Prime Number Theorem

Bent E. Petersen

Contents

1 Introduction

Let $\pi(x)$ be the number of primes $p \leq x$. It was discovered *empirically* by Gauss about 1793 (letter to Enke in 1849, see Gauss [9], volume 2, page 444 and Goldstein [10]) and by Legendre (in 1798 according to [14]) that

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

This statement is the *prime number theorem.* Actually Gauss used the equivalent formulation (see page 10)

$$\pi(x) \sim \int_2^x \frac{dt}{\log t}.$$

For some discussion of Gauss' work see Goldstein [10] and Zagier [45].

In 1850 Čebyšev [3] proved a result far weaker than the prime number theorem — that for certain constants $0 < A_1 < 1 < A_2$

$$A_1 < \frac{\pi(x)}{x/\log x} < A_2.$$

An elementary proof of Čebyšev's theorem is given in Andrews [1]. Čebyšev introduced the functions

$$\begin{aligned} \theta(x) &= \sum_{p \leq x,\, p \text{ prime}} \log p, \quad \text{(Čebyšev theta function)} \\ \psi(x) &= \sum_{p^n \leq x,\, p \text{ prime}} \log p, \quad \text{(Čebyšev psi function).} \end{aligned}$$

Note

$$\psi(x) = \sum_{n=1}^{\infty} \theta(x^{1/n})$$

where the sum is finite for each x since $\theta(x^{1/n}) = 0$ if $x < 2^n$. aČebyšev proved that the prime number theorem is equivalent to either of the relations

$$\theta(x) \sim x, \quad \psi(x) \sim x.$$

In addition Čebyšev showed that if $\lim_{x\to\infty} \frac{\theta(x)}{x}$ exists, then it must be 1, which then implies the prime number theorem. He was, however, unable to establish the existence of the limit.

Like Gauss, Riemann formulated his estimate of $\pi(x)$ in terms of the logarithmic integral

$$\mathrm{Li}(x) = (\mathrm{PV}) \int_0^x \frac{dt}{\log t}, \quad x > 1.$$

In his famous 1859 paper [33] he related the relative error in the asymptotic approximation

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

to the distribution of the *complex* zeros of the *Riemann zeta function*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \left(1 - p^{-s}\right)^{-1}. \tag{1}$$

The Riemann zeta function was actually introduced by Euler as early as 1737. It was used of by Čebyšev (in the real domain) prior to Riemann's use of it. Euler also discovered the *functional equation*

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

which he published in 1749. The functional equation was proved by Riemann in [33].

Riemann does not prove the prime number theorem in his 1859 paper. His object was to find an explicit analytic expression for $\pi(x)$, and he does so. He does comment that $\pi(x)$ is about $\mathrm{Li}(x)$ and that $\pi(x) = \mathrm{Li}(x) + O(x^{1/2})$. This would imply

$$\frac{\pi(x)}{\mathrm{Li}(x)} = 1 + O(x^{-1/2}\log x) = 1 + o(1),$$

which gives the prime number theorem. Thus Hardy's comment, [14], page 352, that Riemann does not even state the prime number theorem, is not strictly accurate. On the other hand, Riemann's assertion about the order of the error is so much stronger than what is required for the prime number theorem, that one could maintain that he does not state the weaker theorem.

In 1896 the prime number theorem was finally proved by Jacques Hadamard [12] and also by Charles–Jean de la Vallée Poussin [6]. The first part of the proof is to show that $\zeta(s) \neq 0$ if $\mathfrak{Re}\, s = 1$. As a general principle, finding zero–free regions for the zeta function in the critical strip leads to better estimates of the error in the $\pi(x) \sim \mathrm{Li}(x)$ (see, for example, theorem 1.2).

Towards the end of his 1859 paper [33] Riemann asserts that $\pi(x) < \mathrm{Li}(x)$. Gauss [9], volume 2, page 444, makes the same assertion. This is known to be true for all $x \leq 10^8$ but was proved false in general in 1914 by Littlewood, [27]. Littlewood showed that $\pi(x) - \mathrm{Li}(x)$ changes sign infinitely often. Indeed Littlewood showed there is a constant $K > 0$ such that

$$\frac{(\pi(x) - \mathrm{Li}(x))\log x}{x^{1/2}\log\log\log x}$$

is greater that K for arbitrarily large x and less than $-K$ for arbitrarily large x. Littlewood's methods yield no information on where the first sign change occurs. In 1933 Skewes [35] showed that there is at least one sign change at x for some

$$x < 10^{10^{10^{34}}}.$$

Skewes proof required the Riemann hypothesis. In 1955 [36] he obtained a bound without using the Riemann hypothesis. This new bound was

$$10^{10^{10^{964}}}.$$

Skewes large bound can be reduced substantially. In 1966 Sherman Lehman [25] showed that between 1.53×10^{1165} and 1.65×10^{1165} there are more than 10^{500} successive integers x for which $\pi(x) > \mathrm{Li}(x)$. Lehman's work suggests there is no sign change before 10^{20}. Perhaps we will see a sign change soon! In 1987 te Riele [37] showed that between 6.62×10^{370} and 6.69×10^{370} there are more than 10^{180} successive integers x for which $\pi(x) > \mathrm{Li}(x)$.

In Ramanujan's second letter to Hardy (in 1913, see [2], page 53) he estimates $\pi(x)$ by

$$\pi(x) \approx \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \,\mathrm{Li}(x^{1/n}) \tag{3}$$

where $\mu(n)$ is the Möbius function. This expression was obtained by Riemann in 1859, except Riemann has additional terms, arising from the complex zeros of $\zeta(s)$. Littlewood points out in a letter to Hardy (discussing Ramanujan's letter, see [2], page 68) that

$$\pi(x) - \mathrm{Li}(x) + \frac{1}{2}\,\mathrm{Li}(x^{1/2}) \neq O\left(\frac{x^{1/2}}{\log x}\right).$$

It follows that

$$\pi(x) - \mathrm{Li}(x) + \frac{1}{2}\,\mathrm{Li}(x^{1/2}) \neq O\left(\mathrm{Li}(x^{1/2})\right). \tag{4}$$

Thus it is clear that equation (3) can not be interpreted as an asymptotic series for $\pi(x)$ (though it is an asymptotic series). Ramanujan says to truncate the series at the first term less than one. This gives an excellent approximation to $\pi(x)$, but it is *empirical*.

The actual expression obtained by Riemann is

$$\begin{aligned} \pi(x) &= \sum_{n=1}^{\infty} \frac{\mu(n)}{n}\, J(x^{1/n}) \\ J(x) &= \mathrm{Li}(x) - \sum_{\rho} \mathrm{Li}(x^{\rho}) - \log 2 + \int_x^{\infty} \frac{dt}{t(t^2-1)\log t} \end{aligned}$$

where ρ runs over the complex roots of the zeta function. The first sum here is actually finite for each x since

$$J(x) = \sum \frac{1}{n}\pi(x^{1/n})$$

is 0 for $x < 2$. The complete proof of Riemann's formula (in a different form) was given by von Mangoldt ([43]) in 1895. In connection with equation (3) we now have

$$\pi(x) - \sum_{n=1}^{N} \frac{\mu(n)}{n}\,\mathrm{Li}(x^{1/n}) = \sum_{n=1}^{N}\sum_{\rho} \mathrm{Li}(x^{\rho/n}) + \text{“other terms”}$$

where the omitted terms are not particularly significant. The terms in the double sum are Riemann's "periodic" terms. Individually they are quite large, but there must be a large amount of cancellation to account for the fact that equation (3) gives a very close estimate of $\pi(x)$.

In the following table the column labelled "Riemann" was computed using the first 6 terms in the series (3). The column labelled "Li(x)" could just as well be labelled "Gauss" as Li(x) differs by about $1.045\ldots$ from $\int_2^x dt/\log t$. The table shows why Li(x) is preferred to the asymptotically equivalent expression $x/\log x$.

x	$\pi(x)$	$x/\log x$	$\mathrm{Li}(x)$	Riemann
500	95	80.4	101.7	94.3
1,000	168	144.7	177.6	168.3
2,000	303	263.1	314.8	303.3
500,000	41,638	38,102.8	41,606.2	41,529.4
1,000,000	78,498	72,382.4	78,627.5	78,527.3
1,500,000	114,155	105,477.9	114,263.0	114,145.7
2,000,000	148,933	137,848.7	149,054.8	148,923.4
2,500,000	183,072	169,700.9	183,244.9	183,101.4
3,000,000	216,816	201,151.6	216,970.5	216,816.2

The prime counts in the table above are taken from Edwards [8] and are due to D. N. Lehmer [26]. The other numbers were computed using Maple V3. While the "Riemann" column looks better it turns out that in the long run Li(x) is just as good – see Littlewood [27] and the discussion in Edwards [8], page 87.

Here is some additional numerical evidence for the prime number theorem:

x	$\pi(x)$	$x/\log x$	$\mathrm{Li}(x)$
10^4	1,229	1085.7	1246.1
10^8	5,761,455	5.42×10^6	5.762×10^6
10^{12}	37,607,912,018	3.61×10^{10}	3.760795×10^{10}
10^{16}	279,238,341,033,925	2.71×10^{14}	$2.79238344 \times 10^{14}$
10^{18}	24,739,954,287,740,860	2.41×10^{16}	$2.4739954309 \times 10^{16}$

The prime counts in the table above are taken from Crandal [5] and are due to Reisel [32] and Odlyzko. The other numbers were computed using Maple V3.

Here's a strong version of the prime number theorem (see Ivić [22]) expressed in terms of the Čebyšev psi function ψ.

Theorem 1.1. *There exists a constant $C > 0$ such that*

$$\psi(x) = x + O\left(x e^{-C(\log x)^{3/5}(\log\log x)^{-1/5}}\right).$$

With regard to the connection between the complex zeros of the zeta function and the estimate of the error in the prime number theorem (see [22]) we have:

Theorem 1.2. *Let $\frac{1}{2} \le \alpha < 1$. Then*

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

if and only if

$$\zeta(s) \ne 0 \text{ for } \mathfrak{Re}\, s > \alpha.$$

2 Asymptotics

Let f and g be functions defined in a neighborhood of a. The notation

$$f(x) = o(g(x)), \quad x \to a$$

means

$$\lim_{x \to a} f(x)/g(x) = 0.$$

The notation

$$f(x) = O(g(x)), \quad x \to a$$

means there is a constant M such that

$$|\, f(x)\,| \le M \;|\, g(x)\,| \text{ for all } x \text{ near } a.$$

The notation

$$f(x) \sim g(x), \quad x \to a$$

means

$$\lim_{x \to a} f(x)/g(x) = 1, \quad \text{that is} \quad f(x) = g(x) + o(g(x)), \quad x \to a.$$

We frequently omit the expression "$x \to a$", especially if $a = \infty$ or if a may be inferred from the context.

Sometimes we want more detailed information about the remainder $o(g(x))$ above. Let $(g(x))_{n \ge 1}$ be a sequence of functions such that

$$g_{n+1}(x) = o(g_n(x))$$

for each $n \geq 1$. Then

$$f(x) \sim \sum_{n=1}^{\infty} c_n \, g_n(x) \tag{5}$$

means

$$f(x) - \sum_{k=1}^{n} c_k \, g_k(x) = o(g_n(x)) \tag{6}$$

for each $n \geq 1$.

By "going out one more term" we see that (6) is equivalent to

$$f(x) - \sum_{k=1}^{n} c_k \, g_k(x) = O(g_{n+1}(x)) \tag{7}$$

for each $n \geq 1$.

The series (5) need not converge. Taking *longer* partial sums need not improve the approximation to $f(x)$, unless we also take larger x. Even if the series (5) converges, it need not converge to the function $f(x)$.

The next result is a useful Tauberian result for estimating an integrand.

Lemma 2.1. *Let f be a function on $[2, \infty)$ and suppose $xf(x)$ is monotone nondecreasing on $[2, \infty)$. Let m and n be real numbers, $n \neq -1$. If*

$$\int_2^x f(t) dt \sim \frac{x^{n+1}}{(\log x)^m}, \quad x \to \infty,$$

then

$$f(x) \sim \frac{(n+1)x^n}{(\log x)^m}, \quad x \to \infty.$$

Proof. Let $x \geq 2$, $\epsilon > 0$ and $x(1-\epsilon) \geq 2$. Then

$$\int_x^{x(1+\epsilon)} f(t) dt = \int_x^{x(1+\epsilon)} t f(t) \frac{dt}{t} \geq x f(x) \int_1^{1+\epsilon} \frac{dt}{t} = x f(x) \log(1+\epsilon)$$

and

$$\int_{x(1-\epsilon)}^x f(t) dt = \int_{x(1-\epsilon)}^x t f(t) \frac{dt}{t} \leq x f(x) \int_{1-\epsilon}^1 \frac{dt}{t} = -x f(x) \log(1-\epsilon).$$

Let $\epsilon > 0$. By hypothesis there exists $A_\epsilon \geq 2$ such that $x \geq A_\epsilon$ implies

$$\frac{x^{n+1}(1-\epsilon^2)}{(\log x)^m} \leq \int_2^x f(t)dt \leq \frac{x^{n+1}(1+\epsilon^2)}{(\log x)^m}.$$

Thus if $0 < \epsilon < 1$ and $x \geq \max\left(\frac{2}{1-\epsilon}, A_\epsilon\right)$ then

$$\begin{aligned} xf(x)\log(1+\epsilon) &\leq \int_x^{x(1+\epsilon)} f(t)dt \\ &= \int_2^{x(1+\epsilon)} f(t)dt - \int_2^x f(t)dt \\ &\leq \frac{x^{n+1}(1+\epsilon)^{n+1}(1+\epsilon^2)}{(\log x(1+\epsilon))^m} - \frac{x^{n+1}(1-\epsilon^2)}{(\log x)^m} \end{aligned}$$

$$\begin{aligned} -xf(x)\log(1-\epsilon) &\geq \int_{x(1-\epsilon)}^{x} f(t)dt \\ &= \int_2^{x} f(t)dt - \int_2^{x(1-\epsilon)} f(t)dt \\ &\geq \frac{x^{n+1}(1-\epsilon^2)}{(\log x)^m} - \frac{x^{n+1}(1-\epsilon)^{n+1}(1+\epsilon^2)}{(\log x(1-\epsilon))^m} \end{aligned}$$

It follows if $0 < \epsilon < 1$ then

$$\limsup_{x\to\infty} \frac{(\log x)^m f(x)}{x^n} \leq \frac{(1+\epsilon)^{n+1}(1+\epsilon^2) - (1-\epsilon^2)}{\log(1+\epsilon)}$$

$$\liminf_{x\to\infty} \frac{(\log x)^m f(x)}{x^n} \geq \frac{(1-\epsilon^2) - (1-\epsilon)^{n+1}(1+\epsilon^2)}{-\log(1-\epsilon)}.$$

Now the two expressions in ϵ above have limit $n+1$ as $\epsilon \to 0$. □

Compare lemma 2.1 with Rademacher [31], page 102, Edwards [8], page 82, and Grosswald [11], page 175.

Note that asymptotic estimates are not preserved by exponentiation. For example, we have

Lemma 2.2. *If $m > 0$ then there exists a continuous increasing function g on $[1,\infty)$ such that*

1. $0 < g(x) < x$

2. $g(x) = o\left(\dfrac{x}{(\log x)^m}\right), \quad x \to \infty$

3. $\log g(x) \sim \log x, \quad x \to \infty$

Proof. $x/(\log x)^{m+1}$ has its minimum $(m+1)^{-(m+1)}\mathrm{e}^{m+1}$ at $x = \mathrm{e}^{m+1}$. Hence if we define

$$g(x) = \begin{cases} x(m+1)^{-(m+1)} & \text{if } 1 \le x \le \mathrm{e}^{m+1} \\ \frac{x}{(\log x)^{m+1}} & \text{if } x \ge \mathrm{e}^{m+1} \end{cases}$$

then g is continuous and the first two conclusions hold. Now note if $x \ge \mathrm{e}^{m+1}$ then

$$\log g(x) = \log x - (m+1)\log(\log x).$$

Since

$$\lim_{x\to\infty} \frac{\log(\log x)}{\log x} = 0$$

the last conclusion holds. □

Lemma 2.2 is adapted from [31], page 96.

3 The Logarithmic Integral

The logarithmic integral $\mathrm{Li}(x)$, $x > 1$, is defined as the Cauchy principal value of the divergent integral $\int_0^\infty dt/\log t$. Explicitly

$$\begin{aligned} \mathrm{Li}(x) &= \lim_{\epsilon\to 0}\left\{\int_0^{1-\epsilon} \frac{dt}{\log t} + \int_{1+\epsilon}^{x} \frac{dt}{\log t}\right\}. \\ &= \lim_{\epsilon\to 0}\left\{\int_0^{1-\epsilon} \left(\frac{1}{\log t} - \frac{1}{t-1}\right) dt + \log\epsilon \right. \\ &\qquad \left. + \int_{1+\epsilon}^{x} \left(\frac{1}{\log t} - \frac{1}{t-1}\right) dt + \log(x-1) - \log\epsilon \right\} \\ &= \log(x-1) + \int_0^x \left(\frac{1}{\log t} - \frac{1}{t-1}\right) dt. \end{aligned}$$

We can avoid the singularity in the integrand by noting that for any $\mu > 1$ we have

$$\mathrm{Li}(x) = \mathrm{Li}(\mu) + \int_\mu^x \frac{dt}{\log t}.$$

Thus

$$\mathrm{Li}(x) = 1.045163780\cdots + \int_2^x \frac{dt}{\log t}$$

which shows that estimate of Gauss for $\pi(x)$ does not differ by more than about 1 from the estimate given by $\mathrm{Li}(x)$. It is fairly easy to see that $\mathrm{Li}(x)$ has exactly one root $\mu > 1$ and therefore

$$\mathrm{Li}(x) = \int_\mu^x \frac{dt}{\log t}$$

where $\mu = 1.45136923488338105\ldots$. This expression is the one Ramanujan used for $\mathrm{Li}(x)$.

Proposition 3.1.

$$\mathrm{Li}(x) \sim \sum_{n=1}^{\infty} (n-1)! \frac{x}{(\log x)^n}, \quad x \to \infty.$$

Proof. If we integrate by parts n times we have

$$\mathrm{Li}(x) = \frac{x}{\log x} + \cdots + (n-1)! \frac{x}{(\log x)^n} + c_n + n! \int_2^x \frac{dt}{(\log t)^{n+1}}$$

where c_n is a constant. It now suffices to show

$$\lim_{x\to\infty} \frac{(\log x)^n}{x} \left(c_n + n! \int_2^x \frac{dt}{(\log t)^{n+1}} \right) = 0.$$

Dividing the interval of integration at $x^{1/2}$ we have

$$\begin{aligned} & \lim_{x\to\infty} \frac{(\log x)^n}{x} \int_2^x \frac{dt}{(\log t)^{n+1}} \\ \le\ & \frac{(\log x)^n}{x} \frac{x^{1/2}-2}{(\log 2)^{n+1}} + \frac{(\log x)^n}{x} \frac{x - x^{1/2}}{(\log x)^{n+1}} 2^{n+1} \\ \le\ & \left(\frac{\log x}{x^{1/2n}} \right)^n + \frac{2^{n+1}}{\log x} \end{aligned}$$

which converges to 0 as $x \to \infty$. □

The idea of dividing the interval of integration at $x^{1/2}$ is a standard trick. See Edwards [8], page 85.

In particular $\mathrm{Li}(x) \sim x/\log x$ so that $\pi(x) \sim \mathrm{Li}(x)$ and $\pi(x) \sim x/\log x$ are equivalent formulations of the prime number theorem.

4 The Čebyšev Functions $\theta(x)$ and $\psi(x)$

The Čebyšev theta function θ is defined by

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

where p runs over primes. There is a simple relationship between $\theta(x)$ and $\pi(x)$:

Proposition 4.1.

$$\pi(x) = \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t\,(\log t)^2}\,dt. \tag{8}$$

Proof. Let $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $\cdots$ be the sequence of primes and let

$$c_k = \sum_{j=1}^{k} \log p_j \quad \text{so} \quad c_k - c_{k-1} = \log p_k.$$

Note if $x \in [p_k, p_{k+1})$ then $\pi(x) = k$ and $\theta(x) = c_k$. Thus

$$\begin{aligned}
\int_2^x \frac{\theta(t)}{t\,(\log t)^2}\,dt &= \sum_{j=1}^{k-1} \int_{p_j}^{p_{j+1}} \frac{c_j}{t\,(\log t)^2}\,dt + \int_{p_k}^{x} \frac{c_k}{t\,(\log t)^2}\,dt \\
&= \sum_{j=1}^{k-1} c_j \left(\frac{1}{\log p_j} - \frac{1}{\log p_{j+1}} \right) + c_k \left(\frac{1}{\log p_k} - \frac{1}{\log x} \right) \\
&= \frac{c_1}{\log p_1} + \sum_{j=2}^{k} \frac{c_j - c_{j-1}}{\log p_j} - \frac{c_k}{\log x} \\
&= 1 + \sum_{j=2}^{k} 1 - \frac{\theta(x)}{\log x} \\
&= \pi(x) - \frac{\theta(x)}{\log x}.
\end{aligned}$$

□

If we work with Stieltjes integrals we see easily that

$$\pi(x) = 1 + \int_2^x \frac{d\theta(t)}{\log t} = \int_\mu^x \frac{d\theta(t)}{\log t}$$

where $1 < \mu < 2$. Then an integration by parts (see Kestelman [23]) yields the proposition. The striking resemblance to the logaritmic integral is very suggestive.

The integral in equation (8) is actually $o(x/\log x)$. This follows immediately from the proof of proposition 3.1 once we know $\theta(x) = O(x)$. The only obvious estimate for $\theta(x)$ however is

$$\theta(x) = \sum_{p \le x} \log p \le \pi(x) \log x \le x \log x \tag{9}$$

Corollary 4.2. *If* $\theta(x) \sim x$ *then* $\pi(x) \sim \dfrac{x}{\log x}$.

Proof. As pointed out above, if $\theta(x) = O(x)$ then

$$\frac{\pi(x)}{x/\log x} = \frac{\theta(x)}{x} + o(1).$$

□

Let

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

Then the Čebyšev psi function ψ is defined by

$$\psi(x) = \sum_{n \le x} \Lambda(n) = \sum_{p^n \le x} \log p$$

where, as usual, p runs over primes. We note

$$\psi(x) = \sum_{n=1}^{\infty} \theta(x^{1/n}) \tag{10}$$

where the sum is actually finite for each x since $\theta(x) = 0$ if $x < 2$.

If n is the largest integer such that $p^n \le x$ then $n \log p$ is the contribution of the powers of p to $\psi(x)$ and $n \le \left[\frac{\log x}{\log p}\right]$, where $[z]$ is the greatest integer in z. Thus

$$\psi(x) \le \sum_{p \le x} \left[\frac{\log x}{\log p}\right] \log p \le \log x \sum_{p \le x} 1 = \pi(x) \log x.$$

Lemma 4.3.

$$\psi(x) = \theta(x) + O\left(x^{1/2}(\log x)^2\right). \tag{11}$$

Proof. In equation (10) there are at most $\frac{\log x}{\log 2}$ nonzero terms and they decrease in magnitude. Thus

$$\theta(x) \le \psi(x) \le \theta(x) + \theta(x^{1/2}) \log x/\log 2.$$

Now apply (9). □

Corollary 4.4. *We have $\theta(x) \sim x$ if and only if $\psi(x) \sim x$.*

Proof.

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

□

Lemma 4.5. *If $\psi(x)$ is the Čebyšev psi function then*

$$\pi(x) \sim \frac{x}{\log x} \text{ if and only if } \psi(x) \sim x.$$

Proof. We have already shown

$$\frac{\psi(x)}{x} \le \frac{\pi(x)}{x/\log x}. \tag{12}$$

Now suppose $\pi(x) \sim x/\log x$. Then by (12) we have $\psi(x) = O(x)$. It follows that $\theta(x) = \psi(x) + O(x^{1/2}(\log x)^2) = O(x)$ and therefore by (8) we have

$$\frac{\pi(x)}{x/\log x} = \frac{\theta(x)}{x} + o(1).$$

Then it follows that $\theta(x) \sim x$ and so by corollary 4.4 we have $\psi(x) \sim x$. Conversely if $\psi(x) \sim x$ then by corollary 4.4 we have $\theta(x) \sim x$ and so by corollary 4.2 we are done.

Alternate proof: Choose an increasing continuous function g on $[1, \infty)$ such that $0 < g(x) < x$, $g(x) = o(x/\log x)$ and $\log g(x) \sim \log x$. Since $\pi(x) \le x$ we have

$$\begin{aligned}\pi(x) &= \pi(g(x)) + (\pi(x) - \pi(g(x))) \\ &\leq g(x) + \sum_{g(x)<p\leq x} 1 \\ &\leq g(x) + \sum_{p\leq x} \frac{\log p}{\log g(x)} \\ &\leq g(x) + \frac{\psi(x)}{\log g(x)}.\end{aligned}$$

Thus

$$\frac{\pi(x)}{x/\log x} \leq \frac{g(x)}{x/\log x} + \frac{\log x}{\log g(x)}\frac{\psi(x)}{x}.$$

If $\psi(x) \sim x$ then it follows that $\limsup \pi(x)/(x/\log x) \leq 1$. If $\pi(x) \sim x/\log x$ then it follows that $\liminf \psi(x)/x \geq 1$. In view of equation (12) the proof is complete. □

The alternate proof of Lemma 4.5 is adapted from [31], page 96. The lemma shows we can express the prime number theorem as $\psi(x) \sim x$. Note at this point the existence of either of the limits $\lim \pi(x)/(x/\log x)$ and $\lim \psi(x)/x$ is not clear.

Exercise 4.6. *Prove the estimates $\theta(x) = O(x)$ and $\psi(x) = O(x)$ by elementary means (see Čebyšev's work) and then use them to simplify all of arguments above. Note we would also have $\psi(x) = \theta(x) + O(x^{1/2}\log x)$ which is somewhat better than equation 11.*

5 Möbius Inversion

We define the Möbius function $\mu(n)$ by

$$\mu(n) = \begin{cases} 1 \text{ if } n = 1 \\ (-1)^m \text{ if } n \text{ is a product of } m \text{ distinct primes} \\ 0 \text{ in all other cases.} \end{cases}$$

Note that $\mu(n) = 0$ if n is not square–free. An important property of $\mu(n)$ is

$$\sum_{d|n} \mu(d) = \begin{cases} 1 \text{ if } n = 1 \\ 0 \text{ otherwise} \end{cases}$$

Let $a = (a_1, a_2, \cdots) \in \ell^1$ and define

$$b_n = \sum_{m \geq 1} a_{mn}.$$

Then $b = (b_1, b_2, \cdots) \in \ell^\infty$. Suppose we have $b \in \ell^1$. Then we can consider

$$\sum_{n \geq 1} \mu(n) b_{mn}.$$

Substituting the definition of b_{mn} and making a change of variable we obtain

$$\sum_{n \geq 1} \sum_{k \geq 1} \mu(n) a_{mnk} = \sum_{h \geq 1} \sum_{d|h} \mu(d) a_{mh} = a_m.$$

Note if $a_n = 0$ for $n > N$ then $b_n = 0$ for $n > N$ and all the sums are finite. In this case the formula

$$a_m = \sum_{n \geq 1} \mu(n) b_{mn}$$

is called the Möbius inversion formula. For a discussion of the extension to ℓ^1 sequences see Hartman and Wintner [18].

Proposition 5.1. *Let $h : \mathbb{N} \to \mathbb{R}$ be totally multiplicative in the sense that $h(mn) = h(m)h(n)$ and $h(1) = 1$. Let f be a function on $[1, \infty)$ such that $f(x) = 0$ if $x < 2$. If*

$$F(x) = \sum_{n=1}^{\infty} h(n) f(x^{1/n}), \quad x \geq 1 \tag{13}$$

then

$$f(x) = \sum_{n=1}^{\infty} \mu(n) h(n) F(x^{1/n}). \tag{14}$$

Proof. Fix x and let

$$a_m = h(m) f(x^{1/m}).$$

Note that $a_m = 0$ if $m > \frac{\log x}{\log 2}$. Now let

$$\begin{aligned} b_n &= \sum_{m \geq 1} a_{mn} \\ &= \sum_{m \geq 1} h(mn) f(x^{1/mn}) \\ &= h(n) \sum_{m \geq 1} h(m) f(x^{1/mn}) \\ &= h(n) F(x^{1/n}). \end{aligned}$$

Then by the Möbius inversion formula

$$f(x) = h(1)f(x) = a_1 = \sum_{n\geq 1} \mu(n) b_n = \sum_{n\geq 1} \mu(n)h(n)F(x^{1/n}).$$

$\square$

Note $h(n) = n^k$ satisfies the multiplicative hypothesis. The cases $k = 0$ and $k = -1$ occur frequently. Recall if $\theta(x)$ is the Čebyšev theta function and $\psi(x)$ is the Čebyšev psi function then

$$\psi(x) = \sum_{n=1}^{\infty} \theta(x^{1/n}).$$

Thus we now have

$$\theta(x) = \sum_{n=1}^{\infty} \mu(n)\psi(x).$$

Likewise if we introduce

$$J(x) = \sum_{n=1}^{\infty} \frac{1}{n}\pi(x^{1/n})$$

then

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} J(x^{1/n}).$$

Note Riemann denoted $\pi(x)$ by $F(x)$ and $J(x)$ by $f(x)$.

6 The Tail of the Zeta Series

To estimate the zeta function we will need an expression for the *tail* of the zeta series, that is,

$$\sum_{n=k}^{\infty} \frac{1}{n^s}.$$

Consider first a partial sum of the tail of the series

$$\zeta(s,k,m) = \sum_{n=k+1}^{m} \frac{1}{n^s}.$$

Clearly

$$\begin{aligned} \zeta(s,k,m) &= \sum_{n=k+1}^{m} \frac{n-(n-1)}{n^s} \\ &= -k^{1-s} + m^{1-s} + \sum_{n=k}^{m-1} n\left(\frac{1}{n^s} - \frac{1}{(n+1)^s}\right). \end{aligned}$$

For the summands here we have

$$n\left(\frac{1}{n^s}-\frac{1}{(n+1)^s}\right)=sn\int_n^{n+1}x^{-s-1}dx=s\int_n^{n+1}[x]x^{-s-1}dx$$

where $[x]$ is the greatest integer in x. It follows that

$$\zeta(s,k,m)=-k^{1-s}+m^{1-s}+s\int_k^m [x]x^{-s-1}dx.$$

Trivially, if $s\neq 1$, we have

$$s\int_k^m x^{-s}dx=\frac{s}{1-s}\left(m^{1-s}-k^{1-s}\right)$$

and therefore

$$\zeta(s,k,m)=s\int_k^m \frac{[x]-x}{x^{s+1}}\,dx+\frac{1}{1-s}\left(m^{1-s}-k^{1-s}\right).$$

Letting $m\to\infty$ we obtain the tail of the zeta series

$$\zeta(s,k)=\sum_{n=k}^{\infty}n^{-s}=s\int_k^\infty \frac{[x]-x}{x^{s+1}}\,dx-\frac{k^{1-s}}{1-s}$$

for $\mathfrak{Re}\, s>1$. It follows that

$$\boxed{\zeta(s)=\sum_{n=1}^{k}n^{-s}-s\int_k^\infty \frac{x-[x]}{x^{s+1}}\,dx+\frac{k^{1-s}}{s-1}} \tag{15}$$

at least for $\mathfrak{Re}\, s>1$. But the integral converges absolutely for $\mathfrak{Re}\, s>0$ and so by uniqueness of analytic continuation we see that (15) holds for all $\mathfrak{Re}\, s>0$, $s\neq 1$. In particular we see that $\zeta(s)$ has a simple pole at $s=1$ and the residue at the pole is 1.

By taking $k=1$ in (15) and inserting the integral $\int_0^1 x^{-s}dx=1/(1-s)$ (for $\mathfrak{Re}\, s<1$) we obtain

$$\zeta(s)=-s\int_0^\infty \frac{x-[x]}{x^{s+1}}\,dx,\quad \text{for } 0<\mathfrak{Re}\, s<1.$$

In particular we have $\zeta(\sigma)<0$ if $0<\sigma<1$.

7 The Logarithm $\log\zeta(s)$

Recall $\zeta(s)$ is meromorphic in the plane with a simple pole at $s=1$ (with residue 1) and

$$\zeta(s)=\prod_p\left(1-p^{-s}\right)^{-1},\quad \mathfrak{Re}\, s>1,\quad p \text{ prime}.$$

Since $\zeta(s)$ has no zeros in $\mathfrak{Re}\, s > 1$ there is an analytic branch of the logarithm $\log\zeta(s)$ defined in the half space $\mathfrak{Re}\, s > 1$. Since $\zeta(\sigma) > 0$ for $\sigma > 1$ we may choose the branch of the logarithm so that $\log\zeta(\sigma)$ is real for $\sigma > 1$. Then for $\sigma > 1$ we have

$$\log\zeta(\sigma) = -\sum_p \log\left(1 - p^{-\sigma}\right). \tag{16}$$

For the principal branch of the logarithm we have

$$\log(1-z) = -\sum_{n=1}^{\infty} \frac{1}{n} z^n, \quad |\, z\,| < 1.$$

Therefore (16) yields

$$\log\zeta(\sigma) = \sum_p \sum_{n=1}^{\infty} \frac{1}{n} p^{-n\sigma}.$$

The double sum consists of positive terms. Hence we can change the order of summation to obtain

$$\log\zeta(\sigma) = \sum_{n=2}^{\infty} c_n n^{-\sigma}$$

where

$$c_n = \begin{cases} \frac{1}{m} \text{ if } n = p^m, & p \text{ prime} \\ 0 \text{ otherwise.} \end{cases}$$

The sum here continues to an analytic function in $\mathfrak{Re}\, s > 1$. Thus we obtain

$$\log\zeta(s) = \sum_{n=2}^{\infty} c_n n^{-s}, \quad \mathfrak{Re}\, s > 1, \tag{17}$$

with absolute convergence in $\mathfrak{Re}\, s > 1$ and uniform convergence in $\mathfrak{Re}\, s \geq 1+\epsilon$ for any $\epsilon > 0$. In particular we have normal convergence and so we can differentiate term–by–term. Thus

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=2}^{\infty} \Lambda(n) n^{-s}, \quad \mathfrak{Re}\, s > 1. \tag{18}$$

Here we have used $c_n \log n = \frac{\log n}{m} = \log p$ if $n = p^m$. Since $\Lambda(n) \leq \log(n)$ we have uniform convergence in $\mathfrak{Re}\, s \geq 1 + \epsilon$ for any $\epsilon > 0$.

From (18) we have

$$-\zeta'(s) = \sum_{m=1}^{\infty} m^{-s} \sum_{n=2}^{\infty} n^{-s}\Lambda(n) = \sum_{n=2}^{\infty} n^{-s} \sum_{d|n} \Lambda(d).$$

But

$$\sum_{d|n} \Lambda(d) = \sum_{p^m|n} \log p = \log n.$$

It follows that

$$\zeta'(s) = -\sum_{n=2}^{\infty} n^{-s} \log n, \quad \mathfrak{Re}\, s > 1. \tag{19}$$

Since $\log n = o(n^t)$ for any $t > 0$ the series converges absolutely in $\mathfrak{Re}\, s > 1$ and uniformly in $\mathfrak{Re}\, s \geq 1 + \epsilon$ for any $\epsilon > 0$. Of course (19) is completely obvious: the Dirichlet series for the zeta function converges normally in $\mathfrak{Re}\, s > 1$ and so may be differentiated term–by–term to yield (19).

Theorem 7.1.

$$\log \zeta(s) = s \int_2^{\infty} \frac{\pi(x)}{x\,(x^s - 1)} dx, \quad \mathfrak{Re}\, s > 1. \tag{20}$$

Proof. Since $\pi(x) \leq x$ the integral is absolutely convergent for $\mathfrak{Re}\, s > 1$ and defines an analytic function in the half-plane $\mathfrak{Re}\, s > 1$. Hence it suffices to verify the identity for $s = \sigma > 1$. The argument can be done by using Stieltjes integrals and integration by parts, but we will proceed by using summation by parts. We note first that

$$\pi(n) - \pi(n-1) = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

Hence from (16) we have for $\sigma > 1$

$$\begin{aligned} \log \zeta(\sigma) &= -\sum_{n=2}^{\infty} \big(\pi(n) - \pi(n-1)\big) \log(1 - n^{-\sigma}) \\ &= \lim_{N\to\infty} \Big(-\sum_{n=2}^{N} \pi(n) \log(1 - n^{-\sigma}) + \sum_{n=2}^{N} \pi(n-1) \log(1 - n^{-\sigma}) \Big) \\ &= \lim_{N\to\infty} \Big(-\sum_{n=2}^{N} \pi(n) \big[\log(1 - n^{-\sigma}) - \log(1 - (n+1)^{-\sigma})\big] \\ &\qquad + \pi(1) \log(1 - 2^{-\sigma}) + \pi(N) \log(1 - (N+1)^{-\sigma}) \Big) \end{aligned}$$

Now $\pi(1) = 0$ and

$$\begin{aligned}
\big|\,\pi(N)\log(1-(N+1)^{-\sigma})\,\big| &\le N\sum_{k=1}^{\infty}\frac{1}{k}(N+1)^{-k\sigma} \\
&\le \frac{N}{(N+1)^{\sigma}}\sum_{k=0}^{\infty}(N+1)^{-k} \\
&= \frac{1}{(N+1)^{\sigma-1}} \to 0
\end{aligned}$$

as $N \to \infty$ since $\sigma > 1$. It follows that

$$\begin{aligned}
\log\zeta(\sigma) &= -\sum_{n=2}^{\infty}\pi(n)\big(\log(1-n^{-\sigma}) - \log(1-(n+1)^{-\sigma})\big) \\
&= \sum_{n=2}^{\infty}\int_{n}^{n+1}\pi(x)\frac{d}{dx}\log(1-x^{-\sigma})dx \\
&= \int_{2}^{\infty}\frac{\sigma\pi(x)}{x(x^{\sigma}-1)}dx.
\end{aligned}$$

$\square$

Theorem 7.2.

$$\frac{\zeta'(s)}{\zeta(s)} = -s\int_0^{\infty} x^{-s-1}\psi(x)dx, \quad \Re\, s > 1, \tag{21}$$

where $\psi(x)$ is the Čebyšev psi function.

Proof. Since $\psi(x) = \sum_{n\le x}\Lambda(n)$ we have $\psi(n)-\psi(n-1) = \Lambda(x)$. Then by (18) we have

$$\begin{aligned}
-\frac{\zeta'(s)}{\zeta(s)} &= \lim_{N\to\infty}\sum_{n=2}^{N}(\psi(n)-\psi(n-1))n^{-s} && (22)\\
&= \lim_{N\to\infty}\Bigg(\sum_{n=2}^{N}\psi(n)(n^{-s}-(n+1)^{-s}) && (23)\\
&\qquad (N+1)^{-s}\psi(N) - \psi(1)\Bigg) && (24)\\
&= \sum_{n=2}^{\infty}\psi(n)(n^{-s}-(n+1)^{-s}) && (25)\\
& && (26)
\end{aligned}$$

since $\psi(1) = 0$ and $\psi(N) \leq \pi(N) \log N \leq N \log N$. Thus

$$-\frac{\zeta'(s)}{\zeta(s)} = -\sum_{n=2}^{\infty} \int_n^{n+1} \psi(x) \frac{d}{dx} x^{-s} dx = s \int_0^{\infty} x^{-s-1} \psi(x) dx$$

where everything is justified for $\mathfrak{Re}\, s > 1$ as in the previous theorem. $\square$

8 The Zeta Function on $\mathfrak{Re}\, s = 1$

Theorem 8.1. *If t is real then $\zeta(1 + \mathrm{i}t) \neq 0$ and $\zeta(\mathrm{i}t) \neq 0$.*

Proof. If $s = \sigma + \mathrm{i}t$, $\sigma > 1$ then

$$\log |\, \zeta(s)\, | = \mathfrak{Re}\, \log \zeta(s) = \sum_{n=2}^{\infty} c_n n^{-\sigma} \cos(t \log n).$$

It follows that

$$\begin{aligned} & \log \left|\, \zeta(\sigma)^3 \zeta(\sigma + \mathrm{i}t)^4 \zeta(\sigma + \mathrm{i}2t)\, \right| \\ = & \sum_{n=2}^{\infty} c_n n^{-\sigma} \left(3 + 4\cos(t \log n) + \cos(2t \log n)\right) \\ \geq & \ 0 \end{aligned}$$

since

$$3 + 4\cos t + \cos 2t = 2 + 4\cos t + 2(\cos t)^2 = 2(1 + \cos t)^2 \geq 0.$$

Thus

$$((\sigma - 1)\zeta(\sigma))^3 \left| \frac{\zeta(\sigma + \mathrm{i}t)}{\sigma - 1} \right|^4 |\, \zeta(\sigma + \mathrm{i}2t)\, | \geq \frac{1}{\sigma - 1} \tag{27}$$

for $\sigma > 1$ and for all t. Since $\zeta(s)$ has a simple pole at $s = 1$ with residue 1 we have

$$\lim_{\sigma \to 1} (\sigma - 1)\zeta(\sigma) = 1.$$

Now suppose $t \neq 0$ and $\zeta(1 + \mathrm{i}t) = 0$. Then

$$\lim_{\sigma \to 1} \frac{\zeta(\sigma + \mathrm{i}t)}{\sigma - 1} = \zeta'(1 + \mathrm{i}t)$$

and therefore (27) implies

$$\lim_{\sigma \to 1} |\, \zeta(\sigma + \mathrm{i}2t)\, | = \infty$$

which contradicts $t \neq 0$. By the functional equation we now also have $\zeta(\mathrm{i}t) \neq 0$ if t is real and $t \neq 0$. Finally the functional equation implies $\zeta(0) = -1/2$. $\square$

Theorem 8.1 is due to Hadamard and de la Vallée Poussin. The proof given above is from Rademacher [31], page 94, who attributes this particular proof to Ingham [21], page 29.

We now proceed to estimate $\zeta(\sigma+\mathrm{i}t)$ and, for us more importantly, $\log\zeta(\sigma+\mathrm{i}t)$ in $\sigma \ge 1$ and $|\,t\,| \ge 2$. The estimates we give are quite rough and unremittingly technical. The arguments here are adapted from [11], but they appear in a number of places. Much stronger (and more technical) results may be found in [22].

Lemma 8.2. *Let $C > 0$. Then there exists a constant $C_1 > 0$ such that*

$$|\,\zeta(s)\,| \le C_1 \log t$$

if $s = \sigma + \mathrm{i}t$, $1 - C/\log t \le \sigma \le 2$ and $t \ge 2$.

Proof. Note

$|\,s\,| = (\sigma^2+t^2)^{1/2} \le (4+t^2)^{1/2} \le \sqrt{2}t \le 2t$ and $|\,s-1\,| = ((\sigma-1)^2+t^2)^{1/2} \ge t$.

Now

$$\left|\int_k^\infty \frac{x-[x]}{x^{s+1}}\,dx\right| \le \int_k^\infty \frac{dx}{x^{\sigma+1}} = \frac{1}{\sigma k^\sigma}.$$

Consider a fixed $t \ge 2$ and choose k such that $k \le t < k+1$. Then for $n \le k$ we have

$$|\,n^{-s}\,| \le n^{-\sigma} \le \mathrm{e}^{(-1+C/\log k)\log n} \le n^{-1}\mathrm{e}^C$$

and so

$$\left|\sum_{n=1}^k n^{-s}\right| \le \mathrm{e}^C \sum_{n=1}^k \frac{1}{n} \le \mathrm{e}^c(1+\log k) \le \mathrm{e}^C(1+\log t).$$

It follows by equation (15) if $t \ge 2$ and we choose k with $k \le t < k+1$ then

$$|\,\zeta(s)\,| \le \frac{2t}{\sigma k^\sigma} + \frac{k^{1-\sigma}}{t} + \mathrm{e}^C(1+\log t) \tag{28}$$

for $1-C/\log t \le \sigma \le 2$. Since $t \ge 2$ then $k \ge 1$ and therefore $t/k \le 1+1/k \le 2$. Since $\sigma \le 2$ it follows that we have $(t/k)^{\sigma-1} \le 2$. Thus

$$\frac{k^{1-\sigma}}{t} \le 2t^{-\sigma}.$$

We have $k > t-1$. In addition $t \ge 2$ implies $1-t^{-1} \ge 1/2$ and so $\sigma \le 2$ implies $(1-t^{-1})^\sigma \ge 1/4$. Thus

$$\frac{2t}{\sigma k^\sigma} \le \frac{2t}{\sigma(t-1)^\sigma} = \frac{2t^{1-\sigma}}{\sigma(1-t^{-1})^\sigma} \le \frac{8t^{1-\sigma}}{\sigma}.$$

From (28) we now have

$$|\,\zeta(s)\,| \le \frac{8t^{1-\sigma}}{\sigma} + 2t^{-\sigma} + \mathrm{e}^{C}(1+\log t) \tag{29}$$

for any $t \ge 2$ and σ with $1 - C/\log t \le \sigma \le 2$. Now choose $t_0 > 2$ such that

$$1 - \frac{C}{\log t_0} > \frac{1}{2} \quad \text{and} \quad 2 + 17\mathrm{e}^{C} \le \log t_0.$$

For $2 \le t \le t_0$, $1 - C/\log t \le \sigma \le 2$ we have

$$|\,\zeta(s)\,| \le C_1 \le \frac{C_1}{\log 2}\log t$$

for some constant C_1 just by continuity. On the other hand $t \ge t_0$ implies

$$\frac{1}{2} < 1 - \frac{C}{\log t} \le \sigma \le 2$$

and so

$$\frac{8t^{1-\sigma}}{\sigma} + 2t^{-\sigma} \le 16t^{C/\log t} + 2t^{-1/2} \le 16\mathrm{e}^{C} + 2.$$

Thus

$$\begin{aligned} |\,\zeta(s)\,| &\le 17\mathrm{e}^{C} + 2 + \mathrm{e}^{C}\log t \\ &\le \log t_0 + \mathrm{e}^{C}\log t \le (1+\mathrm{e}^{C})\log t. \end{aligned}$$

□

Since $1 - 1/\log t \le 1$ if $t \ge 2$ we have:

Corollary 8.3.

$$\zeta(1 + \mathrm{i}t) = O(\log t)$$

Lemma 8.4. *Let $C > 0$. Then there exists a constant $C_1 > 0$ such that*

$$|\,\zeta'(s)\,| \le C_1\,(\log t)^2$$

if $s = \sigma + \mathrm{i}t$, $1 - C/\log t \le \sigma \le 2$ and $t \ge 2$.

Proof. By equation (15) we have

$$\begin{aligned} \zeta'(s) &= -\sum_{n=1}^{k} n^{-s}\log n \\ &- \int_k^\infty \frac{x-[x]}{x^{s+1}}\,dx + s\int_k^\infty \frac{(x-[x])\log x}{x^{s+1}}\,dx \\ &- -\frac{k^{1-s}}{(s-1)^2} - \frac{k^{1-s}\log k}{s-1}. \end{aligned}$$

Now we get much the same estimates as in the previous lemma but with an extra factor of $\log t$. $\square$

Corollary 8.5.
$$\zeta'(1+\mathrm{i}t) = O((\log t)^2)$$

Lemma 8.6.
$$\frac{\zeta'(1+\mathrm{i}t)}{\zeta(1+\mathrm{i}t)} = O((\log t)^9)$$

Proof. From equation (27) we have
$$\frac{1}{|\,\zeta(\sigma+\mathrm{i}t)\,|} \le |\,\zeta(\sigma)\,|^{3/4}\,|\,\zeta(\sigma+2\mathrm{i}t)\,|^{1/4}. \tag{30}$$

Since $\lim_{s\to 1}(s-1)\zeta(s) = 1$ it follows that
$$\frac{1}{|\,\zeta(\sigma+\mathrm{i}t)\,|} \le C_1\frac{(\log t)^{1/4}}{|\,\sigma-1\,|^{3/4}}.$$

Now for $\sigma > 1$ we have
$$\begin{aligned}
|\,\zeta(1+\mathrm{i}t)\,| &= \left|\zeta(\sigma+\mathrm{i}t) - \int_1^\sigma \zeta'(u+\mathrm{i}t)\,du\right| \\
&\ge |\,\zeta(\sigma+\mathrm{i}t)\,| - \left|\int_1^\sigma \zeta'(u+\mathrm{i}t)\,du\right| \\
&\ge \frac{1}{C_1}\frac{(\sigma-1)^{3/4}}{(\log t)^{1/4}} - C_2(\sigma-1)(\log t)^2.
\end{aligned}$$

If $A > 0$ and $B > 0$ then the function
$$g(u) = \frac{4A}{3}u^{3/4} - Bu, \quad u \ge 0,$$

has the maximum $A^4/(3B^3)$ at the point $u = A^4/B^4$. Therefore taking $A = 3/(4C_1(\log t)^{1/4})$, $B = C_2(\log t)^2$ and $\sigma - 1 = A^4/B^4$ we see that
$$|\,\zeta(1+\mathrm{i}t)\,| \ge \frac{3^3}{4^4C_1^4C_2^3}\,(\log t)^{-7}.$$

Since $|\,\zeta'(1+\mathrm{i}t)\,| = O((\log t)^2)$ the proof is complete. $\square$

Lemma 8.7. *There exist constants C and C_1 suct that*
$$\left|\frac{\zeta'(\sigma+\mathrm{i}t)}{\zeta(\sigma+\mathrm{i}t)}\right| \le C_1(\log t)^9$$

if $t \ge 2$ and $\sigma \ge 1 - C/(\log t)^9$.

Proof. Given any $C > 0$ we have constants C_2 and C_3 such that

$$\begin{aligned} |\,\zeta(\sigma+\mathrm{i}t)\,| &\leq |\,\zeta(1+\mathrm{i}t)\,| - \left| \int_1^\sigma \sigma'(u+\mathrm{i}t)\,du \right| \\ &\geq C_2(\log t)^{-7} - C_3\,|\,\sigma-1\,|\,(\log t)^2 \end{aligned}$$

if $t \geq 2$ and $\sigma \geq 1 - C/\log t$. Thus if $\sigma \geq C/(\log t)^9 \geq 1 - C/\log t$ then

$$|\,\zeta(\sigma+\mathrm{i}t)\,| \geq (C_2 - CC_3)(\log t)^{-7}.$$

Once we have found C_2 and C_3 for a given $C > 0$ then making C smaller simply strengthens the hypothesis on σ. Hence we may choose C so that $C_4 = C_2 - CC_3 > 0$. For this choice of C we know there is a constant $C_5 > 0$ such that

$$|\,\zeta'(\sigma+\mathrm{i}t)\,| \leq C_5(\log t)^2$$

for $t \geq 2$ and $\sigma \geq 1 - C/\log t$. □

We can now improve a bit on the theorem of Hadamard and de la Vallée Poussin.

Corollary 8.8. *There exists a constant $C > 0$ such that $\zeta(\sigma+\mathrm{i}t) \neq 0$ if $t \geq 2$ and $\sigma \geq 1 - C(\log t)^{-9}$.*

A much stronger zero–free result, due to H.-E. Richert, is known (see Ivić's book, [22]).

Theorem 8.9. *There exists a constant $C > 0$ such that $\zeta(s) \neq 0$ for*

$$\sigma \geq 1 - C(\log t)^{-2/3}(\log\log t)^{-1/3}, \quad t \geq 2.$$

Remarkably enough no one has ever proved that there is a zero–free region of the form $\mathfrak{Re}\, s \geq \sigma_0$ for some $\frac{1}{2} < \sigma_0 < 1$. Indeed it is even the case that no one has ever proved there is a zero–free *line* $\mathfrak{Re}\, s = \sigma$ for some $\frac{1}{2} < \sigma < 1$. Thus the result of Jacques Hadamard and Charles–Jean de la Vallée Poussin that there are no zeros on the line $\mathfrak{Re}\, s = 1$ is still the only know line result.

We now obtain a uniform estimate for the logarithm of the zeta function. We will use this estimate in the proof of the prime number theorem.

Theorem 8.10. *There exists a constant $C > 0$ such that*

$$|\log\zeta(\sigma+\mathrm{i}t)\,| \leq C(\log t)^9$$

for $t \geq 2$ and $\sigma \geq 1$.

Proof. Let $s = \sigma + \mathrm{i}t$. If $\sigma \geq 2$ then

$$\begin{aligned}\mathfrak{Re}\,\zeta(s) &= \sum_{n=1}^{\infty} \mathfrak{Re}\, n^{-s} \\ &= \sum_{n=1}^{\infty} n^{-\sigma} \cos(t \log n) \\ &\geq 1 - \sum_{n=2}^{\infty} n^{-\sigma} \\ &= 2 - \zeta(\sigma) \\ &\geq \frac{1}{3}\end{aligned}$$

since $\zeta(\sigma) \leq \zeta(2) = \pi^2/6 < 5/3$ for $\sigma \geq 2$. It follows that $|\arg\zeta(s)| \leq \pi/2$ for $\sigma \geq 2$. Then

$$|\log\zeta(s)| = |\log|\zeta(s)| + \mathrm{i}\arg\zeta(s)| \leq \log\zeta(2) + \pi/2 = K \leq K(\log 2)^{-9}(\log t)^9$$

if $\sigma \geq 2$ and $t \geq 2$. If now $1 - C(\log t)^{-9} \leq \sigma \leq 2$ and $t \geq 2$ then

$$|\log\zeta(\sigma + \mathrm{i}t)| \leq |\log\zeta(2 + \mathrm{i}t)| + \int_{\sigma}^{2} \left| \frac{\zeta'(u + \mathrm{i}t)}{\zeta(u + \mathrm{i}t)} \right| du.$$

The first term we estimated in the first part of the proof. Since $2 - \sigma \leq 1$ the integral is bounded by

$$\max_{\sigma \leq u \leq 2} \left| \frac{\zeta'(u + \mathrm{i}t)}{\zeta(u + \mathrm{i}t)} \right| \leq K(\log t)^9.$$

□

9 Mellin Transforms

If f is a function on $(0, \infty)$ its Mellin transform F is formally defined by the integral

$$F(s) = \int_0^{\infty} x^{s-1} f(x) dx. \tag{31}$$

A measurable function f on $(0, \infty)$ is said to be of *type* (α, β) if $x^{\sigma-1}f(x)$ is Lebesgue integrable on $(0, \infty)$ for each σ with $\alpha < \sigma < \beta$. This terminology, from Patterson [30], though it overloads the word *type*, is convenient. If f has type (α, β) then the Mellin transform $F(s)$ is analytic in the strip $\alpha < \Re\mathfrak{e}\, s < \beta$.

A classical example is the Gamma function

$$\Gamma(s) = \int_0^\infty x^{s-1} \mathrm{e}^{-x} dx.$$

Another example is the very first formula that Riemann derives for the zeta function in his 1859 paper,

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{x^{s-1}}{\mathrm{e}^x - 1} dx.$$

In (31) if we make the change of variable $x = \mathrm{e}^t$ we obtain

$$F(s) = \int_{-\infty}^\infty \mathrm{e}^{st} f(\mathrm{e}^t) dt.$$

Thus the Mellin transform F of f is the Fourier–Laplace transform of the composition $f \circ \exp$. In particular

$$F(\sigma - \mathrm{i}u) = \int_{-\infty}^\infty \mathrm{e}^{-\mathrm{i}ut} \left(\mathrm{e}^{\sigma t} f(\mathrm{e}^t)\right) dt$$

is the Fourier transform of $\mathrm{e}^{\sigma t} f(t)$. It follows that we have the inversion formula

$$\mathrm{e}^{\sigma t} f(\mathrm{e}^t) = \frac{1}{2\pi} \int_{-\infty}^\infty \mathrm{e}^{\mathrm{i}ut} F(\sigma - \mathrm{i}u)\, du,$$

that is,

$$f(\mathrm{e}^t) = \frac{1}{2\pi} \int_{-\infty}^\infty \mathrm{e}^{-(\sigma + \mathrm{i}u)} F(\sigma + \mathrm{i}u) du,$$

provided that

$$\int_{-\infty}^\infty |\, F(\sigma + \mathrm{i}u)\, |\, du < \infty.$$

Indeed in this case f is continuous (after correction on a null set, if needed) and the inversion formula folds for each t.

In the absence of the integrability hypothesis the inversion formula may continue to hold pointwise in a suitable principal–value sense, as well as in various non–pointwise senses, see Mellin [29], Hardy [15], [16], Titchmarsh [40], Patterson [30]. For example, by results of Dirichlet, Dini and Jordan, if f is locally of bounded variation then

$$\frac{f(\mathrm{e}^t + 0) + f(\mathrm{e}^t - 0)}{2} = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^T \mathrm{e}^{-(\sigma + \mathrm{i}u)} F(\sigma + \mathrm{i}u) du.$$

Theorem 9.1. *If f is of type (α, β) on $(0, \infty)$ then the Mellin transform*

$$F(s) = \int_0^\infty f(x)\, x^{s-1}\, dx$$

is analytic in the strip $\alpha < \mathfrak{Re}\, s > \beta$. Moreover

$$f(x) = \frac{1}{2\pi \mathfrak{i}} \int_{\sigma - \mathfrak{i}\infty}^{\sigma + \mathfrak{i}\infty} F(s) x^{-s}\, ds$$

for each σ with $\alpha < \sigma < \beta$ for which

$$\int_{-\infty}^{\infty} |\, F(\sigma + \mathfrak{i}u)\, |\, du < \infty.$$

Let f be a nonnegative monotone nondecreasing function on $[0, \infty)$ such that $f(x) = 0$ if $x < 2$ and such that $f(x) = O(x)$. Let

$$g(x) = \sum_{n=1}^{\infty} \frac{1}{n} f(x^{1/n}).$$

Then

$$f(x) \le g(x) \le f(x) + f(x^{1/2}) \frac{\log x}{\log 2}$$

and so

$$g(x) = f(x) + O(x^{1/2} \log x) = O(x).$$

Since $g(x) = 0$ for $x < 2$ we see that the Mellin transform

$$G(-s) = \int_0^\infty g(x) x^{-s-1} dx$$

defines a function analytic in $\mathfrak{Re}\, s > 1$. If we define

$$g_1(x) = \int_1^\infty \frac{g(t)}{t} dt$$

then we can integrate by parts to obtain

$$\frac{G(-s)}{s} = \int_0^\infty g_1(x) x^{-s-1} dx.$$

The integration by parts is justified since $g(x) x^{-s} = o(1)$ for $\mathfrak{Re}\, s > 1$.

We also note

$$\begin{aligned} G(-\sigma) &= \int_0^\infty \sum_{n=1}^{\infty} \frac{1}{n} f(t^{1/n}) t^{-\sigma-1} dt \\ &= \sum_{n=1}^{\infty} \int_0^\infty \frac{1}{n} f(t^{1/n}) t^{-\sigma-1} dt \\ &= \sum_{n=1}^{\infty} f(x) x^{-ns-1} dx, \end{aligned}$$

where the interchange of the summation and integration is justified since all the terms are nonnegative. We now interchange the summation and integration again and note that we have a geometric series in x^{-s}. Since $f(x) = 0$ if $x < 2$ we can sum the series to obtain

$$G(-s) = \int_0^\infty \frac{f(x)}{x(x^s - 1)} dx, \quad \Re s > 1.$$

Indeed, we have this relation for $s = \sigma > 1$, and both expressions are analytic in $\Re s > 1$. If we now apply the inverse Mellin transform, which we can do if we have a suitable estimate for G, we have

$$g(x) = \sum_{n=1}^\infty \frac{1}{n} f(x^{1/n})$$

$$G(-s) = \int_0^\infty g(x) x^{-s-1} dx = \int_0^\infty \frac{f(x)}{x(x^s - 1)} dx, \quad \Re s > 1$$

$$g(x) = \frac{1}{2\pi \mathrm{i}} \int_{\sigma - \mathrm{i}\infty}^{\sigma + \mathrm{i}\infty} G(-s) x^s ds, \quad \sigma > 1$$

$$g_1(x) = \int_1^x \frac{g(t)}{t} dt$$

$$g_1(x) = \frac{1}{2\pi \mathrm{i}} \int_{\sigma - \mathrm{i}\infty}^{\sigma + \mathrm{i}\infty} \frac{G(-s)}{s} x^s ds, \quad \sigma > 1$$

$$f(x) = g(x) + O(x^{1/2} \log x).$$

If we know G we may be able to estimate g and then f. The integral for g_1 in terms of G is better behaved at ∞, so perhaps we estimate g_1 instead. In this case, we use the Tauberian lemma, lemma 2.1, to estimate g and then f as before.

Mellin calisthenics

10 Sketches of the Proof of the PNT

The proof goes by estimating an inverse Mellin transform in accord with the Mellin calesthenics above. In the sketches below we do not concern ourselves with integrability of the Mellin transform. In those cases were integrability holds the inversion formula is valid by theorem 9.1 above. In the other cases, where the integral in the inversion formula is taken in some principal value sense, results of Dirichlet, Dini and Jordan apply.

10.1 Čebyšev function method

Since

$$-\frac{\zeta'(s)}{s\zeta(s)} = \int_0^\infty \psi(x)x^{-s-1}dx$$

we have

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

We now evaluate the contour integral to obtain a formula for $\psi(x)$, the von Mangoldt formula of 1895, [43] (see also Edwards [8]). We then use this formula to deduce

$$\int_1^x \psi(t)dt \sim \frac{x^2}{2}.$$

This part uses the fact that there are no roots on $\mathfrak{Re}\, s = 1$. Then the Tauberian lemma, lemma 2.1, yields $\psi(x) \sim x$.

Naturally there are thematic variations: Hadamard estimated $\displaystyle\int_1^x \frac{\psi(x)}{t^2}dt$ whereas de la Vallée Poussin estimated $\displaystyle\int_1^x \frac{\psi(x)}{t}dt$.

10.2 Modified Čebyšev function method

If

$$\psi_1(x) = \int_1^x \frac{\psi(t)}{t}dt$$

then we have

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

If we estimate the contour integral to obtain $\psi_1(x) \sim x$ then the Tauberian lemma, lemma 2.1, implies $\psi(x)/x \sim 1$.

10.3 Still another Čebyšev function method

Since $\zeta(s)$ is meromorphic with a simple pole at $s = 1$, with residue 1, and no other poles, we have

$$\zeta(s) = \frac{h(s)}{s-1}$$

where $h(s)$ is entire and $h(1) = 1$. Thus

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

is analytic in $\mathfrak{Re}\, s \geq 1$. Now

$$-\frac{1}{s}\left(\frac{\zeta'(s)}{\zeta(s)} + \frac{s}{s-1}\right) = \int_0^\infty (\psi(x) - x)\; x^{-s-1}\, dx$$

and so

$$\frac{\psi(x)}{x} - 1 = -\frac{1}{2\pi\mathfrak{i}} \int_{\sigma-\mathfrak{i}\infty}^{\sigma+\mathfrak{i}\infty} \left(\frac{\zeta'(s)}{\zeta(s)} + \frac{s}{s-1}\right) \frac{1}{s}\, x^{s-1}\, ds$$

for $\sigma > 1$. Since the integrand is well-behaved on $\sigma = 1$ we can try moving the contour of integration to $\sigma = 1$ to estimate $1 - \psi(x)/x$. This is (roughly) the approach taken in Heins [19].

10.4 Yet another Čebyšev function method

We estimate the integral

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

It turns out that

$$\int_1^x \psi(t)dt = \sum_{n\leq x} \Lambda(n)(x-n) = x\Psi(x) = \frac{x^2}{2} + o(x^2)$$

and we use the Tauberian lemma as before. For this proof see Rademacher [31].

10.5 Riemann's method

Riemann did not actually prove the prime number theorem, but he did have the ingredients described here.

Since

$$\frac{\log\zeta(s)}{s} = \int_0^\infty \frac{\pi(x)}{x(x^s-1)} dx$$

we have

$$\frac{\log\zeta(s)}{s} = \int_0^\infty J(x) x^{-s-1} dx$$

where

$$J(x) = \sum_{n=1}^{\infty} \frac{1}{n} \pi(x^{1/n}).$$

Then

$$J(x) = \frac{1}{2\pi \mathrm{i}} \int_{\sigma - \mathrm{i}\infty}^{\sigma + \mathrm{i}\infty} \frac{\log \zeta(s)}{s} x^s ds.$$

This integral is only conditionally convergent. We estimate it very carefully and deduce $J(x) \sim x/\log x$. Then $\pi(x) = J(x) + O(x^{1/2}\log x) = x/\log x + o(x/\log x)$.

10.6 Modified Riemann method

If

$$J_1(x) = \int_1^x \frac{J(t)}{t} dt$$

then

$$J_1(x) = \frac{1}{2\pi \mathrm{i}} \int_{\sigma - \mathrm{i}\infty}^{\sigma + \mathrm{i}\infty} \frac{\log \zeta(s)}{s^2} x^s ds.$$

We estimate the integral to show $J_1(x) \sim x/\log x$. Then the Tauberian lemma, lemma 2.1, implies $J(x)/x \sim 1/\log x$. This is the method of proof used in Grosswald [11] and is the method we will use here.

10.7 Littlewood's Method

Littlewood [28] gives a proof by studying the inverse Mellin transform of

$$\Gamma(s) \frac{\zeta'(s)}{\zeta(s)}$$

and relating it to the Čebyšev ψ function

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

10.8 Ikehara Tauberian Theorem

A very famous result concerning the asymptotic behavior of the inverse Mellin (or inverse Fourier) transform is the Tauberian theorem of Landau, Hardy, Littlewood and Ikehara. It shows that a large part of the proof of the prime number theorem is valid in a more general context. See Landau [24], Hardy–Littlewood [17], Ikehara [20], Wiener [44] or Donoghue [7].

Theorem 10.1. Ikehara. *Let μ be a monotone nondecreasing function on $(0, \infty)$ and let*

$$F(s) = \int_1^\infty x^{-s} d\mu(x).$$

If the integral converges absolutely for $\mathfrak{Re}\, s > 1$ and there is a constant A such that

$$F(s) - \frac{A}{s-1}$$

extends to a continuous function in $\mathfrak{Re}\, s \geq 1$ then

$$\mu(x) \sim Ax.$$

If we apply the theorem to

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

with

$$\mu(x) = \int_1^x \frac{\psi(t)}{t}\, dt$$

then we obtain the prime number theorem.

11 Proof of the Prime Number Theorem

Recall we have seen

$$\frac{\log \zeta(s)}{s} = \int_0^\infty \frac{\pi(x)}{x(x^s - 1)}\, dx, \quad \mathfrak{Re}\, s > 1$$

and therefore if

$$J(x) = \sum_{n=1}^\infty \frac{1}{n} \pi(x^{1/n})$$

then we have a Mellin transform

$$\frac{\log \zeta(s)}{s} = \int_0^\infty J(x) x^{-s-1}\, dx, \quad \mathfrak{Re}\, s > 1.$$

This transform decays only as $t^{-1}(\log t)^9$ for large $|\, t\,|$ and therefore a direct estimate of $J(x)$ from the inversion formula would require a delicate analysis of a conditionally convergent integral If we let

$$J_1(x) = \int_1^x \frac{J(t)}{t}\, dt$$

(note $J(t) = 0$ if $t < 2$) then an integration by parts yields the Mellin transform

$$\frac{\log \zeta(s)}{s^2} = \int_0^\infty J_1(x) x^{-s-1}\, dx, \quad \Re\mathfrak{e}\, s > 1.$$

This transform (for $s = \sigma + \mathrm{i}t$, $\sigma > 1$ fixed) decays like $t^{-2}(\log t)^9$ for large $|\,t\,|$ and so is integrable along vertical lines. As we have seen it follows that the inverse transform formula is valid. Thus

$$J_1(x) = \frac{1}{2\pi\mathrm{i}} \int_{a-\mathrm{i}\infty}^{a+\mathrm{i}\infty} \frac{\log \zeta(s)}{s^2} x^s\, ds$$

for any $a > 1$. We will estimate this contour integral (essentially by moving the contour to $a = 1$).

Let $h(s) = (s-1)\zeta(s)$ so $h(s)$ is an entire analytic function and $h(s) \neq 0$ if $\Re\mathfrak{e}\, s \geq 1$. For $\sigma > 1$ we have

$$\log h(s) = \log(\sigma - 1) + \log \zeta(s)$$

where the principal branch of the logarithm is used. We know $\log \zeta(s)$ is analytic in $\Re\mathfrak{e} s \geq 1$, $s \neq 1$. Thus by uniqueness of analytic continuation we have

$$\log h(s) = \log(s-1) + \log \zeta(s)$$

for $\Re\mathfrak{e}\, s \geq 1$, $s \neq 1$. It follows that $J_1(x) = I_1(x) - H_2(x)$ where

$$I_1(x) = \frac{1}{2\pi\mathrm{i}} \int_{a-\mathrm{i}\infty}^{a+\mathrm{i}\infty} \frac{\log h(s)}{s^2} x^s\, ds \tag{32}$$

$$H_2(x) = \frac{1}{2\pi\mathrm{i}} \int_{a-\mathrm{i}\infty}^{a+\mathrm{i}\infty} \frac{\log(s-1)}{s^2} x^s\, ds. \tag{33}$$

$$\tag{34}$$

The estimate of $I_1(x)$ will depend on $\zeta(s) \neq 0$ when $s = 1 + \mathrm{i}t$. The estimate of $H_2(x)$, which is harder, has nothing to do with zeta functions.

Lemma 11.1. *Let*

$$H_m(x) = \frac{1}{2\pi\mathrm{i}} \int_{a-\mathrm{i}\infty}^{a+\mathrm{i}\infty} \frac{\log(s-1)}{s^m} x^s\, ds,$$

$a > 1$ and $m > 1$. Then $H_m(x)$ is independent of a and

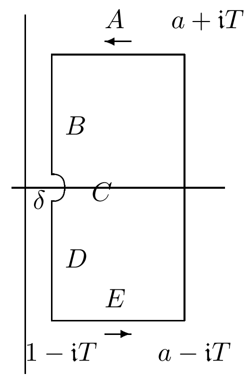
$$H_m(x) = \frac{x}{\log x} + o\left(\frac{x}{\log x}\right).$$

Proof. Consider the contour illustrated at the right. By Cauchy's theorem we have

$$\int_{\sigma-\mathrm{i}T}^{\sigma+\mathrm{i}T} \frac{\log(s-1)}{s^m}\, x^s\, ds = -IA - IB - IC - ID - IE$$

where IX indicates the integral over the segment labeled X in the illustration. Consider first the integral over the segment A:

$$-IA = \int_{1+\mathrm{i}T}^{a+\mathrm{i}T} \frac{\log(s-1)}{s^m}\, x^s\, ds.$$



Along the segment A the integrand is bounded in absolute value by $C_a \dfrac{\log T}{T^m}$ if $T > 2$. The path of integration has length $a-1$. We have the same estimate for IE. Since $m > 0$ we see $IA \to 0$ and $IE \to 0$ as $T \to \infty$. The integral over the semicircle of radius $\delta < 1$ is given by

$$-IC = \int_{-\pi/2}^{\pi/2} \frac{\log(\delta \mathrm{e}^{\mathrm{i}\theta})}{(1+\delta\mathrm{e}^{\mathrm{i}\theta})^m}\, x^{1+\delta\exp(\mathrm{i}\theta)}\, \mathrm{i}\delta\mathrm{e}^{\mathrm{i}\theta}\, d\theta.$$

Thus

$$\begin{aligned}
|\, IC \,| &\le \int_{-\pi/2}^{\pi/2} \frac{\left|\log(\delta\mathrm{e}^{\mathrm{i}\theta})\right|}{(1-\delta)^m}\, x^{1+\delta}\, \delta\, d\theta \\
&\le x^{1+\delta} \frac{\delta}{(1-\delta)^m} \int_{-\pi/2}^{\pi/2} (-\log\delta + \pi^2/4)\, d\theta \\
&\le \frac{\pi x^{1+\delta}}{(1-\delta)^m} (-\delta\log\delta + \delta\pi^2/4) \\
&\to 0 \quad \text{as } \delta \to 0.
\end{aligned}$$

Combining the integrals IB and ID we now have

$$\begin{aligned}
H_m(x) &= -\lim_{\delta\to 0} \lim_{T\to\infty} \frac{1}{\pi} \int_\delta^T \Re\mathfrak{e} \left(\frac{\log(\mathrm{i}t)}{(1+\mathrm{i}t)^m}\, x^{1+\mathrm{i}t} \right) dt \\
&= -\frac{x}{\pi}\, \Re\mathfrak{e}\, \widehat{G}_m(\log x)
\end{aligned}$$

where

$$\widehat{G}_m(y) = \int_0^\infty \frac{\log(\mathrm{i}t)}{(1+\mathrm{i}t)^m}\, \mathrm{e}^{\mathrm{i}ty}\, dt$$

is the Fourier transform of the integrable function

$$G_m(t) = \frac{\log(\mathrm{i}t)}{(1+\mathrm{i}t)^m}\, Y(t)$$

where Y is Heaviside's function. By the Riemann–Lebesgue lemma it follows that $\widehat{G_m}(y) \to 0$ as $y \to \infty$, but we need more. In fact we need

$$\mathfrak{Re}\left(y\widehat{G}_m(y)\right) = -\pi + o(1), \quad y \to \infty. \tag{35}$$

The traditional estimate of Fourier integral goes by integration by parts:

$$\begin{aligned}
\mathrm{i}y\widehat{G}_m(y) &= \lim_{\delta\to 0}\lim_{T\to\infty}\int_\delta^T \frac{\log(\mathrm{i}t)}{(1+\mathrm{i}t)^m}\frac{d}{dt}\mathrm{e}^{\mathrm{i}ty}\,dt \\
&= \lim_{\delta\to 0}\lim_{T\to\infty}\left(\frac{\log(\mathrm{i}T)}{(1+\mathrm{i}T)^m}\mathrm{e}^{\mathrm{i}Ty}\frac{\log(\mathrm{i}\delta)}{(1+\mathrm{i}\delta)^m}\mathrm{e}^{\mathrm{i}\delta y} - \right. \\
&\qquad \left. \int_\delta^T \frac{\mathrm{e}^{\mathrm{i}ty}}{t(1+\mathrm{i}t)^m}dt - \int_\delta^T \frac{m\pi/2 - m\mathrm{i}\log t}{(1+\mathrm{i}t)^{m+1}}\,\mathrm{e}^{\mathrm{i}ty}\,dt\right).
\end{aligned}$$

The first term is $O(\frac{\log T}{T^m})$. In the last integral the integrand is $O(\frac{\log t}{t^{m+1}})$ for large t and $O(-\log t)$ for small t. It follows the integrand is integrable, the limit of the last integral exists and this limit is the Fourier transform of an integrable function, and so has limit 0 at ∞ by the Riemann-Lebesgue lemma. We now have

$$\mathrm{i}y\widehat{G}_m(y) = -\lim_{\delta\to 0}\left(\frac{\log(\mathrm{i}\delta)}{(1+\mathrm{i}\delta)^m}\,\mathrm{e}^{\mathrm{i}\delta y} + \int_\delta^\infty \frac{\mathrm{e}^{\mathrm{i}ty}}{t(1+\mathrm{i}t)^{m+1}}\,dt\right) + o(1), \tag{36}$$

as $y \to \infty$.

Now let

$$u(t) = \frac{1}{t(1+\mathrm{i}t)^{m+1}} - \frac{1}{t} + \frac{\mathrm{i}^{m+1}t^m}{(1+\mathrm{i}t)^{m+1}}.$$

A quick calculation shows

$$\begin{aligned}
u(t) &= -(m+1)\mathrm{i} - \frac{(m+1)(m+2)}{2}t + O(t^2), \quad t \to 0 \\
u(t) &= (m+1)\mathrm{i}t^{-2} + \frac{(m+1)(m+2)}{2}t^{-3} + O(t^{-4}), \quad t \to \infty.
\end{aligned}$$

It follows that u is integrable and therefore by the Riemann–Lebesgue lemma we now have

$$\begin{aligned}
\mathrm{i}y\widehat{G}_m(y) &= o(1) - \lim_{\delta\to 0}\lim_{T\to\infty}\left(\frac{\log(\mathrm{i}\delta)}{(1+\mathrm{i}\delta)^m}\mathrm{e}^{\mathrm{i}\delta y} + \int_\delta^T \frac{\mathrm{e}^{\mathrm{i}ty}}{t}\,dt\right. \\
&- \left.\mathrm{i}^{m+1}\int_\delta^T \frac{t^m\mathrm{e}^{\mathrm{i}ty}}{(1+\mathrm{i}t)^{m+1}}\,dt\right)
\end{aligned}$$

We integrate by parts again

$$\begin{aligned}
& \mathrm{i}y \int_\delta^T \frac{t^m \mathrm{e}^{\mathrm{i}ty}}{(1+\mathrm{i}t)^{m+1}}\, dt \\
= & \int_\delta^T \frac{t^m}{(1+\mathrm{i}t)^{m+1}} \frac{d}{dt}\mathrm{e}^{\mathrm{i}ty}\, dt \\
= & \frac{t^m}{(1+\mathrm{i}t)^{m+1}}\, \mathrm{e}^{\mathrm{i}ty}\Big|_\delta^T - \int_\delta^T \frac{mt^{m-1} - \mathrm{i}t^m}{(1+\mathrm{i}t)^{m+2}}\, \mathrm{e}^{\mathrm{i}ty}\, dt.
\end{aligned}$$

It follows that

$$\mathrm{i}y \int_0^\infty \frac{t^m \mathrm{e}^{\mathrm{i}ty}}{(1+\mathrm{i}t)^{m+1}}\, dt = -\int_0^\infty \frac{mt^{m-1} - \mathrm{i}t^m}{(1+\mathrm{i}t)^{m+2}}\, \mathrm{e}^{\mathrm{i}ty}\, dt.$$

Now $(mt^{m-1} - \mathrm{i}t^m)/(1+\mathrm{i}t)^{m+2}$ is integrable and so by the Riemann–Lebesgue lemma this last integral is $o(1)$ as $y \to \infty$. Thus

$$\begin{aligned}
\mathrm{i}y\widehat{G}_m(y) &= o(1) - \lim_{\delta\to 0}\left(\frac{\log(\mathrm{i}\delta)}{(1+\mathrm{i}\delta)^m}\, \mathrm{e}^{\mathrm{i}\delta y} + \int_\delta^\infty \frac{\mathrm{e}^{\mathrm{i}ty}}{t}\, dt\right) \\
&= o(1) - \lim_{\delta\to 0}\left(\log(\mathrm{i}\delta)\mathrm{e}^{\mathrm{i}\delta y} + \int_\delta^\infty \frac{\mathrm{e}^{\mathrm{i}ty}}{t}\, dt\right).
\end{aligned}$$

Since $\log(\mathrm{i}\delta) = \log\delta + \mathrm{i}\pi/2$ if we take imaginary parts we obtain

$$\begin{aligned}
\mathfrak{Re}\left(y\widehat{G}_m(y)\right) &= o(1) - \lim_{\delta\to 0}\left((\log\delta)(\sin\delta y) + \frac{\pi}{2}\cos\delta y + \int_\delta^\infty \frac{\sin ty}{t}\, dt\right) \\
&= o(1) + \frac{\pi}{2} + \int_0^\infty \frac{\sin ty}{t}\, dt.
\end{aligned}$$

Since the integral is $\pi/2$ we have

$$\mathfrak{Re}\left(\widehat{G}_m(y)\right) = \frac{\pi}{y} + o\left(\frac{1}{y}\right)$$

as required. □

To prove the prime number theorem it remains to show if

$$I_1(x) = \frac{1}{2\pi\mathrm{i}} \int_{a-\mathrm{i}\infty}^{a+\mathrm{i}\infty} \frac{\log h(s)}{s^2}\, x^s\, ds, \quad a > 1,$$

then

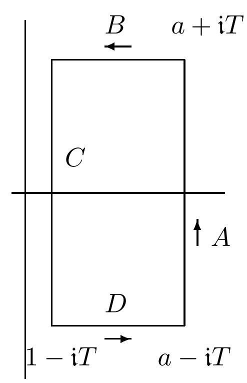
$$I_1(x) = o\left(\frac{x}{\log x}\right), \quad x \to \infty.$$

The critical issue here is that $\log h(s)$, where $h(s) = (s-1)\zeta(s)$, is analytic in $\mathfrak{Re}\, s \geq 1$.

By Cauchy's theorem

$$I_1(x) = \lim_{T\to\infty} IA = \lim_{T\to\infty} -IB - IC - ID$$

where IX indicates the contour integral over the segment labelled X in the illustration in the right. To estimate the integrals we note



$$x^s\, s^{-2}\, \log h(s)$$

is analytic in $\mathfrak{Re}\, s \geq 1$ and, by theorem 8.10, for $|\,t\,|$ large and $s = \sigma + \mathrm{i}t$ we have

$$\begin{aligned}
\left| x^s\, s^{-2}\, \log h(s) \right| &= \left| \frac{x^{\sigma+\mathrm{i}t}}{(\sigma+\mathrm{i}t)^2}\, \log\left((s-1)\zeta(s)\right) \right| \\
&\leq \frac{x^\sigma}{t^2}\, |\log|\,t\,| + \log|\,\zeta(s)\,|\,| \\
&\leq C\, \frac{x^\sigma}{t^2} (\log|\,t\,|)^9.
\end{aligned}$$

It follows the integral over the horizontal segments goes to 0 as $T \to \infty$. Thus

$$\begin{aligned}
I_1(x) &= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{x^{1+\mathrm{i}t}}{(1+\mathrm{i}t)^2}\, \log h(1+\mathrm{i}t)\, dt \\
&= \frac{x}{2\pi} H(\log x)
\end{aligned}$$

where

$$H(y) = \lim_{T\to\infty} \int_{-T}^{T} \mathrm{e}^{\mathrm{i}yt}\, \frac{\log h(1+\mathrm{i}t)}{(1+\mathrm{i}t)^2}\, dt.$$

Integrating by parts we have

$$\mathrm{i}yH(y) = \lim_{T\to\infty} \int_{-T}^{T} \frac{\log h(1+\mathrm{i}t)}{(1+\mathrm{i}t)^2}\, \frac{d}{dt} \mathrm{e}^{\mathrm{i}yt}\, dt \tag{37}$$

$$= -\mathrm{i} \lim_{T\to\infty} \int_{-T}^{T} \frac{(1+\mathrm{i}t)\frac{h'(1+\mathrm{i}t)}{h(1+\mathrm{i}t)} - 2\, \log h(1+\mathrm{i}t)}{(1+\mathrm{i}t)^3}\, \mathrm{e}^{\mathrm{i}yt}\, dt \tag{38}$$

(39)

since $\log h(1+\mathrm{i}t) = O((\log|\,t\,|)^9)$ implies the boundary terms have limit 0.

We have already seen that $\log h(s)$ is analytic in $\mathfrak{Re}\, s \geq 1$. Hence the same is true for $h'(s)/h(s)$. We can also see it directly – since $\zeta(s)$ has a simple pole at $s = 1$ it follows that $\zeta'(s)/\zeta(s)$ has a simple pole with residue -1 at $s = 1$. Thus $h(s) = (s-1)\zeta(s)$ implies

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

is analytic in a neighborhood of $s = 1$ and also away from the zeros of $\zeta(s)$. Thus its analytic in $\mathfrak{Re}\, s \geq 1$. It follows that

$$\left| \frac{h'(1+\mathrm{i}t)}{h(1+\mathrm{i}t)} \right| \leq C \quad \text{for } |\,t\,| \leq 1.$$

From lemma 8.6 we have

$$\begin{aligned} \left| \frac{h'(1+\mathrm{i}t)}{h(1+\mathrm{i}t)} \right| &= \left| \frac{\zeta'(1+\mathrm{i}t)}{\zeta(1+\mathrm{i}t)} + \frac{1}{\mathrm{i}t} \right| \\ &\leq C_1 + c_2 (\log |\,t\,|)^9 \\ &\leq C_3 (\log |\,t\,|)^9 \end{aligned}$$

for $|\,t\,| \geq 1$. It follows that the numerator in the last integrand in equation (37) is bounded by

$$C_3\ |\,t\,| (\log |\,t\,|)^9 + C_4 (\log |\,t\,|)^9.$$

Hence the integrand is

$$O\left(\left| t^{-2} \right| (\log |\,t\,|)^9 \right), \quad |\,t\,| \to \infty$$

and bounded for small $|\,t\,|$. Thus the integrand is integrable and so by the Riemann–Lebesgue lemma we have

$$yH(y) = o(1), \quad \text{as } |\,t\,| \to \infty.$$

Then

$$I_1(x) = \frac{x}{2\pi} H(\log x) = o\left(\frac{x}{\log x} \right)$$

which completes the proof of the prime number theorem.

References

[1] George E. Andrews, *Number theory*, W. B. Saunders Co., Philadelphia, 1971, (Reprinted by Dover, New York, 1994).

[2] Bruce C. Berndt and Robert A. Rankin, *Ramanujan. Letters and commentary*, Amer. Math. Soc. and London Math. Soc., Providence, Rhode Island, 1995.

[3] P. L. Čebyšev, *La totalité des nombres premiers inférieurs a une limite donnée*, J. Math. Pures Appl. **17** (1852), 341–365.

[4] ———, *Œuvres*, Chelsea, New York, 1962, (Reprint).

[5] Richard E. Crandall, *Projects in scientific computation*, Springer–Verlag, New York, Berlin, Heidelberg, 1994.

[6] Charles-Jean de la Vallée Poussin, *Recherches analytiques sur la théorie des nombres; Première partie: La fonction $\zeta(s)$ de Riemann et les nombres premiers en général*, Annales de la Soc. scientifiques de Bruxelles **20 (II)** (1896), 183–256 (Second part 281–297 ??).

[7] William F. Donoghue, *Distributions and Fourier transforms*, Academic Press, New York and London, 1969.

[8] H. M. Edwards, *Riemann's zeta function*, Academic Press, New York, London, 1974.

[9] Carl Friedrich Gauss, *Werke*, 2 ed., Königlichen Gesellschaft der Wissenschaften, Göttingen, 1876, (First edition 1863).

[10] L. J. Goldstein, *A history of the prime number theorem*, Amer. Math. Monthly **80** (1973), 599–615, (Includes Gauss' 1849 letter to Enke, translated by G. Ehrlich).

[11] Emil Grosswald, *Topics from the theory of numbers*, 2 ed., Birkhäuser, Boston, Basel, Stuttgart, 1984.

[12] Jacques Hadamard, *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bull. Soc. Math. France **14** (1896), 199–220, (Also in "Œuvres" Vol. 1).

[13] ———, *Œuvres de Jacques Hadamard*, Centre National de la Recherche Scientifiques, Paris, 1968.

[14] G. H. Hardy, *Prime numbers*, British Association Report, Britist Association, Manchester, 1915, (In "Collected Papers", vol. 2), pp. 350–354.

[15] ———, *On Mellin's inversion formula*, Messenger of Mathematics **47** (1918), 178–184, (In "Collected Papers", vol. 7).

[16] ———, *Further notes on Mellin's inversion formulæ*, Messenger of Mathematics **50** (1921), 165–171, (In "Collected Papers", vol. 7).

[17] G. H. Hardy and J. E. Littlewood, *Contributions to the theory of the Riemann zetafunction and the theory of the distribution of primes*, Acta Math. **41** (1918), 119–196.

[18] Philip Hartman and Aurel Wintner, *On Möbius' inversion*, Amer. J. Math **69** (1947), 853–858.

[19] Maurice Heins, *Complex function theory*, Academic Press, New York and London, 1968.

[20] S. Ikehara, *An extension of Landau's theorem in the analytic theory of numbers*, Journ. Math. Phys. M.I.T. **10** (1931), 1–12.

[21] A. E. Ingham, *The distribution of prime numbers*, Cambridge Tracts in Mathematics and Mathematical Physics, no. 30, Cambridge Univ. Press, London, 1932.

[22] Aleksandar Ivić, *The Riemann zeta–function. The theory of the riemann zeta–function with applications*, John Wiley & Sons, New York, 1985.

[23] H. Kestelman, *Modern theories of integration*, 2nd ed., Dover Publications, New York, 1960 (Orig. Oxford Univ. Press 1937).

[24] E. Landau, *Über dei Konvergenz einiger Klassen von unendlichen Reihen am Rande des Konvergenzgebietes*, Monatsh. f. Math. und Phys. **18** (1907), 8–28.

[25] R. Sherman Lehman, *On the difference* $\pi(x) - \mathrm{li}(x)$, Acta Aritmetica **11** (1966), 397–410.

[26] D. N. Lehmer, *List of prime numbers from 1 to 10,006,721. Publ. no. 165*, Carnegie Inst. of Washington, Washington, D. C., 1913, (Reprinted, Hafner, New York, 1956).

[27] J. E. Littlewood, *Sur la distribution des nombres premiers*, C. R. Acad. Sci. Paris **158** (1914), 263–266, (In "Collected Papers", Vol. 2).

[28] ———, *The quickest proof of the prime number theorem*, Acta Aritmetica **18** (1971), 83–86, (In "Collected Papers", Vol. 2).

[29] H. Mellin, *Abriß einer einheitlichen Theorie der Gamma– und der hypergeometrischen Funktionen*, Math. Ann **68** (1910), 305–337.

[30] S. J. Patterson, *An introduction to the theory of the Riemann zeta–function*, Cambridge Univ. Press, Cambridge New York Melbourne, 1988.

[31] Hans Rademacher, *Topics in analytic number theory*, Springer–Verlag, Berlin, Heidelberg, New York, 1973.

[32] H. Reisel, *Prime numbers and computer methods for factorization*, Birkhauser, Boston, Basel, Stuttgart, 1985.

[33] Bernhard Riemann, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Berliner Akademie (1859), 671–680, (Also in "Gesammelte Werke").

[34] ———, *Gesammelte werke*, Teubner, Leipzig, 1892, (Reprinted by Dover, New York, 1953).

[35] S. Skewes, *On the difference* $\pi(x) - \operatorname{li} x$ *(I)*, Jour. London Math. Soc. **8** (1933), 277–283.

[36] ———, *On the difference* $\pi(x) - \operatorname{li} x$ *(II)*, Proc. London Math. Soc. **5** (1955), 48–70.

[37] Herman J. J. te Riele, *On the sign of the difference* $\pi(x) - \operatorname{li}(x)$, Mathematics of Computation **48** (1987), 323–328.

[38] E. C. Titchmarsh, *The zeta–function of Riemann*, Cambridge Tracts in Mathematics and Mathematical Physics, no. 26, Cambridge Univ. Press, London, 1930.

[39] ———, *The theory of functions*, Oxford Univ. Press, Oxford, 1932.

[40] ———, *Introduction to the theory of Fourier integrals*, Oxford Univ. Press, Oxford, 1937.

[41] ———, *The theory of functions*, 2 ed., Oxford Univ. Press, Oxford, 1939.

[42] ———, *The theory of the Riemann zeta–function*, Oxford Univ. Press, London, Glascow, New York, 1951.

[43] H. von Mangoldt, *Zu Riemann's Abhandlung 'Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse'*, J. Reine Angew. Math. **114** (1895), 255–305.

[44] Norbert Wiener, *The Fourier integral and certain of its applications*, Dover, New York, 1933 (Original edition Cambridge Univ Press).

[45] Don Zagier, *The first 50 million prime numbers*, The Mathematical Intelligencer **0** (1977), 7–19, (Sample volume initiating the Intelligencer).

Bent E. Petersen
Department of Mathematics
Oregon State University
Corvallis, OR 97331-4605

24 hour phone numbers
office (541) 737-5163
home (541) 754-2315
fax (541) 737-0517

petersen@math.orst.edu
http://www.orst.edu/~peterseb