TWO PROOFS OF THE PRIME NUMBER THEOREM

PO-LAM YUNG

1. Introduction

Let $\pi(x)$ be the number of primes $\leq x$. The famous prime number theorem asserts the following:

Theorem 1 (Prime number theorem)**.**

$$\pi(x) \sim \frac{x}{\log x} \tag{1}$$

as $x \to +\infty$. *(This means* $\lim_{x\to+\infty}(\pi(x)\log x)/x = 1$*).*

It has been known since Euclid that there are infinitely many primes. Euler gave an alternative proof of the infinitude of primes based on the divergence of $\sum 1/p$. But it seems that Gauss and Legendre were the first to consider distributions of primes. By studying large tables of primes (primes up to millions!), Gauss noted that the density of primes is approximately $1/\log x$. Chebyshev made some important progress in the 1850's. The landmark paper of Riemann [10] (and his only one on this subject) made clear the connection of the asymptotics of $\pi(x)$ to the ζ function that now bears his name. Subsequently, in 1896, the first complete proof of the prime number theorem was given independently by Hadamard [5] and de la Vallée Poussin [2]. About 50 years later, elementary approaches to the prime number theorem were also discovered, most notably by Erdös [3] and Selberg [8].

Our goal in this article is to elucidate a complex analytic proof of the prime number theorem, given in Chapter 7 of [9]. We will also give a variant of that proof based on the work of D. J. Newman [6] (but proceeds via Chebyshev's ψ function instead of φ; see also the exposition in [12] for another account of Newman's proof.)

Before we begin, we note here that

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log t}\,dt$$

also satisfies $\mathrm{Li}(x) \sim x/\log x$. So the prime number theorem can also be written as

$$\pi(x) \sim \mathrm{Li}(x)$$

as $x \to +\infty$. Indeed, $\mathrm{Li}(x)$ satisfies the following asymptotics: for any $N > 0$, one has

$$\mathrm{Li}(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + 2!\frac{x}{(\log x)^3} + \cdots + (N-1)!\frac{x}{(\log x)^N} + O\left(\frac{x}{(\log x)^{N+1}}\right).$$

$\mathrm{Li}(x)$ turns out to be a better approximation of $\pi(x)$ than $x/\log x$.

Date: October 26, 2016.

2. Chebyshev's ψ function

The proofs of the prime number theorem we will give proceeds via Chebyshev's ψ function:

$$\psi(x) := \sum_{p \leq x} \left[\frac{\log x}{\log p}\right] \log p, \qquad x > 0.$$

The following proposition is well known:[1]

Proposition 2. *The prime number theorem is equivalent to the assertion that*

$$\psi(x) \sim x. \tag{2}$$

Proof. Indeed, assume for the moment that (2) holds. Then since

$$\psi(x) \leq \sum_{p \leq x} \log x = \pi(x) \log x,$$

dividing both sides by x and letting $x \to +\infty$, we get

$$1 \leq \liminf_{x \to \infty} \frac{\pi(x) \log(x)}{x}.$$

Also, for any $\alpha \in (0,1)$, we have

$$\psi(x) \geq \sum_{p \leq x} \log p \geq \sum_{x^\alpha < p \leq x} \log p \geq (\pi(x) - \pi(x^\alpha)) \log(x^\alpha) \geq \alpha(\pi(x) - x^\alpha) \log x.$$

Hence if (2) holds, then dividing the above inequality by x, and letting $x \to \infty$, we get that

$$1 \geq \alpha \limsup_{x \to +\infty} \frac{(\pi(x) - x^\alpha) \log x}{x} = \alpha \limsup_{x \to +\infty} \frac{\pi(x) \log x}{x}.$$

Letting $\alpha \to 1^-$, we get

$$1 \geq \limsup_{x \to +\infty} \frac{\pi(x) \log x}{x}.$$

Together we obtain (1), and the prime number theorem holds.

The converse implication, namely that (1) implies (2), is not much harder. Since we do not need this direction of the implication, we leave this verification to the interested reader. $\square$

Note that ψ can be rewritten as

$$\psi(x) = \sum_{p} \sum_{m \in \mathbb{N}\colon p^m \leq x} \log p = \sum_{n \leq x} \Lambda(n)$$

where Λ is the von Mangoldt function, defined for $n \in \mathbb{N}$ by

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^m \text{ for some prime } p \text{ and some positive integer } m \\ 0, & \text{otherwise} \end{cases}.$$

We are interested in the asymptotics of $\psi(x)$ as $x \to +\infty$. We will see, in the next section, that one can study this by considering the Dirichlet series corresponding to $\{\Lambda(n)\}_{n=1}^{\infty}$, namely

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

[1] See also Proposition 2.1 of Chapter 7 of [9].

This is in turn basically the logarithmic derivative of the Riemann zeta function, and it is ultimately why the Riemann zeta function makes its appearance in this approach towards the prime number theorem.

3. The Mellin transform

We digress a little to discuss three important integral transforms: the Mellin transform, the Laplace transform and the Fourier transform.

Suppose $f\colon (0,\infty) \to \mathbb{C}$ is a measurable function that vanishes on $(0,1)$. Suppose further that there exists some $a \in \mathbb{R}$, $A > 0$ such that

$$|f(x)| \leq Ax^a$$

for all $x \in [1,\infty)$. Let a_0 be the infimum of all $a \in \mathbb{R}$, for which there exists $A > 0$ such that the above estimate holds. Then the Mellin transform of f is defined by

$$\mathcal{M}f(s) = \int_0^\infty f(x)x^{-s}\frac{dx}{x}$$

for all $s \in \mathbb{C}$ with $\operatorname{Re} s > a_0$; indeed the integral defining $\mathcal{M}f(s)$ converges absolutely there, and defines a holomorphic function of s in that half plane.

The Mellin transform is really the Laplace transform in disguise. Indeed, suppose $F\colon \mathbb{R} \to \mathbb{C}$ is a measurable function that vanishes on $(-\infty,0)$. Suppose further that there exists some $a \in \mathbb{R}$, $A > 0$ such that

$$|F(t)| \leq Ae^{at}$$

for all $t \in [0,\infty)$. Let a_0 be the infimum of all $a \in \mathbb{R}$, for which there exists $A > 0$ such that the above estimate holds. Then the Laplace transform of F is defined by

$$\mathcal{L}F(s) = \int_{-\infty}^\infty F(t)e^{-st}dt$$

for all $s \in \mathbb{C}$ with $\operatorname{Re} s > a_0$; indeed the integral defining $\mathcal{L}F(s)$ converges absolutely there, and defines a holomorphic function of s in that half plane.

If $F(t) := f(e^t)$ where f is as in the above definition of the Mellin transform, then $\mathcal{M}f(s) = \mathcal{L}F(s)$ whenever they are defined.

Recall also the Fourier transform on $\mathbb{R}$. If $G \in L^1(\mathbb{R})$, then its Fourier transform is defined by

$$\widehat{G}(\tau) = \int_{-\infty}^\infty G(t)e^{-it\tau}dt$$

for all $\tau \in \mathbb{R}$. If F is as in the above definition of the Laplace transform, then for all $c > a_0$ and all $\tau \in \mathbb{R}$, we have $\mathcal{L}F(c+i\tau) = \widehat{F_c}(\tau)$ where $F_c(t) := F(t)e^{-ct}$.

Our goal was to understand asymptotics of $\psi(x) = \sum_{n\leq x}\Lambda(n)$ as $x \to +\infty$. The strategy we will follow is indeed fairly general, and the initial steps works perfectly well when the sequence $\{\Lambda(n)\}_{n=1}^\infty$ is replaced by any sequence of complex numbers $\{a_n\}_{n=1}^\infty$, as long as $|a_n| = n^{o(1)}$ as $n \to \infty$. We phrase it in the following proposition:

Proposition 3. *Suppose $\{a_n\}_{n=1}^{\infty}$ is a sequence of complex numbers satisfying $|a_n| = n^{o(1)}$ as $n \to \infty$. Let $f\colon (0,\infty) \to \mathbb{C}$ be defined by $f(x) = \sum_{n \le x} a_n$. Then the Mellin transform of f is defined for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$, and is given by*

$$\mathcal{M}f(s) = \frac{1}{s}\mathcal{D}_a(s)$$

for all such s, where

$$\mathcal{D}_a(s) := \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

is the Dirichlet series corresponding to the sequence $\{a_n\}_{n=1}^{\infty}$ (also defined for $\operatorname{Re} s > 1$). In particular, the Mellin transform of ψ is defined for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$, and is given for all such s by

$$\mathcal{M}\psi(s) = \frac{1}{s}\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

Proof. Suppose $|a_n| = n^{o(1)}$ as $n \to \infty$, and $f(x) = \sum_{n \le x} a_n$. Then for any $a > 1$, there exists $A > 0$ such that $f(x) \le Ax^a$. Furthermore, we can rewrite $f(x)$, as

$$f(x) = \sum_{n=1}^{\infty} a_n \chi_{[n,\infty)}(x)$$

where $\chi_{[n,\infty)}$ is the characteristic function of the interval $[n,\infty)$. Hence the Mellin transform of f is defined for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$, and is given by

$$\mathcal{M}f(s) = \sum_{n=1}^{\infty} a_n \int_n^{\infty} x^{-s} \frac{dx}{x} = \frac{1}{s}\sum_{n=1}^{\infty} \frac{a_n}{n^s} = \frac{1}{s}\mathcal{D}_a(s)$$

for all such s. (The interchange of the sum with the integral can be justified using Fubini's theorem.) Since $\Lambda(n) = n^{o(1)}$ as $n \to \infty$, applying the result to $a_n = \Lambda(n)$ yields the desired conclusion for $\mathcal{M}\psi(s)$. □

The proposition suggests that in order to understand $\psi(x) = \sum_{n \le x} \Lambda(n)$ (or more generally $f(x) = \sum_{n \le x} a_n$ where $\{a_n\}$ is as in the proposition), it may be helpful to study the corresponding Dirichlet series $\mathcal{D}_\Lambda(s)$ (or $\mathcal{D}_a(s)$); indeed, if we can invert the Mellin transform, then we can hope to convert information about the Dirichlet series $\mathcal{D}_\Lambda(s)$ (or $\mathcal{D}_a(s)$) into information about $\psi(x)$ (or $f(x)$). The success of this approach ultimately lies with our ability to invert the Mellin transform; we study the latter, by studying how one could invert the Fourier and the Laplace transforms.

First, recall that the Fourier transform can be inverted by the following formula under suitable hypothesis on G. For instance, if both G and $\widehat{G}$ are in $L^1(\mathbb{R})$, then

$$G(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{G}(\tau)e^{it\tau}d\tau$$

for all $t \in \mathbb{R}$.

We will need a slightly different form of the Fourier inversion formula, when $\widehat{G}$ is not necessarily integrable:

Proposition 4. *Suppose $G \in L^1(\mathbb{R})$, and $t_0 \in \mathbb{R}$ is a point where the following limits all exist:*

$$G(t_0^+) := \lim_{t \to t_0^+} G(t), \qquad G(t_0^-) := \lim_{t \to t_0^-} G(t),$$

$$G'(t_0^+) := \lim_{t \to t_0^+} \frac{G(t) - G(t_0^+)}{t - t_0}, \qquad G'(t_0^-) := \lim_{t \to t_0^-} \frac{G(t) - G(t_0^-)}{t - t_0}.$$

Then

$$\lim_{T \to +\infty} \frac{1}{2\pi} \int_{-T}^{T} \widehat{G}(\tau) e^{it_0\tau} d\tau \quad \text{exists, and equals} \quad \frac{G(t_0^+) + G(t_0^-)}{2}.$$

In particular, if $G \in L^1(\mathbb{R})$ is piecewise C^1 (meaning that there exists a strictly increasing sequence $\{t_n\}_{n=-\infty}^{\infty}$ such that for all $n \in \mathbb{Z}$, G is differentiable on (t_n, t_{n+1}), and both $G(t)$ and $G'(t)$ has a limit as $t \to t_n^+$ and $t \to t_{n+1}^-$), then

$$\lim_{T \to +\infty} \frac{1}{2\pi} \int_{-T}^{T} \widehat{G}(\tau) e^{it\tau} d\tau = \frac{G(t^+) + G(t^-)}{2}$$

for all $t \in \mathbb{R}$.

The proof uses the famous lemma of Riemann-Lebesgue:

Lemma 5 (Riemann-Lebesgue)**.** *If $H \in L^1(\mathbb{R})$, then $\widehat{H}(\tau) \to 0$ as $\tau \to \pm\infty$.*

In particular, if $H \in L^1(\mathbb{R})$, then

$$\int_{\mathbb{R}} H(t) \sin(tT) dt = \frac{\widehat{H}(-T) - \widehat{H}(T)}{2i} \to 0 \tag{3}$$

as $T \to +\infty$.

Proof of Lemma 5. If h is a smooth function with compact support on $\mathbb{R}$, then $\widehat{h}$ is rapidly decreasing at infinity; indeed

$$\widehat{h}(\tau) = \int_{\mathbb{R}} h(t) e^{-it\tau} dt = \frac{1}{(i\tau)^N} \int_{\mathbb{R}} h(t) \left(-\frac{d}{dt}\right)^N e^{-it\tau} dt = \frac{1}{(i\tau)^N} \int_{\mathbb{R}} \frac{d^N h}{dt^N} e^{-it\tau} dt$$

for all $N \in \mathbb{N}$, so

$$|\widehat{h}(\tau)| \leq C_N |\tau|^{-N}$$

for all $N \in \mathbb{N}$, where $C_N := \left\| \frac{d^N h}{dt^N} \right\|_{L^1(\mathbb{R})}$. In particular then $\widehat{h}(\tau) \to 0$ as $\tau \to \pm\infty$. Now if $H \in L^1(\mathbb{R})$, then we approximate H by a smooth function with compact support; indeed for any $\varepsilon > 0$, there exists a smooth function h with compact support on $\mathbb{R}$, such that

$$\|H - h\|_{L^1(\mathbb{R})} \leq \varepsilon.$$

It follows that

$$|\widehat{H}(\tau)| \leq |\widehat{H - h}(\tau)| + |\widehat{h}(\tau)| \leq \|H - h\|_{L^1(\mathbb{R})} + |\widehat{h}(\tau)| \leq \varepsilon + |\widehat{h}(\tau)|.$$

Since $\widehat{h}(\tau) \to 0$ as $\tau \to \pm\infty$, we conclude that

$$\limsup_{\tau \to \pm\infty} |\widehat{H}(\tau)| \leq \varepsilon.$$

Since this is true for all $\varepsilon > 0$, we conclude that $\widehat{H}(\tau) \to 0$ as $\tau \to \pm\infty$, as desired. $\square$

Proof of Proposition 4. Suppose t_0 is as in the proposition. For any $T > 0$, we have

$$\begin{aligned}\frac{1}{2\pi}\int_{-T}^{T}\widehat{G}(\tau)e^{it_0\tau}d\tau =&\frac{1}{2\pi}\int_{-T}^{T}\left(\int_{\mathbb{R}}G(t)e^{-it\tau}dt\right)e^{it_0\tau}d\tau\\ =&\frac{1}{2\pi}\int_{\mathbb{R}}G(t)\int_{-T}^{T}e^{i(t_0-t)\tau}d\tau dt\\ =&\frac{1}{\pi}\int_{\mathbb{R}}G(t)\frac{\sin((t_0-t)T)}{t_0-t}dt\\ =&\frac{1}{\pi}\int_{\mathbb{R}}G\left(t_0-t\right)\frac{\sin(tT)}{t}dt\end{aligned}$$

(The interchange of the integrals in the second equality is justified by Fubini's theorem since $G \in L^1(\mathbb{R})$.) We also recall the well-known fact that

$$\int_0^\infty \frac{\sin t}{t}dt = \frac{\pi}{2}. \tag{4}$$

(This can be proved, for instance, using contour integrals.) Thus

$$G(t_0^+) = \frac{1}{\pi}\int_{-\infty}^{0}G(t_0^+)\frac{\sin(tT)}{t}dt,$$

and

$$G(t_0^-) = \frac{1}{\pi}\int_{0}^{\infty}G(t_0^-)\frac{\sin(tT)}{t}dt.$$

It follows that

$$\begin{aligned}&\frac{1}{2\pi}\int_{-T}^{T}\widehat{G}(\tau)e^{it_0\tau}d\tau-\frac{G(t_0^+)+G(t_0^-)}{2}\\ =&\frac{1}{\pi}\int_{-\infty}^{0}\frac{G\left(t_0-t\right)-G(t_0^+)}{t}\sin(tT)dt+\frac{1}{\pi}\int_{0}^{\infty}\frac{G\left(t_0-t\right)-G(t_0^-)}{t}\sin(tT)dt.\end{aligned}$$

Now since $G'(t_0^+)$ and $G'(t_0^-)$ both exist, there exists some $\delta > 0$ such that $[G(t_0-t)-G(t_0^+)]/t$ is bounded for $t \in (-\delta, 0)$, and $[G(t_0-t)-G(t_0^-)]/t$ is bounded for $t \in (0, \delta)$. We define

$$H(t) := \begin{cases}[G\left(t_0-t\right)-G(t_0^+)]/t & \text{if } t\in(-\delta,0)\\ [G\left(t_0-t\right)-G(t_0^-)]/t & \text{if } t\in(0,\delta)\\ G(t_0-t)/t & \text{if } |t|\geq\delta\end{cases}.$$

Then $H \in L^1(\mathbb{R})$ (because $H(t)$ is bounded when $0 < |t| < \delta$, and $|H(t)| \leq \delta^{-1}|G(t_0-t)|$ when $|t| \geq \delta$), and

$$\begin{aligned}&\frac{1}{2\pi}\int_{-T}^{T}\widehat{G}(\tau)e^{it_0\tau}d\tau-\frac{G(t_0^+)+G(t_0^-)}{2}\\ =&\frac{1}{\pi}\int_{\mathbb{R}}H(t)\sin(tT)dt+\frac{G(t_0^+)}{\pi}\int_{-\infty}^{-\delta}\frac{\sin(tT)}{t}dt+\frac{G(t_0^-)}{\pi}\int_{\delta}^{\infty}\frac{\sin(tT)}{t}dt.\end{aligned}$$

As $T \to +\infty$, the first term on the right hand side tends to zero by the lemma of Riemann-Lebesgue (see (3)). The second and the third terms tend to zero as well, since

$$\int_{-\infty}^{-\delta}\frac{\sin(tT)}{t}dt = \int_{\delta}^{\infty}\frac{\sin(tT)}{t}dt = \int_{\delta T}^{\infty}\frac{\sin(t)}{t}dt \to 0$$

as $T \to +\infty$ (see (4)). This concludes the proof of Proposition 4. □

In view of the connection of the Laplace transform to the Fourier transform, we obtain the following corollary of Proposition 4:

Proposition 6. *Suppose $F\colon \mathbb{R} \to \mathbb{C}$ is a piecewise C^1 function that vanishes on $(-\infty, 0)$. Suppose further that there exists some $a \in \mathbb{R}$, $A > 0$ such that*

$$|F(t)| \leq Ae^{at}$$

for all $t \in [0, \infty)$. Let a_0 be the infimum of all $a \in \mathbb{R}$, for which there exists $A > 0$ such that the above estimate holds. Then for all $c > a_0$, and all $t \in \mathbb{R}$, we have

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{L}F(s) e^{st} ds = \frac{F(t^+) + F(t^-)}{2}, \tag{5}$$

in the sense that if $\gamma_{c,T}$ is the vertical contour joining $c - iT$ to $c + iT$, then

$$\lim_{T \to +\infty} \frac{1}{2\pi i} \int_{\gamma_{c,T}} \mathcal{L}F(s) e^{st} ds$$

exists, and is equal to $(F(t^+) + F(t^-))/2$.

The integral on the left-hand side of (5) is called the Bromwich integral. The proposition gives a precise set of conditions under which the Bromwich integral inverts the Laplace transform of a function.

Proof of Proposition 6. Suppose F and c are as above. Then $F_c(t) := F(t)e^{-ct}$ is piecewise C^1, and is in $L^1(\mathbb{R})$ since there exists $A > 0$ such that $|F_c(t)| \leq Ae^{-(c-a_0)t/2}$ for all $t \in [0, \infty)$. Now $\mathcal{L}F(c + i\tau) = \widehat{F_c}(\tau)$ for all $\tau \in \mathbb{R}$. Thus Proposition 4 applied to F_c shows that

$$\begin{aligned} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{L}F(s) e^{st} ds &= \frac{1}{2\pi i} \lim_{T \to +\infty} \int_{-T}^{T} \mathcal{L}F(c + i\tau) e^{(c+i\tau)t} i d\tau \\ &= \frac{e^{ct}}{2\pi} \lim_{T \to +\infty} \int_{-T}^{T} \widehat{F_c}(\tau) e^{it\tau} d\tau \\ &= e^{ct} \frac{F_c(t^+) + F_c(t^-)}{2} \\ &= \frac{F(t^+) + F(t^-)}{2}. \end{aligned}$$

□

By a change of variable $x = e^t$, we obtain the following corollary for the inverse of the Mellin transform as well.

Proposition 7. *Suppose $f\colon (0, \infty) \to \mathbb{C}$ is a piecewise C^1 function that vanishes on $(0, 1)$. Suppose further that there exists some $a \in \mathbb{R}$, $A > 0$ such that*

$$|f(x)| \leq Ax^a$$

for all $x \in [1, \infty)$. Let a_0 be the infimum of all $a \in \mathbb{R}$, for which there exists $A > 0$ such that the above estimate holds. Then for all $c > a_0$, and all $x > 0$, we have

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{M}f(s) x^s ds = \frac{f(x^+) + f(x^-)}{2},$$

in the sense that if $\gamma_{c,T}$ is the vertical contour joining $c - iT$ to $c + iT$, then

$$\lim_{T\to+\infty} \frac{1}{2\pi i} \int_{\gamma_{c,T}} \mathcal{M}f(s) x^s ds$$

exists, and is equal to $(f(x^+) + f(x^-))/2$.

Proof. Apply Proposition 6 to $F(t) := f(e^t)$, noting that $\mathcal{M}f(s) = \mathcal{L}F(s)$. □

The following proposition then follows:[2]

Proposition 8. *Suppose $\{a_n\}_{n=1}^{\infty}$ and $\mathcal{D}_a(s)$ are as in Proposition 3. Then for all $c > 1$, and all $x > 0$, we have*

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{D}_a(s) \frac{x^s}{s} ds = \begin{cases} \sum_{n<x} a_n & \text{if } x \notin \mathbb{N} \\ \frac{a_x}{2} + \sum_{n<x} a_n & \text{if } x \in \mathbb{N}. \end{cases} \tag{6}$$

In particular, for all $c > 1$, and all $x > 0$, we have

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} \frac{x^s}{s} ds = \begin{cases} \psi(x) & \text{if } x \notin \mathbb{N} \\ (\psi(x^+) + \psi(x^-))/2 & \text{if } x \in \mathbb{N}. \end{cases} \tag{7}$$

Equation (6) is sometimes known as Perron's formula.

Proof of Proposition 8. To prove (6), it suffices to apply Proposition 7 to $f(x) := \sum_{n\leq x} a_n$, since $f(x)$ is piecewise constant, and Proposition 3 shows that $\mathcal{M}f(s) = \mathcal{D}_a(s)/s$ whenever $\operatorname{Re} s > 1$. Equation (7) then follows from (6) by setting $a_n = \Lambda(n)$. □

4. Relation to the Riemann ζ function

We now continue to prove the prime number theorem. Our strategy was to prove the asymptotics (2) where $\psi(x) := \sum_{n\leq x} \Lambda(n)$. In the previous section, we have expressed ψ in terms of the Dirichlet series of $\{\Lambda(n)\}_{n=1}^{\infty}$, namely

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}.$$

This is intimately connected to the Riemann ζ function, which we will see as follows.

Recall that the Riemann ζ function is defined by

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

[2] We note here that Proposition 7 also provides an alternative proof of Lemma 2.4 in Chapter 7 of [9]. Indeed, if we take f to be such that $f(x) = 1 - (1/x)$ for $x \geq 1$, and $f(x) = 0$ for $x \in (0, 1)$, then f is bounded, continuous, piecewise C^1, and the Mellin transform of f is $\mathcal{M}f(s) = 1/s(s + 1)$ for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 0$. Thus Proposition 7 implies Lemma 2.4 of Chapter 7 of [9]. Similarly, instead of doing a contour integration, one can work out Exercise 6 of Chapter 7 of [9] by interpreting it as an appropriate instance of Proposition 7. We leave the details to the interested readers.

We will assume known the following product factorization of ζ over all primes:

$$\zeta(s) = \prod_p \frac{1}{1-p^{-s}}, \quad \text{valid for } \operatorname{Re} s > 1.$$

Taking logarithmic derivative, we get

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_p \frac{p^{-s}\log p}{1-p^{-s}} = -\sum_p \sum_m \frac{\log p}{p^{ms}},$$

valid for $\operatorname{Re} s > 1$, so in view of the definition of $\Lambda(n)$, we see that

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}, \quad \text{valid for } \operatorname{Re} s > 1. \tag{8}$$

Thus (7) can now be rewritten as

$$-\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{\zeta'(s)}{\zeta(s)} \frac{x^s}{s} ds = \begin{cases} \psi(x) & \text{if } x \notin \mathbb{N} \\ (\psi(x^+) + \psi(x^-))/2 & \text{if } x \in \mathbb{N}. \end{cases} \tag{9}$$

We may thus hope to obtain asymptotics of $\psi(x)$, by studying bounds for ζ' and ζ. Observe that $|x^s| = x^c$ if s is on the contour of integration in the integral in (7) or (9). Thus to show that $\psi(x)$ remains small as $x \to +\infty$, we should shift the contour of integration $\{\operatorname{Re} s = c\}$ to the left as far as possible. A technical point arises here: the integrand in (7) or (9), namely $\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}\frac{x^s}{s}$, is $O(1/|s|)$ only on the contour of integration. So the integrals in (7) or (9) may not converge absolutely, and this is inconvenient when we shift the contour integrals. As a result, in the next section, we show that instead of studying asympotics of $\psi(x)$, it suffices to study the asymptotics of a smoothed out version of $\psi(x)$, that we denote by $\psi_1(x)$. This $\psi_1(x)$ has a Mellin transform that decays more rapidly at infinity, and the analog of (7) or (9) for ψ_1 would converge absolutely, making it easier to deal with.

5. A technical point

Continuing from the last section, let $\psi_1 \colon (0,\infty) \to \mathbb{R}$ be defined by

$$\psi_1(x) = \int_0^x \psi(y)dy$$

for all $x > 0$.

Proposition 9. *$\psi(x) \sim x$ if and only if $\psi_1(x) \sim x^2/2$.*

Thus in view of Proposition 2, to prove the prime number theorem, it suffices to prove that

$$\psi_1(x) \sim \frac{x^2}{2}. \tag{10}$$

Proof of Proposition 9. Indeed, suppose (10) holds. Then for any $\alpha \in (0,1)$, we have

$$\psi_1(x) - \psi_1(\alpha x) = \int_{\alpha x}^x \psi(y)dy \leq (1-\alpha)x\psi(x),$$

so

$$\frac{\psi(x)}{x} \leq \frac{\psi_1(x) - \psi_1(\alpha x)}{(1-\alpha)x^2} = \frac{\frac{\psi_1(x)}{x^2} - \frac{\psi_1(\alpha x)}{(\alpha x)^2}\alpha^2}{1-\alpha}.$$

Letting $x \to +\infty$, we see that

$$\limsup_{x\to+\infty} \frac{\psi(x)}{x} \leq \frac{1}{2} \cdot \frac{1-\alpha^2}{1-\alpha} = \frac{1+\alpha}{2}.$$

Letting $\alpha \to 1^-$, we see that

$$\limsup_{x\to+\infty} \frac{\psi(x)}{x} \leq 1.$$

Similarly, for any $\beta \in (1, \infty)$,

$$\psi_1(\beta x) - \psi_1(x) = \int_x^{\beta x} \psi(y) dy \geq (\beta - 1) x \psi(x),$$

so

$$\frac{\psi(x)}{x} \geq \frac{\psi_1(\beta x) - \psi_1(x)}{(\beta-1)x^2} = \frac{\frac{\psi_1(\beta x)}{(\beta x)^2}\beta^2 - \frac{\psi_1(x)}{x^2}}{\beta - 1}.$$

Letting $x \to +\infty$, we see that

$$\liminf_{x\to+\infty} \frac{\psi(x)}{x} \geq \frac{1}{2} \cdot \frac{\beta^2-1}{\beta-1} = \frac{\beta+1}{2}.$$

Letting $\beta \to 1^+$, we see that

$$\liminf_{x\to+\infty} \frac{\psi(x)}{x} \geq 1.$$

Together we see that $\psi(x) \sim x$, as desired.

The converse implication is similar. Since we do not need this direction of the implication, we leave this verification to the interested reader. □

So from now on, we concentrate on proving asymptotics (10) for $\psi_1(x)$. Note that for any $a > 1$, there exists a constant $A > 0$ such that $\psi_1(x)/x \leq Ax^a$. Thus the Mellin transform of $\psi_1(x)/x$ is defined for all $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$, and is given for all such s by

$$\begin{aligned}
\int_0^\infty \frac{\psi_1(x)}{x} x^{-s} \frac{dx}{x} &= -\frac{1}{s+1} \int_0^\infty \psi_1(x) \frac{d}{dx} x^{-(s+1)} dx \\
&= \frac{1}{s+1} \mathcal{M}\psi(s) \\
&= \frac{1}{s(s+1)} \sum_{n=1}^\infty \frac{\Lambda(n)}{n^s} \\
&= -\frac{1}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)}.
\end{aligned}$$

Now since $\psi(x) = \sum_{n\leq x} \Lambda(n)$ is piecewise constant on $(0, \infty)$, $\psi_1(x) = \int_0^x \psi(y) dy$ is continuous and piecewise linear there. Hence $\psi_1(x)/x$ is continuous and piecewise C^1 on $(0, \infty)$. Proposition 7 then shows that for all $c > 1$ and all $x > 0$, we have[3]

$$\frac{\psi_1(x)}{x} = -\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{1}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} x^s ds.$$

i.e.

$$\psi_1(x) = -\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{1}{s(s+1)} \frac{\zeta'(s)}{\zeta(s)} x^{s+1} ds. \tag{11}$$

[3]This is precisely Proposition 2.3 of Chapter 7 of [9].

We note that the integrand above, namely $\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}x^{s+1}$, is $O(1/|s|^2)$ on the contour of integration, thanks to the appearance of the quadratic factor $s(s+1)$ in the denominator (contrary to the linear factor s in the integrand of (7) or (9)). This makes it easier for us to shift the contour of integration $\{\operatorname{Re} s = c\}$ in a moment. To accomplish the latter, we will need to know that ζ continues meromorphically past the line $\{\operatorname{Re} s = 1\}$. We summarize in the next section the facts we will need about the continuation of ζ.

6. Analytic continuation of ζ

We will assume known that ζ has a meromorphic continuation to the half-space

$$S_\eta := \{s \in \mathbb{C} \colon \operatorname{Re} s > \eta\}$$

for some $\eta < 1$, so that the only singularity of ζ in this strip is a simple pole at $s = 1$. In other words, there exists $\eta < 1$, and a holomorphic function $h(s)$ on S_η, such that

$$\zeta(s) = \frac{h(s)}{s-1} \quad \text{on } S_\eta. \tag{12}$$

We will also assume known the following upper bound for ζ': for any $\varepsilon > 0$, there exists $A > 0$, such that

$$|\zeta'(s)| \leq A|\operatorname{Im} s|^{\varepsilon/2} \tag{13}$$

whenever $s \in \mathbb{C}$ with $\operatorname{Re} s \geq 1$ and $|\operatorname{Im} s| \geq 1$. All these can be proved, for instance, by comparing $\zeta(s) = \sum_{n=1}^\infty n^{-s}$ to the corresponding integral $\int_1^\infty x^{-s} dx = 1/(s-1)$ (and using Cauchy's estimate for the bound on ζ')[4]. Indeed, with more work (for instance by establishing the functional equation of $\xi(s) := \pi^{-s/2}\Gamma(s/2)\zeta(s)$), one can take η all the way to $-\infty$ in the above claims.

7. Non-vanishing of ζ on $\operatorname{Re} s = 1$

Recall that our strategy towards proving the prime number theorem is to establish asymptotics (10) for $\psi_1(x)$. We have represented $\psi_1(x)$ as a contour integral in (11). With the analytic continuation of ζ in the above section in mind, we would like to shift the contour of integration to the left as far as possible (just like what is typically done when one computes the Bromwich integral, in inverting the Laplace transform). This relies on knowing where the zeroes of $\zeta(s)$ are, since every zero of ζ contributes a pole in s of the integrand in (11). We now prove the following theorem.

Theorem 10. *ζ has no zeroes on the vertical line where $\operatorname{Re} s = 1$.*

Proof. We need three observations.

First, note that ζ is real-valued on $\{s \in \mathbb{R} \colon s > 1\}$. Hence

$$\overline{\zeta(s)} = \zeta(\overline{s}) \quad \text{for all } s \in \mathbb{C}.$$

This shows that the (non-real) zeroes of ζ comes in conjugate pairs: if s is a zero of ζ, then so is $\overline{s}$, and ζ vanishes to the same order at both s and $\overline{s}$.

[4]See Proposition 2.5, Corollary 2.6, Proposition 2.7 in Chapter 6 of [9].

Next, recall that if f is a meromorphic function near a point z_0, then the order of of f at z_0 (which is positive if f vanishes there, negative if f has a pole there) can be computed via the residue of the logarithmic derivative of f at z_0. In particular, for any $t \in \mathbb{R}$, the order of ζ at $1+it$ is

$$\mathrm{ord}_{1+it}\zeta = \lim_{\epsilon \to 0} \frac{\epsilon\,\zeta'(1+it+\epsilon)}{\zeta(1+it+\epsilon)}. \tag{14}$$

Finally, recall the logarithmic derivative of ζ, given by (8). What will be important for us is that $\Lambda(n)$ is real and non-negative for all positive integers n.

Now we are ready to put all these together. First, since ζ has a pole at $s = 1$, it cannot have a zero at $s = 1$. Next, suppose ζ is zero at $s = 1 + it$ for some $t \in \mathbb{R}$. We show that this is impossible by considering the orders of ζ at $1 \pm 2it$, $1 \pm it$ and 1: Combining (14) and (8), we see that

$$\begin{aligned}
\mathrm{ord}_{1+2it}\zeta &= -\lim_{\epsilon \to 0^+} \epsilon \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\epsilon}} n^{-2it} \\
\mathrm{ord}_{1+it}\zeta &= -\lim_{\epsilon \to 0^+} \epsilon \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\epsilon}} n^{-it} \\
\mathrm{ord}_{1}\zeta &= -\lim_{\epsilon \to 0^+} \epsilon \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\epsilon}} \\
\mathrm{ord}_{1-it}\zeta &= -\lim_{\epsilon \to 0^+} \epsilon \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\epsilon}} n^{it} \\
\mathrm{ord}_{1-2it}\zeta &= -\lim_{\epsilon \to 0^+} \epsilon \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^{1+\epsilon}} n^{2it}
\end{aligned}$$

We now multiply these five equations by 1, 4, 6, 4, 1 respectively, and add them all up. Observe that

$$n^{-2it} + 4n^{-it} + 6 + 4n^{it} + n^{2it} = (n^{it/2} + n^{-it/2})^4 = (2\cos(t \log n/2))^4 \geq 0. \tag{15}$$

Since $\Lambda(n) \geq 0$ for all n, we then see that

$$\mathrm{ord}_{1+2it}\zeta + 4\,\mathrm{ord}_{1+it}\zeta + 6\,\mathrm{ord}_{1}\zeta + 4\,\mathrm{ord}_{1-it}\zeta + \mathrm{ord}_{1-2it}\zeta \leq 0.$$

But

$$\mathrm{ord}_{1}\zeta = -1,$$

$$\mathrm{ord}_{1+it}\zeta = \mathrm{ord}_{1-it}\zeta,$$

and

$$\mathrm{ord}_{1+2it}\zeta = \mathrm{ord}_{1-2it}\zeta \geq 0.$$

Hence

$$8\,\mathrm{ord}_{1+it}\zeta - 6 \leq 0,$$

which contradicts our assumption that $\zeta(1+it) = 0$.[5] □

[5] We remark that (15) is really Lemma 1.4 of Chapter 7 of [9] in disguise. Also, by rewriting ζ'/ζ as the derivative of $\log \zeta$, and undoing the derivative, the above argument essentially gives Corollary 1.5 of Chapter 7 of [9].

By making the above argument more quantitative, we can establish[6] the following lower bound of ζ: for any $\varepsilon > 0$, there exists a constant $B > 0$, such that

$$|\zeta(s)| \geq B|\mathrm{Im}\, s|^{-\varepsilon/2} \tag{16}$$

whenever $s \in \mathbb{C}$ with $\mathrm{Re}\, s \geq 1$ and $|\mathrm{Im}\, s| \geq 1$. In particular, combining with our earlier bound (13) for ζ', we see that for any $\varepsilon > 0$, there exists a constant $C > 0$, such that

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right| \leq C|\mathrm{Im}\, s|^{\varepsilon} \tag{17}$$

whenever $s \in \mathbb{C}$ with $\mathrm{Re}\, s \geq 1$ and $|\mathrm{Im}\, s| \geq 1$.

8. The Proof of the prime number theorem in [9]

We can now finish the proof of the prime number theorem as in Chapter 7 of [9]. Fix some $c > 1$. Then (11) says that

$$\psi_1(x) = -\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)} x^{s+1} ds,$$

where the integration is along the vertical contour $\{c + i\tau \colon \tau \in \mathbb{R}\}$. The integrand is a holomorphic function of s on an open half-space $\{\mathrm{Re}\, s > \eta\}$ for some $\eta < 1$. In view of estimate (17) for $|\zeta'/\zeta|$, and that

$$\left|\frac{x^{s+1}}{s(s+1)}\right| \leq \frac{|x|^{\mathrm{Re}\, s+1}}{|\mathrm{Im}\, s|^2},$$

we can shift the contour of integration, and obtain, for any $T > 0$, that

$$\psi_1(x) = -\frac{1}{2\pi i}\int_{\gamma(T)} \frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)} x^{s+1} ds,$$

where $\gamma(T)$ is the contour consisting of 5 straight line segments, joining the following points in order: $1 - i\infty$, $1 - iT$, $c - iT$, $c + iT$, $1 + iT$, and $1 + i\infty$. (See p. 195 of [9] for a picture of $\gamma(T)$.)

Suppose now $\varepsilon > 0$ is given. In view of estimate (17) for $|\zeta'/\zeta|$ again, we may choose $T > 0$ large enough, so that

$$\frac{1}{2\pi}\left(\int_{1-i\infty}^{1-iT} + \int_{1+iT}^{1+i\infty}\right)\left|\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}\right| |ds| < \varepsilon.$$

Then since $|x^{s+1}| = x^2$ on the contours of integration in the above two integrals, we get

$$\left|-\frac{1}{2\pi i}\left(\int_{1-i\infty}^{1-iT} + \int_{1+iT}^{1+i\infty}\right) \frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)} x^{s+1} ds\right| < \varepsilon x^2.$$

Hence for this choice of T, we have

$$\psi_1(x) = -\frac{1}{2\pi i}\left(\int_{1-iT}^{c-iT} + \int_{c-iT}^{c+iT} + \int_{c+iT}^{1+iT}\right) \frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)} x^{s+1} ds + O(\varepsilon x^2) \tag{18}$$

for all $x > 0$, where $O(\varepsilon x^2)$ is a term bounded in absolute value by εx^2. Having chosen T, let $\delta \in (0, \min\{1-\eta, 1\})$ be sufficiently small, where η is as in the above, so that ζ has no zeroes in the closed rectangle $\{1 - \delta \leq \mathrm{Re}\, s \leq 1, |\mathrm{Im}\, s| \leq T\}$. Such δ exists because ζ extends meromorphically on S_η as in the description just before (12) (so the zeroes of ζ has no accumulation points in S_η), and because ζ has no zeroes on the line segment $\{\mathrm{Re}\, s = 1, |\mathrm{Im}\, s| \leq T\}$. Let $\gamma(T, \delta)$ be the contour consisting of 5 straight line segments, joining the following points in order: $1 - i\infty$, $1 - iT$, $1 - \delta - iT$,

[6] See Proposition 1.6 in Chapter 7 of [9].

$1-\delta+iT$, $1+iT$, and $1+i\infty$. (See p. 195 of [9] for a picture of $\gamma(T,\delta)$.) Then $\gamma(T)-\gamma(T,\delta)$ is a rectangular contour with vertices $1-\delta-iT$, $c-iT$, $c+iT$, and $1-\delta+iT$. On and inside this rectangular contour, ζ has a pole at $s=1$, and no zeroes anywhere. Thus ζ'/ζ is meromorphic on and inside this rectangular contour, and has a residue -1 at $s=1$. It follows that for each $x>0$, the function $s \mapsto -\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}x^{s+1}$ has a simple pole at $s=1$, and nowhere else inside the rectangle whose vertices are $1-\delta-iT$, $c-iT$, $c+iT$, and $1-\delta+iT$. The residue of this function at $s=1$ is just

$$\frac{1}{1(1+1)}x^{1+1}=\frac{1}{2}x^2.$$

Thus from (18), we see that

$$\psi_1(x)=\frac{1}{2}x^2-\frac{1}{2\pi i}\left(\int_{1-iT}^{1-\delta-iT}+\int_{1-\delta-iT}^{1-\delta+iT}+\int_{1-\delta+iT}^{1+iT}\right)\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}x^{s+1}ds+O(\varepsilon x^2). \tag{19}$$

Now the function $s\mapsto \frac{1}{2\pi}\frac{1}{s(s+1)}\frac{\zeta'(s)}{\zeta(s)}$ is continuous on the above 3 contours of integration. Hence its modulus is bounded above by some constant $C_{T,\delta}$ there. Also,

$$\int_{1-iT}^{1-\delta-iT}|x^{s+1}||ds|=\int_{1-\delta}^{1}x^{\sigma+1}d\sigma\leq\frac{x^2}{\log x};$$

similarly

$$\int_{1-\delta+iT}^{1+iT}|x^{s+1}||ds|\leq\frac{x^2}{\log x}.$$

Furthermore,

$$\int_{1-\delta-iT}^{1-\delta+iT}|x^{s+1}|ds=2Tx^{2-\delta}.$$

All in all, we see that

$$\left|\psi_1(x)-\frac{1}{2}x^2\right|\leq 2C_{T,\delta}\frac{x^2}{\log x}+2Tx^{2-\delta}+\varepsilon x^2.$$

Dividing by $x^2/2$, we see that

$$\left|\frac{\psi_1(x)}{x^2/2}-1\right|\leq 4C_{T,\delta}\frac{1}{\log x}+4Tx^{-\delta}+2\varepsilon.$$

Now T and δ are fixed once we fix ε. If we pick x sufficiently large, the right hand side can be made smaller than 3ε. This proves that $\psi_1(x)\sim x^2/2$, as desired in (10), and completes the proof of the prime number theorem in Chapter 7 of [9].

9. A variant of Newman's proof

The above proof of the prime number theorem is based on analysis of the Chebychev's ψ function:

$$\psi(x)=\sum_p\sum_{m\in\mathbb{N}\colon p^m\leq x}\log p.$$

Some work was needed in obtaining quantitative estimates of ζ near the line $\operatorname{Re}s=1$ (more precisely, an upper bound for $|\zeta'(s)/\zeta(s)|$). On the other hand, in [6], Newman gave another complex analytic proof of the prime number theorem, using only the vanishing of ζ on the line $\operatorname{Re}s=1$ (and no asymptotics of ζ there), by considering the φ function:

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

Below we try to combine the two approaches, and adapt Newman's argument so that it works through Chebychev's ψ function (rather than the φ function).

Recall that by Proposition 2, to prove the prime number theorem, it suffices to verify asymptotics (2) for ψ. It may help to first verify a weaker statement, namely that $\psi(x)/x$ remains bounded as $x \to +\infty$. This is what we are going to do next, via an essentially elementary argument.

Proposition 11. $\dfrac{\psi(x)}{x}$ *remains bounded as* $x \to +\infty$.

Proof. First, we claim that there exists a constant C, such that for any positive integer n, we have

$$\psi(2n) - \psi(n) \leq Cn. \tag{20}$$

To prove this claim, note that

$$\psi(2n) - \psi(n) = \sum_p \sum_{\{m\colon n<p^m\leq 2n\}} \log p = \log\left(\prod_{m=1}^{\infty} \prod_{\{p\colon n<p^m\leq 2n\}} p\right). \tag{21}$$

In the product inside the logarithm, consider first the term corresponding to $m = 1$. We have

$$\prod_{\{p\colon n<p\leq 2n\}} p \leq \binom{2n}{n}. \tag{22}$$

Indeed

$$\binom{2n}{n} = \frac{(2n)(2n-1)\dots(n+1)}{n!}$$

is an integer, so that $n!$ is a factor of $(2n)(2n-1)\dots(n+1)$; also each prime p with $n < p \leq 2n$ is a factor of $(2n)(2n-1)\dots(n+1)$. Since each such prime p is relatively prime with $n!$, we see that $(n!)\prod_{\{p\colon n<p\leq 2n\}} p$ divides $(2n)(2n-1)\dots(n+1)$, i.e. $\prod_{\{p\colon n<p\leq 2n\}} p$ divides $\binom{2n}{n}$. In particular, (22) holds. This further implies

$$\prod_{\{p\colon n<p\leq 2n\}} p \leq 2^{2n}, \tag{23}$$

since

$$\binom{2n}{n} \leq \sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 2^{2n}.$$

This completes our estimate of the product inside the logarithm on the right hand side of (21).

Next, we consider those terms in the same product corresponding to $m \geq 2$. If $m \geq 2$, and $n < p^m \leq 2n$, then $p \leq \sqrt{2n}$. Also, for each prime p, there is at most one power of p that lies in $(n, 2n]$. Hence

$$\prod_{m=2}^{\infty} \prod_{\{p\colon n<p^m\leq 2n\}} p \leq \prod_{p\leq\sqrt{2n}} p \leq (\sqrt{2n})^{\sqrt{2n}}.$$

Hence, together with (23), we obtain

$$\psi(2n) - \psi(n) \leq \log\left(2^{2n}(\sqrt{2n})^{\sqrt{2n}}\right) \leq 2n\log 2 + \frac{\sqrt{2n}\log(2n)}{2}.$$

This establishes our claim (20).

Now that we have the claim (20), we see that there exists a constant C' such that

$$\psi(2x) - \psi(x) \le C'x$$

for all $x \ge 1$. In fact it suffices to prove this for x large. To do so, take n to be the integer closest to x. Then the sum defining $\psi(x)$ and $\psi(n)$ differ in at most one term, and $|\psi(x) - \psi(n)| \le C \log x \le C''x$. Similarly, $|\psi(2x) - \psi(2n)| \le C''x$. Hence together with the bound for $\psi(2n) - \psi(n)$ we already established, we see that $\psi(2x) - \psi(x) \le C'x$, as desired.

Now we just iterate this estimate:

$$\begin{aligned} \psi(x) - \psi(x/2) &\le C'(x/2) \\ \psi(x/2) - \psi(x/4) &\le C'(x/4) \\ &\vdots \end{aligned}$$

and sum up a geometric series on the right. Then

$$\psi(x) \le C'x$$

as $x \to +\infty$, as desired. $\square$

Now recall that the Mellin transform of $\psi(x)$ was given in Proposition 3, which in view of (8) can be written as

$$\int_1^\infty \frac{\psi(x)}{x} \frac{dx}{x^s} = -\frac{\zeta'(s)}{s\zeta(s)}, \quad \text{valid for } \operatorname{Re} s > 1. \tag{24}$$

We are interested in showing $\dfrac{\psi(x)}{x} - 1 \to 0$ as $x \to +\infty$. Hence we are led to consider the following identity:

Proposition 12.

$$\int_1^\infty \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x^s} = -\frac{\zeta'(s)}{s\zeta(s)} - \frac{1}{s-1}, \quad \textit{valid for } \mathit{Re}\, s > 1.$$

Proof. This follows from (24) by simply noting that

$$\int_1^\infty \frac{dx}{x^s} = \frac{1}{s-1}, \quad \text{valid for } \operatorname{Re} s > 1.$$

(This could be interpreted as the Mellin transform of $x\chi_{[1,\infty)}(x)$.) $\square$

From the meromorphic continuation (12) of ζ to some half-space $\{\operatorname{Re} s > \eta\}$ with $\eta < 1$, we see that $\frac{\zeta'(s)}{\zeta(s)} + \frac{1}{s-1}$ extends to a homormorphic function on the same half-space. Hence the right hand side of the identity in Proposition 12 extends to a holomorphic function on an open set containing the closed half plane $\{\operatorname{Re} s \ge 1\}$. This shows that the following Tauberian theorem applies:

Proposition 13. *Let $f(x)$ be a bounded function on $[1, \infty)$, and define*

$$g(s) = \int_1^\infty f(x) \frac{dx}{x^s} \quad \textit{for } \mathit{Re}\, s > 1.$$

Then g is holomorphic on $\mathit{Re}\, s > 1$. If g extends to a holomorphic function on an open set containing the closed half plane $\mathit{Re}\, s \ge 1$, then $\int_1^\infty f(x) \frac{dx}{x}$ exists, and is equal to $g(1)$.

Indeed
$$f(x) := \frac{\psi(x)}{x} - 1$$
is a bounded function by Proposition 11, and the integral
$$\int_1^\infty f(x) \frac{dx}{x^s} = \int_1^\infty \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x^s}$$
extends to a holomorphic function on an open set containing the closed half plane $\{\operatorname{Re} s \geq 1\}$ by Proposition 12. Hence assuming Proposition 13 for the moment, we obtain the following proposition:

Proposition 14. *The improper integral* $\int_1^\infty \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x}$ *converges.*

The prime number theorem would then follow.

Proof of Theorem 1. As observed before, it suffices to verify (2). We argue by contradiction. Suppose (2) is false. Then either there exists $\alpha > 1$ such that $\psi(x_n) > \alpha x_n$ for a sequence $\{x_n\}$ with $x_n \to +\infty$, or there exists $\beta < 1$ such that $\psi(y_n) < \beta y_n$ for a sequence $\{y_n\}$ with $y_n \to +\infty$. In the first case, since ψ is an increasing function, we have $\psi(x) \geq \psi(x_n) \geq \alpha x_n$ whenever $x \geq x_n$. In particular,
$$\int_{x_n}^{\alpha x_n} \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x} \geq \int_{x_n}^{\alpha x_n} \left(\frac{\alpha x_n}{x} - 1 \right) \frac{dx}{x} = \int_1^\alpha \left(\frac{\alpha}{x} - 1 \right) \frac{dx}{x}, \tag{25}$$
the last integral being strictly positive, and independent of n. This contradicts Proposition 14: in fact, Proposition 14 implies that
$$\int_{x_n}^{\alpha x_n} \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x} = \int_1^{\alpha x_n} \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x} - \int_1^{x_n} \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x} \to 0$$
as $n \to \infty$, and this is not compatible with the lower bound we have obtained in (25).

Similarly, in the second case, we use $\psi(x) \leq \psi(y_n) < \beta y_n$ whenever $x \leq y_n$, to conclude that
$$\int_{\beta y_n}^{y_n} \left(\frac{\psi(x)}{x} - 1 \right) \frac{dx}{x} \leq \int_{\beta y_n}^{y_n} \left(\frac{\beta y_n}{x} - 1 \right) \frac{dx}{x} = \int_\beta^1 \left(\frac{\beta}{x} - 1 \right) \frac{dx}{x} < 0$$
independent of n. This contradicts Proposition 14. □

It remains to prove Proposition 13.

Proof of Proposition 13. Suppose f is bounded, say $|f(x)| \leq M$ for all $x \geq 1$. Suppose also that
$$g(s) := \int_1^\infty f(x) \frac{dx}{x^s}$$
extends holomorphically to an open set containing the closed half plane $\{\operatorname{Re} s \geq 1\}$. Let
$$g_t(s) = \int_1^t f(x) \frac{dx}{x^s}.$$
Then g_t is entire for all t, and our goal is to show that $g_t(1)$ converges to $g(1)$ as $t \to +\infty$. For $\varepsilon > 0$ and $\delta > 0$, let $\Lambda_{\varepsilon,\delta}$ be the positively oriented closed contour, given by
$$\Lambda_{\varepsilon,\delta} = C_\varepsilon + L_\delta^{(1)} + L_\delta^{(2)} + L_\delta^{(3)} \tag{26}$$

where

- C_ε be the semicircle in the right half plane $\{\operatorname{Re} s > 1\}$, that is centered at 1 and of radius $1/\varepsilon$;
- $L_\delta^{(1)}$ is the horizontal straight line joining $1 + i\varepsilon^{-1}$ to $1 - \delta + i\varepsilon^{-1}$;
- $L_\delta^{(2)}$ is the vertical straight line joining $1 - \delta + i\varepsilon^{-1}$ to $1 - \delta - i\varepsilon^{-1}$; and
- $L_\delta^{(3)}$ is the horizontal straight line joining $1 - \delta - i\varepsilon^{-1}$ to $1 - i\varepsilon^{-1}$.

Then for any $\varepsilon > 0$, as long as δ is sufficiently small, we have, by Cauchy integral formula, that

$$g_t(1) - g(1) = \frac{1}{2\pi i} \int_{\Lambda_{\varepsilon,\delta}} [g_t(s) - g(s)] \frac{ds}{s-1}.$$

For various technical reasons, we will actually use the following identity instead (which also follow from Cauchy's integral formula, since the extra factor $t^{s-1}(1+\varepsilon^2(s-1)^2)$ is entire in s, and equals 1 when $s = 1$):

$$g_t(1) - g(1) = \frac{1}{2\pi i} \int_{\Lambda_{\varepsilon,\delta}} [g_t(s) - g(s)] t^{s-1}(1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} \tag{27}$$

Now we decompose the above path integral into 4 parts, according to (26). For $s \in C_\varepsilon$, we have

$$\begin{aligned} |g_t(s) - g(s)| &\leq \int_t^\infty |f(x)| \frac{dx}{x^{\operatorname{Re} s}} \leq \frac{M t^{1-\operatorname{Re} s}}{\operatorname{Re} s - 1}, \\ |t^{s-1}| &\leq t^{\operatorname{Re} s - 1} \\ |1 + \varepsilon^2(s-1)^2| &= \frac{|s - (1 - i\varepsilon^{-1})||s - (1 + i\varepsilon^{-1})|}{\varepsilon^{-2}} \leq C \frac{|\operatorname{Re} s - 1|}{\varepsilon^{-1}} \\ \frac{1}{|s-1|} &= \frac{1}{\varepsilon^{-1}}. \end{aligned}$$

Hence

$$\left| \frac{1}{2\pi i} \int_{C_\varepsilon} [g_t(s) - g(s)] t^{s-1}(1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} \right| \leq CM\varepsilon \tag{28}$$

Next, let $\tilde{C}_\varepsilon$ be the semi-circle in the left half plane $\{\operatorname{Re} s < 1\}$, that is centered at 1 and of radius $1/\varepsilon$. Then we integrate the part concerning g in (27), over $\tilde{C}_\varepsilon$ instead of over $L_\delta^{(1)} + L_\delta^{(2)} + L_\delta^{(3)}$. (This is possible because g_t is entire.) Hence

$$\int_{L_\delta^{(1)} + L_\delta^{(2)} + L_\delta^{(3)}} g_t(s) t^{s-1}(1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} = \int_{\tilde{C}_\varepsilon} g_t(s) t^{s-1}(1+\varepsilon^2(s-1)^2) \frac{ds}{s-1}.$$

But on $\tilde{C}_\varepsilon$, we have

$$|g_t(s)| \leq \int_1^t |f(x)| \frac{dx}{x^{\operatorname{Re} s}} \leq \frac{M t^{1-\operatorname{Re} s}}{1 - \operatorname{Re} s}.$$

Similarly as before, on $\tilde{C}_\varepsilon$, we have

$$|t^{s-1}(1+\varepsilon^2(s-1)^2)| \leq C \frac{t^{\operatorname{Re} s - 1}(1 - \operatorname{Re} s)}{\varepsilon^{-1}}.$$

Hence

$$\left| \frac{1}{2\pi i} \int_{L_\delta^{(1)} + L_\delta^{(2)} + L_\delta^{(3)}} g_t(s) t^{s-1}(1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} \right| \leq CM\varepsilon. \tag{29}$$

Finally, the contribution of $g(s)$ to the contour integral over $L_\delta^{(1)} + L_\delta^{(3)}$ is given by

$$\left| \frac{1}{2\pi i} \int_{L_\delta^{(1)}+L_\delta^{(3)}} g(s) t^{s-1} (1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} \right| \leq C\delta\varepsilon \tag{30}$$

where $|g| \leq C$ on $L_\delta^{(1)} \cup L_\delta^{(3)}$. This is because $L_\delta^{(1)}$ and $L_\delta^{(3)}$ both have lengths $\leq \delta$, and that $1/|s-1| \simeq \varepsilon$ on $L_\delta^{(1)} \cup L_\delta^{(3)}$. (Note also that $|t^{s-1}| \leq 1$ since $\operatorname{Re} s < 1$, and $|1+\varepsilon^2(s-1)^2| \leq C$ on $L_\delta^{(1)} \cup L_\delta^{(3)}$.) Now the contribution of $g(s)$ to the contour integral over $L_\delta^{(2)}$ is given by

$$\left| \frac{1}{2\pi i} \int_{L_\delta^{(2)}} g(s) t^{s-1} (1+\varepsilon^2(s-1)^2) \frac{ds}{s-1} \right| \leq C \frac{t^{-\delta}}{\varepsilon\delta}. \tag{31}$$

This is because the length of $L_\delta^{(2)}$ is $2/\varepsilon$, the function $|g|$ is bounded by C on $L_\delta^{(2)}$, and $|t^{s-1}| = t^{-\delta}$ on $L_\delta^{(2)}$; also, $|1+\varepsilon^2(s-1)^2| \leq C$ on $L_\delta^{(2)}$, and $1/|s-1| \leq 1/\delta$ on $L_\delta^{(2)}$.

Altogether, by (27), (28), (29), (30) and (31), we see that for any $\varepsilon > 0$, there exists a small $\delta > 0$, such that

$$|g_t(1) - g(1)| \leq CM\varepsilon + \frac{Ct^{-\delta}}{\varepsilon\delta}.$$

Letting $t \to +\infty$, we see that

$$\limsup_{t\to+\infty} |g_t(1) - g(1)| \leq CM\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows that $g_t(1) \to g(1)$ as $t \to +\infty$, as desired. □

10. Concluding remarks

We end by mentioning some comparisons of the two proofs of the prime number theorem given above.

The Tauberian proof based on the work of Newman is shorter, and does not involve the use of any quantitative estimates of ζ on the line $\{\operatorname{Re} s = 1\}$, whereas the proof given in [9] requires knowing such estimates. Nevertheless, the proof given in [9] works in more general context, when one wants to obtain asymptotics for $\sum_{n\leq x} a_n$ for any appropriate sequences a_n; also, the proof is more powerful, in the sense that if we had known further information about the zeroes of ζ (say we know that ζ has no zeroes on $\{\operatorname{Re} s > \eta\}$ for some particular $\eta > 0$), then we can use that to our advantage, and obtain lower order correction terms to the asymptotics of $\pi(x)$. (This would be hard to do with the Tauberian argument of Newman.) Hence it is useful to know the meromorphic continuation of ζ to a region in the complex plane that is as large as possible, and to understand its zeroes there. Indeed, one of the famous "explicit formulas" in the theory of primes says

$$\psi_1(x) = \frac{x^2}{2} - \sum_\rho \frac{x^\rho}{\rho(\rho+1)} - E(x)$$

where the sum is taken over all zeroes ρ of the ζ function in the critical strip $\{0 \leq \operatorname{Re} s \leq 1\}$, and $E(x) = O(x)$ is an error term.[7] (There is also a corresponding explicit formula for $\psi(x)$, except that that sum does not converge absolutely, contrary to the one for $\psi_1(x)$ given above.) This highlights, for instance, the importance of the famous Riemann hypothesis, that all zeroes of ζ on the critical

[7] See Problem 2 in Chapter 7 of [9].

strip $\{0 < \operatorname{Re} s < 1\}$ is on the line $\operatorname{Re} s = 1/2$. It has been shown in [7] that the Riemann hypothesis implies that

$$|\psi(x) - x| < \frac{1}{8\pi}\sqrt{x}(\log x)^2$$

for all $x \geq 73.2$, and that

$$|\pi(x) - \mathrm{Li}(x)| < \frac{1}{8\pi}\sqrt{x}\log x$$

for all $x \geq 2657$. (Note that the power of x on the right hand side is essentially $1/2$.) Also, the Riemann hypothesis can be shown [11] to be equivalent to the estimate

$$|\pi(x) - \mathrm{Li}(x)| \leq C\sqrt{x}\log x$$

for some constant C. There are actually many other equivalent forms of Riemann hypothesis, which is beyond our scope of discussion here.

References

[1] Tom M. Apostol, *A centennial history of the prime number theorem*, Number theory, Trends Math., Birkhäuser, Basel, 2000, pp. 1–14. MR1764793

[2] C.-L. de la Vallée-Poussin, *Recherches analytiques sur la théorie des nombres premiers,* Annales de la Société Scientifiques de Bruxelles **20B** (1896), 183–256 (French).

[3] P. Erdös, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Sci. U. S. A. **35** (1949).

[4] P. Garrett, *Simple Proof of the Prime Number Theorem*, online lecture notes (January 20, 2015).

[5] J. Hadamard, *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bull. Soc. Math. France **24** (1896), 199–220 (French). MR1504264

[6] D. J. Newman, *Simple analytic proof of the prime number theorem*, Amer. Math. Monthly **87** (1980), no. 9, 693–696.

[7] Lowell Schoenfeld, *Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II*, Math. Comp. **30** (1976), no. 134, 337–360. MR0457374

[8] Atle Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313.

[9] Elias M. Stein and Rami Shakarchi, *Complex analysis*, Princeton Lectures in Analysis, vol. 2, Princeton University Press, Princeton, NJ, 2003.

[10] B. Riemann, *Bestimmung einer Function einer veränderlichen complexen Größe durch Grenz- und Unstetigkeitsbedingungen*, J. Reine Angew. Math. **54** (1857), 111–114, DOI 10.1515/crll.1857.54.111 (German). MR1579034

[11] Helge von Koch, *Sur la distribution des nombres premiers*, Acta Math. **24** (1901), no. 1, 159–182, DOI 10.1007/BF02403071 (French). MR1554926

[12] D. Zagier, *Newman's short proof of the prime number theorem*, Amer. Math. Monthly **104** (1997), no. 8, 705–708.