

since $0 < y < 1$ and $\mathrm{Re}(s) \geq c$. This finishes the case $0 < y < 1$.

Next, let $y > 1$ and consider a different rectangle R, this time with vertices $c + iT$, $-X + iT$, $-X - iT$, and $c - iT$ where $X > 0$. We observe that

$$\frac{1}{2\pi i} \int_R \frac{y^s}{s}\, ds = 1,$$

by the residue theorem and so

$$I(y,T) = 1 - \frac{1}{2\pi i}\left(\int_{c+iT}^{-X+iT} \frac{y^s}{s}\, ds + \int_{-X+iT}^{-X-iT} \frac{y^s}{s}\, ds + \int_{-X-iT}^{c-iT} \frac{y^s}{s}\, ds\right).$$

Similar to the calculation above, we have

$$\left|\frac{1}{2\pi i} \int_{-X+iT}^{-X-iT} \frac{y^s}{s}\, ds\right| \leq \frac{Ty^{-X}}{\pi X} \to 0 \text{ as } X \to \infty.$$

Thus,

$$|I(y,T) - \delta(y)| = \frac{1}{2\pi}\left|\int_{c+iT}^{-\infty+iT} \frac{y^s}{s}\, ds + \int_{-\infty-iT}^{c-iT} \frac{y^s}{s}\, ds\right|.$$

Then, as before, we have

$$\left|\int_{c+iT}^{-\infty+iT} \frac{y^s}{s}\, ds\right| < \frac{1}{T}\int_{-\infty}^{c} y^\sigma\, d\sigma = \frac{y^c}{T\log y}$$

and similarly for the other integral. Again as before, we arrive at

$$|I(y,T) - \delta(y)| < \frac{y^c}{T\log y},$$

which gives one inequality in the case $y > 1$.

For the other, we take

$$\mathcal{C} = \{z \in \mathbb{C} : |z| = \sqrt{c^2 + T^2}, \mathrm{Re}(z) \leq c\}$$

to get

$$I(y,T) = 1 - \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{y^s}{s}\, ds.$$

Then, letting $R = \sqrt{c^2 + T^2}$ we have

$$|I(y,T) - \delta(y)| \leq \frac{1}{2\pi R}\int_{\mathcal{C}} |y^s|\, ds \leq \frac{y^c}{2} < y^c,$$

since $y > 1$ and $\mathrm{Re}(s) \leq c$. This finishes the case $y > 1$.

We are left with the case $y = 1$. In this case, we write

$$\frac{1}{s} = \frac{1}{c+it} = \frac{c-it}{(c-it)(c+it)} = \frac{c-it}{c^2+t^2},$$

giving us

$$\begin{aligned}
I(1,T) &= \frac{1}{2\pi i}\int_{c-iT}^{c+iT} \frac{1}{s}\,ds = \frac{1}{2\pi i}\int_{-T}^{T} \frac{c-it}{c^2+t^2}\,idt \\
&= \frac{1}{2\pi}\left(\int_{-T}^{0} \frac{c-it}{c^2+t^2}\,dt + \int_{0}^{T} \frac{c-it}{c^2+t^2}\,dt\right) \\
&= \frac{1}{2\pi}\left(-\int_{0}^{-T} \frac{c-it}{c^2+t^2}\,dt + \int_{0}^{T} \frac{c-it}{c^2+t^2}\,dt\right) \\
&= \frac{1}{2\pi}\left(\int_{0}^{T} \frac{c+it}{c^2+t^2}\,dt + \int_{0}^{T} \frac{c-it}{c^2+t^2}\,dt\right) \\
&= \frac{1}{\pi}\int_{0}^{T} \frac{c}{c^2+t^2}\,dt \\
&= \frac{1}{\pi}\int_{0}^{T} \frac{dt/c}{1+(t/c)^2} \\
&= \frac{1}{\pi}\int_{0}^{T/c} \frac{du}{1+u^2} \\
&= \frac{1}{\pi}\left(\int_{0}^{\infty} \frac{du}{1+u^2} - \int_{T/c}^{\infty} \frac{du}{1+u^2}\right).
\end{aligned}$$

Now, $\int 1/(1+u^2)\,du = \arctan u$ plus a constant and $\arctan u \to \pi/2$ as $u \to \infty$. Also,

$$\int_{T/c}^{\infty} \frac{du}{1+u^2} < \int_{T/c}^{\infty} \frac{du}{u^2} = \frac{c}{T}.$$

Thus,

$$I(1,T) = \frac{1}{2} - \frac{1}{\pi}\int_{T/c}^{\infty} \frac{du}{1+u^2} < \frac{1}{2} + \frac{c}{T}$$

and this proves the desired inequality for the case $y = 1$. □