16 Riemann's zeta function and the prime number theorem

We now divert our attention from algebraic number theory to talk about zeta functions and L-functions. As we shall see, every global field has a zeta function that is intimately related to the distribution of its primes. We begin with the zeta function of the rational field $\mathbb{Q}$, which we will use to prove the prime number theorem.

We will need some basic results from complex analysis, all of which can be found in any introductory textbook (such as [1, 2, 3, 7, 12]). A short glossary of terms and a list of the basic theorems we will use can be found at the end of these notes.[1]

16.1 The Riemann zeta function

Definition 16.1. The *Riemann zeta function* is the complex function defined by the series

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

for $\mathrm{Re}(s) > 1$, where n varies over positive integers. It is easy to see that this series converges absolutely and locally uniformly for $\mathrm{Re}(s) > 1$, thus by Theorem 16.17, it defines a holomorphic function on $\mathrm{Re}(s) > 1$, since each of term $n^{-s} = e^{-s\log n}$ is holomorphic in this region (and on the entire complex plane).

Theorem 16.2 (Euler product). *For* $\mathrm{Re}(s) > 1$ *we have*

$$\zeta(s) = \sum_{n\geq 1} n^{-s} = \prod_p (1-p^{-s})^{-1},$$

where the product converges absolutely. In particular, $\zeta(s) \neq 0$ *for* $\mathrm{Re}(s) > 1$.

The product in the theorem above ranges over primes p. This is a standard practice in analytic number theory that we will follow: the symbol p always denotes a prime, and any sum or product over p is understood to be over primes, even if this is not explicitly stated.

Proof. The one-line proof is that unique factorization and absolute convergence imply

$$\sum_{n\geq 1} n^{-s} = \sum_{e_2,e_3,\ldots\geq 0} (2^{e_2}3^{e_3}\cdots)^{-s} = \prod_p \sum_{e\geq 0} p^{-es} = \prod_p (1-p^{-s})^{-1}.$$

However the middle equality deserves some justification, and we should verify that the infinite product is absolutely convergent.

For each integer $m \geq 1$, let S_m be the set of m-smooth numbers: positive integers with prime factors $p \leq m$. Now define

$$\zeta_m(s) := \sum_{n\in S_m} n^{-s},$$

which converges absolutely and locally uniformly on $\mathrm{Re}(s) > 1$. If $p_1, \ldots, p_k$ are the primes up to m, then we may write the absolutely convergent sum as

$$\zeta_m(s) = \sum_{n\in S_m} n^{-s} = \sum_{e_1,\ldots,e_k\geq 0} (p_1^{e_1}\cdots p_k^{e_k})^{-s} = \sum_{e_1\geq 0} p_1^{-e_1 s} \sum_{e_2\geq 0} p_2^{-e_2 s} \cdots \sum_{e_k\geq 0} p_k^{-e_k s}.$$

[1] Those familiar with this material should still peruse §16.3.2 which touches on some convergence issues particularly relevant to number theoretic applications.

Lecture by Andrew Sutherland

For $\mathrm{Re}(s) > 1$ we have $\sum_{e\ge 0} p^{-es} = 1 + p^{-s} + p^{-2s} + \cdots = (1-p^{-s})^{-1}$, for any prime p. Applying this k times yields the finite product

$$\zeta_m(s) = \prod_{p\le m}(1-p^{-s})^{-1}.$$

We now note that for any $\delta > 0$ the sequence of functions $\zeta_m(s)$ converges uniformly on $\mathrm{Re}(s) > 1+\delta$ to $\zeta(s)$; indeed, for any $\epsilon > 0$ and any such s we have

$$|\zeta_m(s) - \zeta(s)| \le \left|\sum_{n\ge m} n^{-s}\right| \le \sum_{n\ge m} |n^{-s}| = \sum_{n\ge m} n^{-\mathrm{Re}(s)} \le \int_m^\infty x^{-1-\delta}dx \le \frac{1}{\delta} m^{-\delta} < \epsilon$$

for all sufficiently large m. It follows that the sequence $\zeta_m(s)$ converges locally uniformly to $\zeta(s)$ on $\mathrm{Re}(s) > 1$, and therefore the sequence of functions $P_m(s) := \prod_{p\le m}(1-p^{-s})^{-1}$ does as well. The sequence $(\log P_m)$ converges locally uniformly to $\log \prod_p (1-p^{-s})^{-1}$, and

$$\sum_p |\log(1-p^{-s})^{-1}| = \sum_p \left|\sum_{e\ge 1} \frac{1}{e} p^{-es}\right| \le \sum_p \sum_{e\ge 1} |p^{-s}|^e = \sum_p (|p^s|-1)^{-1} < \infty$$

is absolutely convergent (hence finite), thus $\prod_p (1-p^{-s})^{-1}$ is absolutely convergent (hence nonzero); here we have used the identity $\log(1-z) = -\sum_{n\ge 1} z^n$, valid for $|z| < 1$. □

Theorem 16.3 (Analytic continuation I)**.** *For* $\mathrm{Re}(s) > 1$ *we have*

$$\zeta(s) = \frac{1}{s-1} + \phi(s),$$

where $\phi(s)$ *is a holomorphic function on* $\mathrm{Re}(s) > 0$. *Thus* $\zeta(s)$ *extends to a meromorphic function on* $\mathrm{Re}(s) > 0$ *that has a simple pole at* $s = 1$ *with residue* 1 *and no other poles.*

Proof. For $\mathrm{Re}(s) > 1$ we have

$$\zeta(s) - \frac{1}{s-1} = \sum_{n\ge 1} n^{-s} - \int_1^\infty x^{-s}dx = \sum_{n\ge 1}\left(n^{-s} - \int_n^{n+1} x^{-s}dx\right) = \sum_{n\ge 1}\int_n^{n+1}\left(n^{-s} - x^{-s}\right)dx.$$

For each $n \ge 1$ the function $\phi_n(s) := \int_n^{n+1}(n^{-s} - x^{-s})dx$ is holomorphic on $\mathrm{Re}(s) > 0$. For each fixed s in $\mathrm{Re}(s) > 0$ and $x \in [n, n+1]$ we have

$$|n^{-s} - x^{-s}| = \left|\int_n^x st^{-s-1}dt\right| \le \int_n^x \frac{|s|}{|t^{s+1}|} = \int_n^x \frac{|s|}{t^{1+\mathrm{Re}(s)}}dt \le \frac{|s|}{n^{1+\mathrm{Re}(s)}},$$

and therefore

$$|\phi_n(s)| \le \int_n^{n+1} \left|n^{-s} - x^{-s}\right| ds \le \frac{|s|}{n^{1+\mathrm{Re}(s)}}.$$

For any s_0 with $\mathrm{Re}(s_0) > 0$, if we put $\epsilon := \mathrm{Re}(s_0)/2$ and $U := B_{<\epsilon}(s_0)$, then for each $n \ge 1$,

$$\sup_{s\in U} |\phi_n(s)| \le M_n := \frac{|s_0| + \epsilon}{n^{1+\epsilon}},$$

and $\sum_n M_n = (|s_0| + \epsilon)\zeta(1+\epsilon)$ converges. The series $\sum_n \phi_n$ thus converges locally normally on $\mathrm{Re}(s) > 0$. By the Weierstrass M-test (Theorem 16.19), $\sum_n \phi_n$ converges to a function $\phi(s) = \zeta(s) - \frac{1}{s-1}$ holomorphic on $\mathrm{Re}(s) > 0$. □

We now show that $\zeta(s)$ has no zeros on $\operatorname{Re}(s) = 1$; this fact is crucial to the prime number theorem. For this we use the following ingenious lemma, attributed to Mertens.[2]

Lemma 16.4 (Mertens)**.** *For $x, y \in \mathbb{R}$ with $x > 1$ we have $|\zeta(x)^3\zeta(x+iy)^4\zeta(x+2iy)| \ge 1$.*

Proof. From the Euler product $\zeta(s) = \prod_p (1-p^{-s})^{-1}$, we see that for $\operatorname{Re}(s) > 1$ we have

$$\log|\zeta(s)| = -\sum_p \log|1-p^{-s}| = -\sum_p \operatorname{Re}\log(1-p^{-s}) = \sum_p \sum_{n\ge1} \frac{\operatorname{Re}(p^{-ns})}{n},$$

since $\log|z| = \operatorname{Re}\log z$ and $\log(1-z) = -\sum_{n\ge1}\frac{z^n}{n}$ for $|z| < 1$. Plugging in $s = x+iy$ yields

$$\log|\zeta(x+iy)| = \sum_p \sum_{n\ge1} \frac{\cos(ny\log p)}{np^{nx}},$$

since $\operatorname{Re}(p^{-ns}) = p^{-nx}\operatorname{Re}(e^{-iny\log p}) = p^{-nx}\cos(-ny\log p) = p^{-nx}\cos(ny\log p)$. Thus

$$\log|\zeta(x)^3\zeta(x+iy)^4\zeta(x+2iy)| = \sum_p \sum_{n\ge1} \frac{3 + 4\cos(ny\log p) + \cos(2ny\log p)}{np^{nx}}.$$

We now note that the trigonometric identity $\cos(2\theta) = 2\cos^2\theta - 1$ implies

$$3 + 4\cos\theta + \cos(2\theta) = 2(1+\cos\theta)^2 \ge 0,$$

Taking $\theta = ny\log p$ yields $\log|\zeta(x)^3\zeta(x+iy)^4\zeta(x+2iy)| \ge 0$, which proves the lemma. □

Corollary 16.5. *$\zeta(s)$ has no zeros on $\operatorname{Re}(s) \ge 1$.*

Proof. We know from Theorem 16.2 that $\zeta(s)$ has no zeros on $\operatorname{Re}(s) > 1$, so suppose $\zeta(1+iy) = 0$ for some $y \in \mathbb{R}$. Then $y \ne 0$, since $\zeta(s)$ has a pole at $s = 1$, and we know that $\zeta(s)$ does not have a pole at $1 + 2iy \ne 1$, by Theorem 16.3. We therefore must have

$$\lim_{x\to1} |\zeta(x)^3\zeta(x+iy)^4\zeta(x+2iy)| = 0, \tag{1}$$

since $\zeta(s)$ has a simple pole at $s = 1$, a zero at $1 + iy$, and no pole at $1 + 2iy$. But this contradicts Lemma 16.4. □

16.2 The Prime Number Theorem

The prime counting function $\pi\colon \mathbb{R} \to \mathbb{Z}_{\ge0}$ is defined by

$$\pi(x) := \sum_{p\le x} 1;$$

it counts the number of primes up to x. The prime number theorem (PNT) states that

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

[2] If this lemma strikes you as pulling a rabbit out of a hat, well, it is. For a slight variation, see [15, IV], which uses an alternative approach due to Hadamard.

Here the notation $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$, and one says that f is *asymptotic to*; in other words, the functions f and g grow at the same rate, asymptotically.

This conjectured growth rate for $\pi(x)$ dates back to Gauss and Legendre in the late 18th century. In fact Gauss believed the asymptotically equivalent but more accurate statement[3]

$$\pi(x) \sim \mathrm{Li}(x) := \int_2^\infty \frac{dx}{\log x}.$$

However it was not until a century later that the prime number theorem was independently proved by Hadamard [5] and de la Vallée Poussin [9] in 1896. Their proofs are both based on the work of Riemann [10], who in 1860 showed that there is a precise connection between the zeros of $\zeta(s)$ and the distribution of primes (we shall say more about this later), but was unable to prove the prime number theorem.

The proof we will give is more recent and due to Newman [8], but it relies on the same properties of the Riemann zeta function that were exploited by both Hadamard and de la Vallée, the most essential of which is the fact that $\zeta(s)$ has no zeros on $\mathrm{Re}(s) \geq 1$ (Corollary 16.5). A concise version of Newman's proof by Zagier can be found in [15]; we will follow Zagier's outline but will be slightly more expansive in our presentation. We should note that there are also "elementary" proofs of the prime number theorem independently obtained by Erdös [4] and Selberg [11] in the 1940s that do not use the Riemann zeta function, but they are elementary only in the sense that they do not use complex analysis; these elementary proofs are actually much more complicated than those that use complex analysis.

Rather than work directly with $\pi(x)$, it is more convenient to work with the log-weighted prime-counting function defined by Chebyshev[4]

$$\vartheta(x) := \sum_{p\leq x} \log p,$$

whose growth rate differs from that of $\pi(x)$ by a logarithmic factor.

Theorem 16.6 (Chebyshev)**.** *We have $\pi(x) \sim \frac{x}{\log x}$ if and only $\vartheta(x) \sim x$.*

Proof. We clearly have $0 \leq \vartheta(x) \leq \pi(x)\log x$, thus

$$\frac{\vartheta(x)}{x} \leq \frac{\pi(x)\log x}{x}.$$

For every $\epsilon > 0$ we have

$$\begin{aligned}\vartheta(x) \geq \sum_{x^{1-\epsilon}<p\leq x} \log p &\geq (1-\epsilon)(\log x)\big(\pi(x) - \pi(x^{1-\epsilon})\big)\\ &\geq (1-\epsilon)(\log x)(\pi(x) - x^{1-\epsilon}),\end{aligned}$$

and therefore

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

[3]More accurate in the sense that $|\pi(x) - \mathrm{Li}(x)|$ grows more slowly than $|\pi(x) - \frac{x}{\log x}|$ as $x \to \infty$.

[4]As with most Russian names, there is no canonical way to write Chebyshev in the latin alphabet and one finds many variations in the literature; in English, the spelling Chebyshev is now the most widely used.

Thus for all $\epsilon > 0$ we have

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

The last term on the RHS tends to 0 as $x \to \infty$, and the lemma follows: by choosing ϵ sufficiently small we can make the ratios of $\vartheta(x)$ to x and $\pi(x)$ to $x/\log x$ arbitrarily close together as $x \to \infty$, and if one of them tends to 1, then so must the other. □

In view of Chebyshev's result, the prime number theorem is equivalent to the statement $\vartheta(x) \sim x$, which is what we will prove. The first step is to show that the asymptotic growth rate of $\vartheta(x)$ is at most linear in x.

Lemma 16.7 (Chebyshev)**.** *For $x \ge 1$ we have $\vartheta(x) \le (4\log 2)x$, thus $\vartheta(x) = O(x)$.*

Proof. For any integer $n \ge 1$, the binomial theorem implies

$$2^{2n} = (1+1)^{2n} = \sum_{m=0}^{2n}\binom{2n}{m} \ge \binom{2n}{n} = \frac{(2n)!}{n!n!} \ \ge \prod_{n<p\le 2n} p = \exp(\vartheta(2n) - \vartheta(n)),$$

since $(2n)!$ is divisible by every prime $p \in (n, 2n]$ but $n!$ is not divisible by any such p. Taking logarithms on both sides and reversing the inequality yields

$$\vartheta(2n) - \vartheta(n) \le 2n\log 2,$$

valid for all integers $n \ge 1$. For any integer $m \ge 1$ we have

$$\vartheta(2^m) = \sum_{n=1}^{m}\big(\vartheta(2^n) - \vartheta(2^{n-1})\big) \le \sum_{n=1}^{m} 2^n\log 2 \ \le\ 2^{m+1}\log 2.$$

For any real $x \ge 1$ we can choose an integer $m \ge 1$ so that $2^{m-1} \le x < 2^m$, and then

$$\vartheta(x) \le \vartheta(2^m) \le 2^{m+1}\log 2 = (4\log 2)2^{m-1} \le (4\log 2)x,$$

as claimed. □

In order to prove $\vartheta(x) \sim x$, we will use a general analytic criterion that is applicable to any non-decreasing real function $f(x)$.

Lemma 16.8. *Let $f\colon \mathbb{R}_{\ge 1} \to \mathbb{R}$ be a nondecreasing function. If the integral $\int_1^\infty \frac{f(t)-t}{t^2}dt$ converges then $f(x) \sim x$.*

Proof. Let $F(x) := \int_1^x \frac{f(t)-t}{t^2}dt$. The hypothesis is that $\lim_{x\to\infty} F(x)$ exists. This implies that for all $\lambda > 1$ and all $\epsilon > 0$ we must have $|F(\lambda x) - F(x)| < \epsilon$ for all sufficiently large x.

Fix $\lambda > 1$ and suppose there is an unbounded sequence (x_n) such that $f(x_n) \ge \lambda x_n$ for all $n \ge 1$. For each x_n we have

$$F(\lambda x_n) - F(x_n) = \int_{x_n}^{\lambda x_n} \frac{f(t)-t}{t^2}dt \ge \int_{x_n}^{\lambda x_n} \frac{\lambda x_n - t}{t^2}dt = \int_1^\lambda \frac{\lambda - t}{t^2}dt = c,$$

for some $c > 0$, where we used the fact that f is non-decreasing to get the middle inequality. Taking $\epsilon < c$, we have $|F(\lambda x_n) - F(x_n)| = c > \epsilon$ for arbitrarily large x_n, a contradiction. Thus $f(x) < \lambda x$ for all sufficiently large x. A similar argument shows that $f(x) > \frac{1}{\lambda}x$ for all sufficiently large x. These inequalities hold for all $\lambda > 1$, so $\lim_{x\to\infty} f(x)/x = 1$. Equivalently, $f(x) \sim x$. □

In order to show that the hypothesis of Lemma 16.8 is satisfied for $f = \vartheta$, we will work with the function $H(t) = \vartheta(e^t)e^{-t} - 1$; the change of variables $t = e^u$ shows that

$$\int_1^\infty \frac{\vartheta(t) - t}{t^2} dt \text{ converges} \qquad \Longleftrightarrow \qquad \int_0^\infty H(u) du \text{ converges .}$$

We now recall the Laplace transform.

Definition 16.9. Let $h\colon \mathbb{R}_{>0} \to \mathbb{R}$ be a piecewise continuous function. The *Laplace transform* $\mathcal{L}h$ of h is the complex function defined by

$$\mathcal{L}h(s) := \int_0^\infty e^{-st} h(t) dt,$$

which is holomorphic function $\operatorname{Re}(s) > c$ for any $c \in \mathbb{R}$ for which $h(t) = O(e^{ct})$.

The following properties of the Laplace transform are easily verified:

- $\mathcal{L}(g + h) = \mathcal{L}g + \mathcal{L}h$, and for any $a \in \mathbb{R}$ we have $\mathcal{L}(ah) = \alpha\mathcal{L}h$.
- If $h(t) = a \in \mathbb{R}$ is constant then $\mathcal{L}h(s) = \frac{a}{s}$.
- $\mathcal{L}(e^{at}h(t))(s) = \mathcal{L}(h)(s - a)$ for all $a \in \mathbb{R}$.

We now define the auxiliary function

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

which is related to $\vartheta(x)$ by the following lemma.

Lemma 16.10. $\mathcal{L}(\vartheta(e^t))(s) = \frac{\Phi(s)}{s}$ *is holomorphic on* $\operatorname{Re}(s) > 1$.

Proof. By Lemma 16.7, $\vartheta(e^t) = O(e^t)$, so $\mathcal{L}(\vartheta(e^t))$ is holomorphic on $\operatorname{Re}(s) > 1$. Let p_n be the nth prime, and put $p_0 := 0$. The function $\vartheta(e^t)$ is constant on $t \in (\log p_n, \log p_{n+1})$, so

$$\int_{\log p_n}^{\log p_{n+1}} e^{-st}\vartheta(e^t) dt = \vartheta(p_n) \int_{\log p_n}^{\log p_{n+1}} e^{-st} dt = \frac{1}{s}\vartheta(p_n)\bigl(p_n^{-s} - p_{n+1}^{-s}\bigr).$$

We then have

$$\begin{aligned}
(\mathcal{L}\vartheta(e^t))(s) = \int_0^\infty e^{-st}\vartheta(e^t) dt &= \frac{1}{s}\sum_{n=1}^\infty \vartheta(p_n)\bigl(p_n^{-s} - p_{n+1}^{-s}\bigr)\\
&= \frac{1}{s}\sum_{n=1}^\infty \vartheta(p_n)p_n^{-s} - \frac{1}{s}\sum_{n=1}^\infty \vartheta(p_{n-1})p_n^{-s}\\
&= \frac{1}{s}\sum_{n=1}^\infty \bigl(\vartheta(p_n) - \vartheta(p_{n-1})\bigr)p_n^{-s}\\
&= \frac{1}{s}\sum_{n=1}^\infty p_n^{-s}\log p_n = \frac{\Phi(s)}{s}.
\end{aligned}$$

□

Let us now consider the function $H(t) := \vartheta(e^t)e^{-t} - 1$. It follows from the lemma and standard properties of the Laplace transform that on $\operatorname{Re}(s) > 0$ we have

$$(\mathcal{L}H)(s) = \mathcal{L}(\vartheta(e^t)e^{-t})(s) - (\mathcal{L}1)(s) = \mathcal{L}(\vartheta(e^t))(s+1) - \frac{1}{s} = \frac{\Phi(s+1)}{s+1} - \frac{1}{s}.$$

Lemma 16.11. *The function* $\Phi(s) - \frac{1}{s-1}$ *extends to a meromorphic function on* $\mathrm{Re}(s) > \frac{1}{2}$ *that is holomorphic on* $\mathrm{Re}(s) \ge 1$.

Proof. By Theorem 16.3, $\zeta(s)$ extends to a meromorphic function on $\mathrm{Re}(s) > 0$ (which we also denote $\zeta(s)$) that has only a simple pole at $s = 1$ and no zeros on $\mathrm{Re}(s) \ge 1$, by Corollary 16.5. It follows that the logarithmic derivative $\zeta'(s)/\zeta(s)$ of $\zeta(s)$ is meromorphic on $\mathrm{Re}(s) > 0$, and the only pole $\zeta'(s)/\zeta(s)$ has on $\mathrm{Re}(s) \ge 1$ is a simple pole at $s = 1$ with residue -1 (see §16.3.1 for standard facts about the logarithmic derivative of a meromorphic function). In terms of the Euler product we have

$$\begin{aligned}
-\frac{\zeta'(s)}{\zeta(s)} &= (-\log\zeta(s))' = \left(-\log\prod_p (1-p^{-s})^{-1}\right)' = \left(\sum_p \log(1-p^{-s})\right)' \\
&= \sum_p \frac{p^{-s}\log p}{1-p^{-s}} \;=\; \sum_p \frac{\log p}{p^s - 1} = \sum_p \left(\frac{1}{p^s} + \frac{1}{p^s(p^s-1)}\right)\log p \\
&= \Phi(s) + \sum_p \frac{\log p}{p^s(p^s-1)}.
\end{aligned}$$

The sum on the RHS converges absolutely and locally uniformly to a holomorphic function on $\mathrm{Re}(s) > 1/2$. The LHS is meromorphic on $\mathrm{Re}(s) > 0$, and on $\mathrm{Re}(s) \ge 1$ is has only a simple pole at $s = 1$ with residue 1. It follows that $\Phi(s) - \frac{1}{s-1}$ extends to a meromorphic function on $\mathrm{Re}(s) > \frac{1}{2}$ that is holomorphic on $\mathrm{Re}(s) \ge 1$. ☐

Corollary 16.12. *The functions* $\Phi(s+1) - \frac{1}{s}$ *and* $(\mathcal{L}H)(s) = \frac{\Phi(s+1)}{s+1} - \frac{1}{s}$ *both extend to meromorphic functions on* $\mathrm{Re}(s) > -\frac{1}{2}$ *that are holomorphic on* $\mathrm{Re}(s) \ge 0$.

Proof. The first statement follows immediately from the lemma. For the second, note that

$$\frac{\Phi(s+1)}{s+1} - \frac{1}{s} = \frac{1}{s+1}\left(\Phi(s+1) - \frac{1}{s}\right) - \frac{1}{s+1}$$

is meromorphic on $\mathrm{Re}(s) > -\frac{1}{2}$ and holomorphic on $\mathrm{Re}(s) \ge 0$, since it is a sum of products of such functions. ☐

The final step of the proof relies on the following analytic result due to Newman [8].

Theorem 16.13. *Let* $f\colon \mathbb{R}_{\ge 0} \to \mathbb{R}$ *be a bounded piecewise continuous function, and suppose its Laplace transform extends to a holomorphic function* $g(s)$ *on* $\mathrm{Re}(s) \ge 0$. *Then the integral* $\int_0^\infty f(t)dt$ *converges and is equal to* $g(0)$.

Proof. Without loss of generality we assume $f(t) \le 1$ for all $t \ge 0$. For $\tau \in \mathbb{R}_{>0}$, define $g_\tau(s) := \int_0^\tau f(t)e^{-st}dt$, By definition $\int_0^\infty f(t)dt = \lim_{\tau\to\infty} g_\tau(0)$, thus it suffices to prove

$$\lim_{\tau\to\infty} g_\tau(0) = g(0).$$

For $r > 0$, let γ_r be the boundary of the region $\{s : |s| \le R \text{ and } \mathrm{Re}(s) \le \delta_r\}$ with $\delta_r > 0$ chosen so that g is holomorphic on γ_r; such a δ_r exists because g is holomorphic on $\mathrm{Re}(s) \ge 0$, hence on some open ball $B_{\le 2\delta(y)}(iy)$ for each $y \in [-r, r]$, and we may take $\delta_r := \inf\{\delta(y) : y \in [r, -r]\}$, which is positive because $[-r, r]$ is compact. Each γ_r is a simple closed curve, and for each $\tau > 0$ the function $h(s) := (g(s) - g_\tau(s))e^{-s\tau}(1 + \frac{s^2}{r^2})$ is holomorphic on a region containing γ_r. Using Cauchy's integral formula (Theorem 16.26) to evaluate $h(0)$ yields

$$g(0) - g_\tau(0) = h(0) = \frac{1}{2\pi i}\int_{\gamma_r} \big(g(s) - g_\tau(s)\big)e^{s\tau}\left(\frac{1}{s} + \frac{s}{r^2}\right) ds. \tag{2}$$

We will show the LHS tends to 0 as $\tau \to \infty$ by showing that for any $\epsilon > 0$ we can set $r = 3/\epsilon > 0$ so that the absolute value of the RHS is less than ϵ for all sufficiently large τ.

Let γ_r^+ denote the part of γ_r in $\mathrm{Re}(s) > 0$, a semicircle of radius r. The integrand is absolutely bounded by $1/r$ on γ_r^+, since for $|s| = r$ and $\mathrm{Re}(s) > 0$ we have

$$\begin{aligned}
\big|g(s) - g_\tau(s)\big| \cdot \left|e^{s\tau}\left(\frac{1}{s} + \frac{s}{r^2}\right)\right| &= \left|\int_\tau^\infty f(t)e^{-st}dt\right| \cdot \frac{e^{\mathrm{Re}(s)\tau}}{r} \cdot \left|\frac{r}{s} + \frac{s}{r}\right| \\
&\le \int_\tau^\infty e^{-\mathrm{Re}(s)t}dt \cdot \frac{e^{\mathrm{Re}(s)\tau}}{r}\frac{2\,\mathrm{Re}(s)}{r} \\
&= \left(\frac{1}{\mathrm{Re}(s)}e^{-\mathrm{Re}(s)\tau}\right)\frac{e^{\mathrm{Re}(s)\tau}}{r}\frac{2\,\mathrm{Re}(s)}{r} \\
&= 2/r^2.
\end{aligned}$$

Therefore

$$\left|\frac{1}{2\pi i}\int_{\gamma_r^+} \big(g(s) - g_\tau(z)\big)e^{z\tau}\left(\frac{1}{s} + \frac{s}{r^2}\right) ds\right| \le \frac{1}{2\pi}\cdot \pi r \cdot \frac{2}{r^2} = \frac{1}{r} \tag{3}$$

Now let γ_r^- be the part of γ_r in $\mathrm{Re}(s) < 0$, the left half of the perimeter of a rectangle of height $2r$ and width $2\delta_r$. For any fixed r, the forst term $g(s)e^{s\tau}(s^{-1} + sr^{-2})$ in the integrand of (2) tends to 0 as $\tau \to \infty$ for $\mathrm{Re}(s) < 0$ and $|s| \le r$. For the second term we note that since $g_\tau(s)$ is holomorphic on $\mathbb{C}$, it makes no difference if we instead integrate over the semicircle of radius r in $\mathrm{Re}(s) < 0$. For $|s| = r$ and $\mathrm{Re}(s) < 0$ we then have

$$\begin{aligned}
\left|g_\tau(s)e^{s\tau}\left(\frac{1}{s} + \frac{s}{r^2}\right)\right| &= \left|\int_0^\tau f(t)e^{-st}dt\right| \cdot \frac{e^{\mathrm{Re}(s)\tau}}{r} \cdot \left|\frac{r}{s} + \frac{s}{r}\right| \\
&\le \int_0^\tau e^{-\mathrm{Re}(s)t}dt \cdot \frac{e^{\mathrm{Re}(s)\tau}}{r}\frac{(-2\,\mathrm{Re}(s))}{r} \\
&= \left(1 - \frac{1}{\mathrm{Re}(s)}e^{-\mathrm{Re}(s)\tau}\right)\frac{e^{\mathrm{Re}(s)\tau}}{r}\frac{(-2\,\mathrm{Re}(s))}{r} \\
&= 2/r^2 \cdot (1 - e^{\mathrm{Re}(s)\tau}\,\mathrm{Re}(s)),
\end{aligned}$$

where the factor $(1 - e^{\mathrm{Re}(s)T}\,\mathrm{Re}(s))$ on the RHS tends to 1 as $\tau \to \infty$ since $\mathrm{Re}(s) < 0$. We thus obtain the bound $1/r + o(1)$ when we replace γ_r^+ with γ_r^- in (3), and the RHS of (2) is bounded by $2/r + o(1)$ as $\tau \to \infty$. It follows that for any $\epsilon > 0$, for $r = 3/\epsilon > 0$ we have

$$|g(0) - g_\tau(0)| < 3/r = \epsilon$$

for all sufficiently large τ. Therefore $\lim_{\tau\to\infty} g_\tau(0) = g(0)$ as desired. □

Remark 16.14. Theorem 16.13 is an example of what is known as a *Tauberian theorem.* The Laplace transform

$$(\mathcal{L}f)(s) := \int_0^\infty e^{-st}f(t)dt,$$

is in general not defined on $\operatorname{Re}(s) \le c$, where c is the least c for which $f(t) = O(e^{ct})$. It may happen that the function $\mathcal{L}f$ has an analytic continuation to a larger domain; for example, if $f(t) = e^t$ then $(\mathcal{L}f)(s) = \frac{1}{s-1}$ extends to a holomorphic function on $\mathbb{C} - \{1\}$. But plugging values of s with $\operatorname{Re}(s) \le c$ into the integral usually does not work; in our $f(t) = e^t$ example, the integral diverges on $\operatorname{Re}(s) \le 1$. The theorem says that when $\mathcal{L}f$ extends to a holomorphic function on the entire half-plane $\operatorname{Re}(s) \ge 0$, its value at $s = 0$ is exactly what would get by naively plugging 0 into the integral defining $\mathcal{L}f$.

More generally, Tauberian theorems refer to results related to transforms $f \to \mathcal{T}(f)$ that allow us to deduce properties of f (such as the convergence of $\int_0^\infty f(t)dt$) from properties of $\mathcal{T}(f)$ (such as analytic continuation to $\operatorname{Re}(s) \ge 0$). The term "Tauberian" was coined by Hardy and Littlewood and refers to Alfred Tauber, who proved a theorem of this type as a partial converse to a theorem of Abel.

Theorem 16.15 (Prime Number Theorem). $\pi(x) \sim \frac{x}{\log x}$.

Proof. $H(t) = \vartheta(e^t)e^{-t} - 1$ is piecewise continuous and bounded, by Lemma 16.7, and its Laplace transform extends to a holomorphic function on $\operatorname{Re}(s) \ge 0$, by Corollary 16.12. Theorem 16.13 then implies that the integral

$$\int_0^\infty H(t)dt = \int_0^\infty \big(\vartheta(e^t)e^{-t} - 1\big)dt$$

converges. Replacing t with $\log x$, we see that

$$\int_1^\infty \left(\vartheta(x)\frac{1}{x} - 1\right)\frac{dx}{x} = \int_1^\infty \frac{\vartheta(x) - x}{x^2}dx$$

converges. Lemma 16.8 implies $\vartheta(x) \sim x$, equivalently, $\pi(x) \sim \frac{x}{\log x}$, by Theorem 16.6. □

One disadvantage of our proof is that it does not give us an error term. Using more sophisticated methods, Korobov [6] and Vinogradov [14] independently obtained the bound

$$\pi(x) = \operatorname{Li}(x) + O\left(\frac{x}{\exp\left((\log x)^{3/5+o(1)}\right)}\right),$$

in which we note that the error term is bounded by $O(x/(\log x)^n)$ for all n but not by $O(x^{1-\epsilon})$ for any $\epsilon > 0$. Assuming the Riemann Hypothesis, which states the zeros of $\zeta(s)$ in the critical strip $0 < \operatorname{Re}(s) < 1$ all lie on the line $\operatorname{Re}(s) = \frac{1}{2}$, one can prove

$$\pi(x) = \operatorname{Li}(x) + O(x^{1/2+o(1)}).$$

There thus remains a large gap between what we can prove about the distribution of prime numbers and what we believe to be true. Remarkably, other than refinements to the $o(1)$ term appearing in the Korobov-Vinogradov bound, essentially no progress has been made in this direction in the past 50 years.

16.3 A quick recap of some basic complex analysis

The complex numbers $\mathbb{C}$ are a topological field under the distance metric $d(x, y) = |x - y|$ induced by the standard absolute value $|z| := \sqrt{z\bar{z}}$, which is also a norm on $\mathbb{C}$ as an $\mathbb{R}$-vector space; all references to the topology on $\mathbb{C}$ (open, compact, convergence, limits, etc...) are made with this understanding.

16.3.1 Glossary of terms and basic theorems

Let f and g denote complex functions defined on an open subset of $\mathbb{C}$.

- f is *differentiable* at z_0 if $\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ exists.
- f is *holomorphic* at z_0 if it is differentiable on an open neighborhood of z_0.
- f is *analytic* at z_0 if there is an open neighborhood of z_0 in which f can be defined by a power series $f(z) = \sum_{n=0} a_n(z-z_0)^n$; equivalently, f is infinitely differentiable and has a convergent Taylor series on an open neighborhood of z_0.
- **Theorem**: f is holomorphic at z_0 if and only if it is analytic at z_0.
- **Theorem**: If C is a connected set containing a nonempty open set U and f and g are holomorphic on C with $f_{|U} = g_{|U}$, then $f_{|C} = g_{|C}$.
- With U and C as above, if f is holomorphic on U and g is holomorphic on C with $f_{|U} = g_{|U}$, then g is the (unique) *analytic continuation* of f to C and f *extends* to g.
- If f is holomorphic on a punctured open neighborhood of z_0 and $|f(z)| \to \infty$ as $z \to z_0$ then z_0 is a *pole* of f; note that the set of poles of f is necessarily a discrete set.
- f is *meromorphic* at z_0 if it is holomorphic at z_0 or has z_0 as a pole.
- **Theorem**: If f is meromorphic at z_0 then it can be defined by a Laurent series $f(z) = \sum_{n\ge n_0} a_n(z-z_0)^n$ that converges on an open punctured neighborhood of z_0.
- The *order of vanishing* $\operatorname{ord}_{z_0}(f)$ of a nonzero function f that is meromorphic at z_0 is the least index n of any nonzero coefficient a_n in its Laurent series expansion at z_0. Then z_0 is a pole of f iff $\operatorname{ord}_{z_0}(f) < 0$ and z_0 is a zero of f iff $\operatorname{ord}_{z_0}(f) > 0$.
- If $\operatorname{ord}_{z_0}(f) = 1$ then z_0 is a *simple zero* of f, and if $\operatorname{ord}_{z_0}(f) = -1$ it is a *simple pole*.
- The *residue* $\operatorname{res}_{z_0}(f)$ of a function f meromorphic at z_0 is the coefficient a_{-1} in its Laurent series expansion $f(z) = \sum_{n\ge n_0} a_n(z-z_0)^n$ at z_0.
- **Theorem**: If z_0 is a simple pole of f then $\operatorname{res}_{z_0}(f) = \lim_{z\to z_0}(z-z_0)f(z)$.
- **Theorem**: If f is meromorphic on a set S then so is its *logarithmic derivative* f/f', and f/f' has only simple poles in S and $\operatorname{res}_{z_0}(f/f') = \operatorname{ord}_{z_0}(f)$ for all $z_0 \in V$. In particular the poles of f/f' are precisely the zeros and poles of f.

16.3.2 Convergence

Recall that a series $\sum_{n=1}^\infty a_n$ of complex numbers *converges absolutely* if the series $\sum_n |a_n|$ of nonnegative real numbers converges. An equivalent definition is that the function $a(n) := a_n$ is integrable with respect to the counting measure μ on the set of positive integers $\mathbb{N}$. Indeed, if the series is absolutely convergent then

$$\sum_{n=1}^{\infty} a_n = \int_{\mathbb{N}} a(n)\mu,$$

and if the series is not absolutely convergent, the integral is not defined. Absolute convergence is effectively built-in to the definition of the Lebesgue integral, which requires that in order for the function $a(n) = x(n) + iy(n)$ to be integrable, the positive real functions $|x(n)|$ and $|y(n)|$ must both be integrable (summable), and separately computes sums of the positive and negative subsequences of $(x(n))$ and $(y(n))$ as suprema over finite subsets.

The measure-theoretic perspective has some distinct advantages. It makes it immediately clear that we may replace the index set $\mathbb{N}$ with any set of the same cardinality, since the counting measure depends only on the cardinality of $\mathbb{N}$, not its ordering. We are thus free to sum over any countable index set, including $\mathbb{Z}$, $\mathbb{Q}$, any finite product of countable sets, and any countable coproduct of countable sets (such as countable direct sums of $\mathbb{Z}$); such sums are ubiquitous in number theory and many cannot be meaningfully interpreted as limits of partial sums in the usual sense, since this assumes that the index set is well ordered (not the case with $\mathbb{Q}$, for example). The measure-theoretic view makes also makes it clear that we may convert any absolutely convergent sum of the form $\sum_{X\times Y}$ into an iterated sum $\sum_X\sum_Y$ (or vice versa), via Fubini's theorem.

We say that an infinite product $\prod_n a_n$ of nonzero complex numbers is *absolutely convergent* when the sum $\sum_n \log a_n$ is, in which case $\prod_n a_n := \exp(\sum_n \log a_n)$.[5] This implies that an absolutely convergent product cannot converge to zero, and the sequence (a_n) must converge to 1 (no matter how we order the a_n). All of our remarks above about absolutely convergent series apply to absolutely convergent products as well.

A series or product of complex functions $\{f_n(z)\}$ is *absolutely convergent on S* if the series or product of complex numbers $\{f_n(z_0)\}$ is absolutely convergent for all $z_0 \in S$.

Definition 16.16. A sequence of complex functions (f_n) *converges uniformly on S* if there is a function f such that for every $\epsilon > 0$ there is an integer N for which $\sup_{z\in S}|f_n(z)-f(z)| < \epsilon$ for all $n \ge N$. The sequence (f_n) *converges locally uniformly on S* if every $z_0 \in S$ has an open neighborhood U for which (f_n) converges uniformly on $U \cap S$. When applied to a series of function these terms refer to the sequence of partial sums.

Because $\mathbb{C}$ is locally compact, locally uniform convergence is the same thing as compact convergence: a sequence of functions converges locally uniformly on S if and only if it converges uniformly on every compact subset of S.

Theorem 16.17. *A sequence or series of holomorphic functions f_n that converges locally uniformly on an open set U converges to a holomorphic function f on U, and the sequence or series of derivatives f_n' then converges locally uniformly to f' (and if none of the f_n has a zero in U and $f \ne 0$, then f has no zeros in U).*

Proof. See [3, Thm. III.1.3] and [3, Thm. III.7.2] □

Definition 16.18. A series of complex functions $\sum_n f_n(z)$ converges *normally* on a set S if $\sum_n \|f_n\| := \sum_n \sup_{z\in S}|f_n(z)|$ converges. The series $\sum_n f_n(z)$ converges *locally normally* on S if every $z_0 \in S$ has an open neighborhood U on which $\sum_n f_n(z)$ converges normally.

Theorem 16.19 (Weierstrass M-test)**.** *Every locally normally convergent series of functions converges absolutely and locally uniformly. Moreover, a series $\sum_n f_n$ of holomorphic functions on S that converges locally normally converges to a holomorphic function f on S, and then $\sum_n f_n'$ locally normally to f'.*

Proof. See [3, Thm. III.1.6]. □

Remark 16.20. To show a series $\sum_n f_n$ is locally normally convergent on a set S amounts to proving that for every $z_0 \in S$ there is an open neighborhood U of z_0 and a sequence of real numbers (M_n) such that $|f_n(z)| \le M_n$ for $z \in U \cap S$ and $\sum_n M_n < \infty$, whence the term "M-test".

[5] In this definition we use the principal branch of $\log z := \log|z| + i \operatorname{Arg} z$ with $\operatorname{Arg} z \in (-\pi, \pi)$.

16.3.3 Contour integration

We shall restrict our attention to integrals along contours defined by piecewise-smooth parameterized curves; this covers all the cases we shall need.

Definition 16.21. A *parameterized curve* is a continuous function $\gamma\colon [a,b] \to \mathbb{C}$ whose domain is a compact interval $[a,b] \subseteq \mathbb{R}$. We say that γ is *smooth* if it has a continuous nonzero derivative on $[a,b]$, and *piecewise-smooth* if $[a,b]$ can be partitioned into finitely many subintervals on which the restriction of γ is smooth. We say that γ is *closed* if $\gamma(a) = \gamma(b)$, and *simple* if it is injective on $[a,b)$ and $(a,b]$. Henceforth we will use the term *curve* to refer to any piecewise-smooth parameterized curve γ, or to its oriented image of in the complex plane (directed from $\gamma(a)$ to $\gamma(b)$), which we may also denote γ.

Definition 16.22. Let $f\colon \Omega \to \mathbb{C}$ be a continuous function and let γ be a curve in Ω. We define the *contour integral*

$$\int_\gamma f(z)dz := \int_a^b f(\gamma(t))\gamma'(t)dt,$$

whenever the integral on the RHS (which is defined as a Riemann sum in the usual way) converges. Whether $\int_\gamma f(z)dz$ converges, and if so, to what value, does not depend on the parameterization of γ: if γ' is another parameterized curve with the same (oriented) image as γ, then $\int_{\gamma'} f(z)dz = \int_\gamma f(z)dz$.

We have the following analog of the fundamental theorem of calculus.

Theorem 16.23. *Let $\gamma\colon [a,b] \to \mathbb{C}$ be a curve in an open set Ω and let $f\colon \Omega \to \mathbb{C}$ be a holomorphic function Then*

$$\int_\gamma f'(z)dz = f(\gamma(b)) - f(\gamma(a)).$$

Proof. See [2, Prop. 4.12]. □

Recall that the Jordan curve theorem implies that every simple closed curve γ partitions $\mathbb{C}$ into two components, one of which we may unambiguously designate as the *interior* (the one on the left of our *positively oriented* curves). We say that γ is *contained* in an open set U if both γ and its interior lie in U. The interior of γ is a simply connected set, and if an open set U contains γ then it contains a simply connected open set that contains γ.

Theorem 16.24 (Cauchy's Theorem)**.** *Let U be an open set containing a simple closed curve γ. For any function f that is holomorphic on U we have*

$$\int_\gamma f(z)dz = 0.$$

Proof. See [2, Thm. 8.6] (we can restrict U to a simply connected set). □

Cauchy's theorem generalizes to meromorphic functions.

Theorem 16.25 (Cauchy Residue Formula)**.** *Let U be an open set containing a simple closed curve γ. Let f be a function that is meromorphic on U, let $z_1, \ldots, z_n$ be the poles of f that lie in the interior of γ, and suppose that no pole of f lies on γ. Then*

$$\int_\gamma f(z)dz = 2\pi i \sum_{i=1}^n \operatorname{res}_{z_i}(f).$$

Proof. See [2, Thm. 10.5] (we can restrict U to a simply connected set). □

To see where the $2\pi i$ comes from, consider $\int_\gamma \frac{dz}{z}$ with $\gamma(t) = e^{it}$ for $t \in [0, 2\pi]$. In general one weights residues by a corresponding winding number, but the winding number of a simple closed curve about a point in its interior is always 1.

Theorem 16.26 (Cauchy's Integral Formula)**.** *Let U be an open set containing a simple closed curve γ. For any function f holomorphic on U and a in the interior of γ,*

$$f(a) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-a}dz.$$

Proof. Apply Cauchy's residue formula to $g(z) = f(z)/(z-a)$; the only poles of g in the interior of γ are a simple pole at $z = a$ with $\operatorname{res}_a(g) = f(a)$. □

Cauchy's residue formula can also be used to recover the coefficients $f^{(n)}(a)/n!$ appearing in the Laurent series expansion of a meromorphic function at a (apply it to $f(z)/(z-a)^{n+1}$). One of many useful consequences of this is Liouville's theorem, which can be proved by showing that the Laurent series expansion of a bounded holomorphic function on $\mathbb{C}$ about any point has only one nonzero coefficient (the constant coefficient).

Theorem 16.27 (Liouville's theorem)**.** *Bounded entire functions are constant.*

Proof. See [2, Thm. 5.10]. □

We also have the following converse of Cauchy's theorem.

Theorem 16.28 (Morera's Theorem)**.** *Let f be a continuous function and on an open set U, and suppose that for every simple closed curve γ contained in U we have*

$$\int_\gamma f(z)dz = 0.$$

Then f is holomorphic on U.

Proof. See [3, Thm. II.3.5]. □

References

[1] Lars V. Ahlfors, *Complex analysis: an introduction to the theory of analytic functions of one complex variable*, 3rd edition, McGraw-Hill, 1979.

[2] Joseph Bak and Donald J. Newman, *Complex analysis*, Springer, 2010.

[3] Rolf Busam and Eberhard Freitag, *Complex analysis*, 2nd edition, Springer 2009.

[4] Paul Erdös, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Scis. U.S.A. **35** (1949), 373–384.

[5] Jacques Hadamard, *Sur la distribution des zéros de la function $\zeta(s)$ et ses conséquences arithmétique*, Bull. Soc. Math. France **24** (1896), 199–220.

[6] Nikolai M. Korobov, *Estimates for trigonometric sums and their applications*, Uspechi Mat. Nauk **13** (1958), 185–192.

[7] Serge Lange, *Complex analysis*, 4th edition, Springer, 1985.

[8] David J. Newman, *Simple analytic proof of the Prime Number Theorem*, Amer. Math. Monthly **87** (1980), 693–696.

[9] Charles Jean de la Vallée Poussin, *Reserches analytiques sur la théorie des nombres premiers*, Ann. Soc. Sci. Bruxelles **20** (1896), 183–256.

[10] Bernhard Riemann, *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsberichte der Berliner Akademie, 1859.

[11] Alte Selberg, *An elementary proof of the Prime-Number Theorem*, Ann. Math. **50** (1949), 305–313.

[12] Elias M. Stein and Rami Shakarchi, *Complex analysis*, Princeton University Press, 2003.

[13] Alfred Tauber, *Ein Satz aus der Theorie der unendlichen Reihen*, Monatsh f. Mathematik und Physik **8** (1897), 273–277.

[14] Ivan M. Vinogradov, *A new estimate of the function* $\zeta(1+it)$, Izv. Akad. Nauk SSSR. Ser. Mat. **22** (1958), 161–164.

[15] Don Zagier, *Newman's short proof of the Prime Number Theorem*, Amer. Math. Monthly **104** (1997), 705–708.